Implementation and Partial Provability[1]

Elchanan Ben-Porath[2] Barton L. Lipman[3]

First Draft
September 2008

Current Draft
April 2009

[1]We thank Eddie Dekel, Jacob Glazer, Sean Horan, Phil Reny, and various seminar audiences for helpful comments and the US–Israel Binational Science Foundation for supporting this research. Lipman also thanks the National Science Foundation for support.

[2]Department of Economics and Center for Rationality, Hebrew University. Email: benporat@math.huji.ac.il.

[3]Department of Economics, Boston University. Email: blipman@bu.edu.

Abstract

We extend implementation theory by allowing the social choice function to depend on more than just the profile of preferences of the agents and by allowing agents to support their statements with hard evidence. We show that a simple condition on the evidence structure which is necessary for the implementation of a social choice function f when the preferences of the agents are state independent is also sufficient for implementation for any preferences (including state dependent) if the social planner can perform small monetary transfers and there are at least three players. If transfers can be large, f can be implemented in a game with perfect information when there are at least two players under an additional boundedness assumption. In both cases, transfers only occur off the equilibrium path. Finally, in the special but important case of allocation problems, under weak conditions, f can be implemented in a perfect information game with at least two players and *no* transfers. In all cases, the use of evidence enables implementation which is robust in the sense that the social planner needs very little information about the preferences, beliefs, and evidence of the agents and the agents need little information about each others' preferences.

1 Introduction

This paper addresses a pair of related limitations present in the theory of implementation and mechanism design since the seminal work of Hurwicz (1972) and Maskin (1977, 1999). The first limitation is the assumption that the social alternative to be implemented depends *only* on the profile of preferences of the agents. As we argue below, this excludes many important situations of economic interest. The second limitation is the assumption that the agents cannot provide hard evidence. Obviously, hard evidence can establish the veracity of statements which would not be incentive compatible. Furthermore, use of hard evidence is very common in real world institutions, so the omission of this possibility from mechanism design is clearly an important gap to fill. As will become clear shortly these two issues are related in that it is impossible to address the first limitation without introducing evidence. In addition, we will show that implementation with evidence has robustness properties which are not possible in standard models.

The general goal of our research is to extend the standard theory by considering a more general model that is not restricted by these two assumptions. In this paper we study complete information implementation, primarily focusing on subgame perfect equilibrium. By implementation, we refer to what is sometimes called full implementation — that is, the requirement that *every* equilibrium in the game that is induced by the mechanism leads to the social alternative specified by the social choice function.

Let $\mathcal{I} = \{1, \ldots, I\}$ be a set of agents, A a set of social alternatives, and S a set of states. A state $s \in S$ specifies all the parameters that are relevant for the determination of the alternative $f(s)$ that a social planner (henceforth, SP) would like to implement. The function $f : S \to A$ is called the social choice function (SCF). We assume that SP does not know the true state but that every agent $i \in \mathcal{I}$ does know it. (This is the assumption of complete information.) As in the usual model, each agent $i \in \mathcal{I}$ in each state $s \in S$ has a preference $\succeq_{i,s}$ over the set of alternatives. (This is phrased differently in the usual model, but this difference is one of notation, not substance.) The more significant change is that we also assume each agent i has a proof technology, $M_i = \{M_i(s)\}_{s \in S}$, where $M_i(s)$ is the set of events (subsets of S) that player i can prove in state s. So $E \in M_i(s)$ if in state s, player i has some evidence he can present which proves that the true state lies in the event E.

To see that the standard model is a special case of this model, note that the assumption that players cannot prove anything corresponds to $M_i(s) = S$ for every $i \in \mathcal{I}$ and $s \in S$. Similarly, the assumption that the social choice depends only on the preference profile translates to the requirement that if $\succ_{i,s} = \succ_{i,s'}$ for every $i \in \mathcal{I}$, then $f(s) = f(s')$. In our model, this condition will often be violated, implying that none of the standard implementation results (*e.g.*, Maskin (1977), Moore and Repullo (1988), Palfrey and Srivastava (1991), or Abreu and Matsushima (1992)) apply. To see this, simply note that if the preferences of the agents do not vary across states and if agents cannot present evidence (as the usual model assumes), then the game induced by any mechanism in the usual model will not vary across states, so the set of equilibria cannot vary across states. Hence the social choice cannot depend on more than the preference profile if there is no evidence available.

Our goal is to study the conditions on the proof technology, the preferences of the players, and the social choice function which make implementation of the social choice function possible. We identify a simple condition on the relationship between the proof technologies of the agents and the social choice function, f, which we call *measurability*, which is necessary for implementation of f when the preferences are state independent. We then show that measurability is also sufficient for implementation of f for *every* preference structure if SP can perform monetary transfers among the players. More specifically, Theorem 1 establishes that implementation is possible in a mechanism with perfect information if there are at least two players under a boundedness assumption. Theorem 2 shows that if there are at least three players, then implementation is possible with arbitrarily small monetary transfers in a mechanism which involves an integer game. This mechanism has only one stage and hence it implements in Nash equilibrium as well. For both theorems, transfers only occur off the equilibrium path. We also show in Theorem 3 that for the problem of allocating a set of goods among a set of agents, we can achieve implementation in a perfect information mechanism without monetary transfers under weak conditions.

We emphasize that the measurability condition puts no restrictions on how the preferences of the agents relate to the evidence available and the social alternative to be implemented. Thus, if measurability is satisfied, SP can implement without knowing anything about the preference structure and its relation to the proof technologies and social choice function. We also show that the agents do not need more than minimal information about the preferences of other agents to play equilibrium strategies. As a result, the planner does not need to know whether the agents know each other's preferences or what they believe about the preferences of others.

To further motivate the issues, we present three examples. In the first two examples, the preferences of the players are independent of the state. In such cases, the standard theory of implementation is irrelevant: if players cannot present any evidence, then for any mechanism the set of equilibrium outcomes is independent of the state. The third example illustrates a situation where SP has limited information about the preferences of the players. Clearly, in such a situation SP would like to use a mechanism which will implement for every profile of preferences of the players.

1. Consider a personal injury trial where the problem of SP is to determine the level of compensation that the defendant should pay the plaintiff. Suppose that SP (the judge)

wants to set the compensation equal to the damage that has been caused to the plaintiff but does not know what that damage is. Thus, the set of alternatives is the set of possible compensations and the state specifies the level of damages. Clearly, the preferences of the two players over the set of alternatives are independent of the state — the plaintiff always prefers a higher compensation and the defendant a lower one.

2. Consider the problem of an organization allocating a fixed budget among a set of individuals or departments. Here again it is reasonable to assume that each player wants to get as much as possible regardless of the state. On the other hand, SP wants to implement an allocation which depends on which department can most efficiently use the organization's resources to further its goals, an objective which depends on the state of the world.

3. Consider a situation where the set of players is a public commission that has to decide how to allocate money between different projects. Each member of the commission may have a personal preference among the projects. For instance, some may be close to his home or may benefit family members or friends. Clearly, personal benefits that members stand to gain from the projects are irrelevant for the selection rule which SP would like to implement. Thus, from the point of view of SP, the members of the committee are experts who may have personal biases and from whom information should be extracted. Theorems 1 and 2 show that if the social choice function satisfies measurability, then SP can implement with little or no information about the preferences of the players.

The literature on communication with verifiable information is quite extensive, so our account must be brief and very partial. Starting with Grossman (1981), Grossman and Hart (1980), and Milgrom (1981), there is a growing literature which examines persuasion games. In these games, a set of agents (often only one agent) make verifiable statements to a "receiver" who takes an action. See, for example, Fishman and Hagerty (1990), Glazer and Rubinstein (2001, 2004, 2006), Lipman and Seppi (1995), Milgrom and Roberts (1986), Seidman and Winter (1997), Sher (2008), and Shin (1994). See also Forges and Koessler (forthcoming) and Okuno-Fujiwara, Postlewaite, and Suzumura (1990) for other interesting models of games where players make verifiable statements.

There has been some work that examines communication of verifiable information in the context of mechanism design (Alger and Ma (2003), Bull and Watson (2004, 2007), Deneckere and Severinov (2008), Green and Laffont (1986)) and also more recently communication of verifiable information with a mediator (Forges and Koessler (2005)). These papers analyze environments with incomplete information and primarily focus on the question of identifying an appropriate form of the Revelation Principle — that is, when some form of a direct revelation mechanism suffices for achieving an equilibrium with a particular outcome. By contrast, we focus on (full) implementation, where we seek games for which *every* equilibrium has the desired outcome.

There are a few papers that discuss implementation with evidence in specific contexts. Hurwicz, Maskin, and Postlewaite (1995) and Postlewaite and Wettstein (1989) consider the implementation of the Walrasian social choice function in a model where a player may misrepresent his endowment by destroying or hiding some of it. In this setting, showing (part of) one's endowment proves that the player has at least that much. Bull and Watson (2004) study Nash implementation of a "zero–sum" social choice function that specifies monetary transfers between a set of agents.

Perhaps the closest predecessor to the current research is Lipman and Seppi (1995). They consider a game of persuasion and obtain necessary and sufficient conditions for robust inference of the true state. While their results are phrased in terms of equilibrium outcomes of a fixed game, many of the results can be directly translated into statements of implementation of a social choice function. However, their analysis is limited by several assumptions. First, they assume that all the players have the same proof technology. Second, they assume that players have conflicting preferences. Finally, for the most part, they restrict attention to mechanisms where each player sends a message only once and players move in a sequence. The results which we present in Section 3 cannot be obtained by such mechanisms.

We discuss another related paper, Kartik and Tercieux (2009), in Section 3.

In Section 2, we define the model. The implementation results are presented in Section 3. We discuss the sense in which the implementation is robust in Section 4. In Section 5, we provide a series of examples demonstrating that our results are tight. Section 6 concludes. Proofs are contained in the Appendix.

2 The Model

A social environment Ψ is a tuple $\langle \mathcal{I}, A, S, (\succeq_{i,s}, M_i(s))_{i\in I, s\in S}\rangle$. $\mathcal{I} = \{1, .., I\}$ is the set of agents and A is a set of social alternatives. S is the set of states of the world. $\succeq_{i,s}$ is the preference relation of agent i over A at state s.

The more novel part of the model is $M_i(s)$. This is the set of "hard evidence" messages i has available in state s. When agent i presents a message $m_i \in \cup_{s\in S} M_i(s)$, he presents proof of the event $E(m_i) = \{s \,| m_i \in M_i(s)\}$. That is, he proves the event that consists of the states in which m_i can be presented. It is convenient to identify the message or hard evidence m_i with the event $E(m_i)$ and thus we define $M_i(s)$ as the set of events that i can prove at state s.

This definition implies two important properties of $M_i(s)$. First, for all $E \in M_i(s)$, we must have $s \in E$. That is, agent i cannot prove something that is false. Second, $E \in M_i(s)$ implies $E \in M_i(s')$ for every $s' \in E$. In words, if the agent can prove event E in some state, he must be able to prove it in every state in E. In subsequent discussions, we refer to this property as *consistency*.

In addition, we assume that there are no restrictions on the amount of his evidence any agent can present. Formally, for every i and every s, we assume that

$$\bigcap_{E \in M_i(s)} E \in M_i(s).$$

That is, it is always feasible for i to prove the event that corresponds to presenting all the evidence in his possession. Thus this assumption implies that there are no effective constraints on the time and effort that are required to establish a claim.[1]

It would be very natural to also assume that $S \in M_i(s)$ for all s. That is, any player i can always present a trivial message that proves nothing. However, such an assumption is not needed for our results.

Let $M_i = \cup_{s \in S} M_i(s)$. We say that a message $m_i \in M_i$ *refutes* state $s' \in S$ if $m_i \notin M_i(s')$. That is, m_i could not have been sent at state s'. Equivalently, m_i refutes s' if $s' \notin m_i$. Note that player i can refute state s' at state s iff $M_i(s) \not\subseteq M_i(s')$.

For a simple example of what this framework allows, suppose there is one agent and two states of the world, p (where the agent can play the piano) and np (where he cannot). Then a natural specification of the evidence available to player 1 is $M_1(p) = \{S, \{p\}\}$ and $M_1(np) = \{S\}$. To see this, note that if player 1 cannot play the piano, all he can do is make random noises with it. However, this does not prove he cannot play the piano, since he could make the same noises if he does know how to play. Hence in state np, the only event player 1 can prove is the trivial event S. On the other hand, in state p, the agent could provide only the trivial proof of S by making random noises or could play a piece which demonstrates his ability, hence proving $\{p\}$.[2] The example represents a general phenomenon — in many situations, proving a negative proposition ("agent i cannot do x" or "agent i does not have y") is difficult or impossible.[3]

[1]Lipman and Seppi (1995) call this assumption the full reports condition, while Bull and Watson (2007) call it normality. See Lipman and Seppi (1995), Glazer and Rubinstein (2001, 2004, 2006), Bull and Watson (2007), and Kartik and Tercieux (2009) for models which relax this assumption. We thank Kartik and Tercieux for pointing out that a previous draft of this paper stated an unnecessarily strong version of this property.

[2]The piano player example appears in Lipman and Seppi (1995) who attribute it to Mike Peters.

[3]For example, Hurwicz, Maskin, and Postlewaite (1995) and Postlewaite and Wettstein (1989) assume that a player who has x units of a good can prove that he has *at least* y units for every $y \leq x$ by simply showing y units. However, the player cannot prove that he has *exactly* x units because he might be hiding some units. See Okuno-Fujiwara, Postlewaite, and Suzumura (1990) for a similar model. See also

We refer to the set $\{\succeq_{i,s} | i \in I, s \in S\}$ as the *preference structure*. We say that a preference structure has *state independent preferences* if the preference of each player over A is independent of s. That is, for every $i \in \mathcal{I}$ and $s, s' \in S$, we have $\succeq_{i,s} = \succeq_{i,s'}$.

A *social choice function* (SCF) for Ψ is a function $f : S \to A$. The social choice function represents the outcome a social planner (SP) wishes to achieve as a function of the state. The social planner does not know the true state but every player in $\mathcal{I}$ does know it, as in the standard complete information implementation problem.

We think of a state $s \in S$ as a specification of all facts about the world which are relevant for the implementation problem. In particular, a state specifies every fact that is relevant for the determination of the alternative that SP wishes to implement, whether these facts affect the preferences of the agents or not. As in the examples in the introduction, SP's preferred outcome may vary across states even when the preferences of the agents do not. In addition, as is clear from the definition, a state specifies the preference of each player over the social alternatives and the evidence that he can present.

There is one conceptual question we must address before defining mechanisms in this environment, namely, whether the mechanism can forbid or require an agent to present certain evidence. In the usual definitions, SP is completely free to determine the set of messages an agent may use in any given situation. However, when the messages include hard evidence, can the mechanism forbid use of some messages? Alternatively, can it require use of some messages (when the agent has them available)?

The issue of whether SP can forbid some hard evidence messages or not is irrelevant to the results. Since we assume that SP is committed to the mechanism, he can commit to simply ignoring certain evidence and not allowing other players to observe it. Hence he has no need to forbid any particular messages.[4]

The issue of compelling the agent to provide evidence is quite different. One can imagine situations where the social planner knows that the evidence an agent has is kept in a specific location (say, documents in a bank safe) and has legal authority to seize this evidence. In a scenario like this, the social planner can force the agent to reveal the evidence he has. Clearly, in such a case, there are no interesting incentive questions

Shin (1994) for a model where an agent can prove that he has evidence by providing it, but can never prove that he does not have evidence.

[4]To be more precise, this argument implies that there is no need to forbid *informative* evidence since it can always be treated as if no evidence were presented. Kartik and Tercieux (2009) give an example where there is a value to forbidding agents from presenting no information at all. Thus, forbidding *uninformative* evidence can be useful. However, this possibility does not affect our results since our necessary condition is necessary whether or not this is allowed and our sufficiency results do not need such mechanisms. On the other hand, forbidding uninformative evidence is the same as compelling informative evidence, an assumption we argue against below.

surrounding such evidence — the social planner can always learn it whether the agent has an incentive to reveal it or not. In reality, there are many situations where the planner does not know whether the agent has evidence and, if so, where it could be found. In such a case, even the most draconian legal requirements do not *force* presentation of evidence, but merely provide very strong incentives for agents to *choose* to provide it. To see the point concretely, consider the piano player example above. Can the mechanism say that in state p, player 1 has no option other than to present the evidence $\{p\}$? We assume that this is not possible. The mechanism can give incentives for the agent to provide this evidence by, for example, punishing the agent severely if he does not provide it. However, this is quite different from constructing a mechanism in which in state p, it is not feasible for player 1 to avoid providing this proof. For one thing, punishing the agent for not providing evidence implies that he is punished in state np as well.

In short, we assume that whenever an agent has an opportunity to present evidence in a mechanism, the mechanism must allow him to present as much or as little of his available evidence as he chooses. Of course, the way outcomes depend on his choices is unconstrained.

More specifically, a mechanism is a multistage game form which specifies at each history h of the game a subset of players who act simultaneously at h after observing h. An action for a player i is a pair that consists of a cheap talk message c_i (that is, c_i can be played in every state s) and a hard evidence message m_i. The assumption that player i can present any evidence he has means that at state s, player i can play any message m_i in the set $M_i(s_i)$. In addition, the mechanism defines an outcome function which specifies the social alternative $a \in A$ that SP chooses as a function of the terminal history.

For notational simplicity, our definition of a mechanism imposes more restrictions than this. In particular, we consider only multistage games where the set of "nonevidence" messages available to a player is constant across his information sets. However, we emphasize that this is only to avoid excessive notation. As will be obvious from the proof, our necessity result would hold for any class of games and any equilibrium notion. Since we can prove sufficiency by constructing any game, the restriction to multistage games is without loss of generality for sufficiency results.

Formally,[5] a *mechanism* consists of the following objects. For each player i, there is a set C_i, interpreted as the "cheap talk" messages available for i. We also have a set of histories, H, where a history $h \in H$ is a finite sequence $(h_1, \ldots, h_k)$ for some k. We assume that the empty sequence, denoted $\emptyset$, is an element of H and that if $(h_1, \ldots, h_k, h_{k+1}) \in H$, then $(h_1, \ldots, h_k) \in H$. A history $h = (h_1, \ldots, h_k) \in H$ is *terminal* if there is no h_{k+1} such that $(h_1, \ldots, h_k, h_{k+1}) \in H$. Let H_T denote the set of terminal histories. We also

[5]Our definition is a slight variation on that of Osborne and Rubinstein (1994), Section 6.3.2.

have a function $P : H \setminus H_T \to 2^{\mathcal{I}} \setminus \{\emptyset\}$ where $P(h)$ is the set of players who choose actions at history h where the interpretation is that these players move simultaneously after having observed history h. We require that if $(h_1, \ldots, h_k, h_{k+1}) \in H$, then $h_{k+1} \in \prod_{i \in P(h_1,\ldots,h_k)} C_i \times M_i$. (Recall that $M_i = \cup_{s \in S} M_i(s)$). That is, each component of a history is a set of cheap talk messages and evidence statements by some subset of the players. In line with this, we write $h_{k+1} = \{(c_i, m_i)\}_{i \in P(h_1,\ldots,h_k)}$

Since, at state s, player i can only provide evidence in the set $M_i(s)$, the set of actions available to each player i at each history $h \in H$, $i \in P(h)$, is restricted to the set $C_i \times M_i(s)$. Say that a history $(h_1, \ldots, h_k)$ is *feasible* in state s if for $n = 1, \ldots, k$, letting $h_n = \{(c_{in}, m_{in})\}_{i \in P(h_1,\ldots,h_{n-1})}$, we have $m_{in} \in M_i(s)$ for all $i \in P(h_1, \ldots, h_{n-1})$. Formally, our assumption that players cannot be compelled or forbidden to present evidence takes the following form: If $(h_1, \ldots, h_k, h_{k+1}) \in H$ is feasible in state s and $h'_{k+1} \in \prod_{i \in P(h_1,\ldots,h_k)} C_i \times M_i(s)$, then $(h_1, \ldots, h_k, h'_{k+1}) \in H$.

Finally, we have a function $g : H_T \to A$ specifying the outcome selected by the mechanism for each terminal node. A mechanism Γ is the tuple $\langle H, P, g, (C_i)_{i \in \mathcal{I}} \rangle$.

The mechanism Γ is a *one–stage mechanism* if every nonempty history has length one. The mechanism has *perfect information* if for every nonterminal $h \in H$, $P(h)$ is a singleton.

For every $s \in S$, a mechanism Γ induces a game, denoted by $\Gamma(s)$, where the set of histories is the set of histories that are feasible in state s and the preference of player i over the set of terminal histories is the preference that is induced by $\succeq_{i,s}$.

We say that a mechanism Γ *implements* a social choice function f if for every $s \in S$ and every profile of pure strategies $\sigma^s = (\sigma_1^s, ..., \sigma_I^s)$ that is a subgame perfect equilibrium in the game $\Gamma(s)$, we have $g(\sigma^s) = f(s)$ (where $g(\sigma^s)$ is the alternative that is selected by g at the terminal history that is reached under σ^s). Obviously, if Γ is a one–stage mechanism, then the set of subgame perfect equilibria of the induced game is the same as the set of Nash equilibria.

The definitions for our robustness results are given in Section 4.

Remark 1 In the text, we focus on implementation in pure strategy equilibria. Most of the proofs cover mixed strategy equilibria also and Appendix H shows that all results carry over to implementation in mixed equilibria.

3 Main Results

This section contains our main implementation results, while robustness is discussed in Section 4.

In this section, we first give a simple condition, *measurability*, which is necessary for the implementation of a social choice function when the preferences are state independent. We then turn to environments where the social planner can perform monetary transfers among the players and show that this condition is also sufficient for implementation of f for *every* preference structure, state independent or otherwise. Theorem 1 establishes that it is possible to implement any SCF f satisfying measurability with a perfect information mechanism when there are at least two players under a boundedness assumption. Theorem 2 establishes that, with three or more players, any measurable f can be implemented with only "small" (epsilon) monetary transfers though at the cost of using a mechanism which involves an integer game. Furthermore, in this case, f can be implemented in a one–stage mechanism, establishing that it can be implemented in Nash equilibrium as well. Finally, in the special but important case of allocation problems, Theorem 3 establishes implementation with no transfers at all under weak conditions.

We turn now to a formal statement of our results.

Definition 1 *Given a social environment* Ψ*, we say that the SCF* f *satisfies* measurability *if for every pair of states* s *and* s' *with* $M_i(s) = M_i(s')$ *for all* i*, we have* $f(s) = f(s')$*.*

In other words, define two states to be equivalent if every agent has the same set of available evidence in the two states. Then f satisfies measurability if f is measurable with respect to the partition of S induced by this equivalence relation.

Put differently, measurability says that if the alternatives that are selected by f at the states s and s' are different from each other, then there exists some player i who can either refute state s' when the true state is s (*i.e.*, $M_i(s) \not\subseteq M_i(s')$) or can refute state s at s' ($M_i(s') \not\subseteq M_i(s)$).

Proposition 1 *Let* Ψ *be a social environment with state independent preferences and let* f *be an SCF for* Ψ*. If* f *can be implemented, then* f *satisfies measurability.*

Proof: Suppose f is implemented by the mechanism Γ. Fix any pair of states $s, s' \in S$ such that $M_i(s) = M_i(s')$ for all i. It is easy to see, then, that a history h is feasible in s iff it is feasible in s'. Since the preferences in states s and s' are also the same, we have $\Gamma(s) = \Gamma(s')$. Hence the set of subgame perfect equilibrium outcomes in $\Gamma(s)$ and $\Gamma(s')$ are the same. Since Γ implements f, then, we must have $f(s) = f(s')$, so f is measurable. ▌

Clearly, this result is not driven by our restriction to multistage games or our focus on subgame perfect equilibrium. The result holds for any class of mechanisms and any equilibrium concept simply because a mechanism cannot induce different games in states s and s' unless some agent's preference or evidence differs in the two states. The result highlights the obvious fact that for the outcome to vary across states, either preference variation or evidence variation is necessary.

Definition 2 *We say that Ψ is a social environment with* monetary transfers *if the following conditions hold. First, there is a set $\hat{A}$ such that the set of alternatives A can be written as*

$$A = \{(\hat{a}, t_1, \ldots, t_I) \in \hat{A} \times \mathbf{R}^I \mid \sum_{i \in \mathcal{I}} t_i \leq 0\}.$$

Second, for each i, the preference relations $\succeq_{i,s}$ for $s \in S$ can be represented by a utility function $u_i : A \times S \to \mathbf{R}$ of the form

$$u_i((\hat{a}, t_1, \ldots, t_n), s) = v_i(\hat{a}, s) + t_i$$

for some function $v_i : \hat{A} \times S \to \mathbf{R}$. Finally, player i's preferences over lotteries over A are represented by the expectation of u_i.

As we explain below, the linearity of u_i in t_i is not critical for our results. The assumption on preferences over lotteries is made in order to establish that our results are valid not only for *pure* subgame perfect equilibrium but also for mixed equilibria.

The following terminology will be convenient. We say that an SCF $f : S \to A$ is *essential* if for every $s \in S$, there exists an alternative $\hat{a}(s) \in \hat{A}$ such that $f(s) = (\hat{a}(s), 0, \ldots, 0)$. We have the following interpretation in mind. The social planner is interested in implementing a function which maps S to $\hat{A}$, not A. To obtain this goal he uses monetary transfers as incentives to induce revelation of the true state. We emphasize that the social choice function f may itself call for transfers and that the t_i's are "above and beyond" what f calls for. That is, $\hat{A}$ itself may have a product structure which allows for transfers. In this sense, the restriction to essential f's is without loss of generality.

Given any $\varepsilon > 0$, we say that a mechanism uses ε *monetary transfers* if for every terminal history h, if $g(h) = (\hat{a}, t_1, \ldots, t_I)$, we have $|t_i| \leq \varepsilon$ for all i. We say that a mechanism is *budget–balanced* if for every terminal history h, $\sum_{i \in \mathcal{I}} t_i = 0$ where $g(h) = (\hat{a}, t_1, \ldots, t_I)$.

Let

$$V(\Psi) = \max_{i,s} \sup_{\hat{a},\hat{a}' \in \hat{A}} v_i(\hat{a}, s) - v_i(\hat{a}', s).$$

Thus $V(\Psi)$ is an upper bound on the monetary value of a change in the selection of an alternative in $\hat{A}$ for every player i and every state s. We say that an environment is *bounded* if $V(\Psi)$ is finite.[6]

Theorem 1 *Let Ψ be a bounded social environment with monetary transfers and at least two agents. Let f be an essential SCF for Ψ that satisfies measurability. Then there exists a perfect information mechanism Γ_f that implements f. If there are at least three agents, there is such a mechanism satisfying budget–balance.*

Remark 2 Theorem 1 has the following simple consequence. Consider a bounded social environment Ψ with monetary transfers and at least two agents such that for every s and s', with $s \neq s'$, there is some i with $M_i(s) \neq M_i(s')$. Then *every* essential social choice function f can be implemented with a perfect information mechanism. If there are at least three agents, the same is true restricting to budget–balanced mechanisms.

To see the intuition for Theorem 1, consider the following example.

Example 1.

There are two players, 1 and 2, and two states, s_1 and s_2. Let $f(s_i) = a_i = (\hat{a}_i, 0, 0)$, $i = 1, 2$, where $\hat{a}_1 \neq \hat{a}_2$. Let $M_1(s_1) = \{S, \{s_1\}\}$ and $M_1(s_2) = M_2(s_1) = M_2(s_2) = \{S\}$. That is, in state s_1, 1 can prove that s_1 is the true state, but 1 cannot prove anything in s_2 and 2 can never prove anything.

We can implement f with no assumptions on the preferences aside from the possibility of monetary transfers and boundedness (which is implied by finiteness of the state space). To see this, let F_1 and F_2 be numbers such that $F_2 > F_1 > V(\Psi)$. Consider the following perfect information mechanism. First, player 1 provides whatever evidence he likes. If he provides the evidence $\{s_1\}$, the game ends and the outcome is $f(s_1)$.

If 1 does not provide evidence, this is interpreted as a claim that the state is s_2. In this case, player 2 chooses between cheap talk messages s_1 and s_2. The former is interpreted as claiming that 1 lied, while the latter supports 1's claim. If 2's message is s_1, again, the game ends, this time with an outcome of $(\hat{a}_2, -F_1, -F_1)$. That is, if 2 says that 1 lied, both agents pay a fine of F_1. As will be clear shortly, the choice of $\hat{a}_1$ versus $\hat{a}_2$ in this situation plays no role, but $\hat{a}_2$ is used for concreteness.

[6] Obviously, Ψ is bounded if the set of states is finite.

If 2's message is s_2, so that he supports 1's initial claim, then 1 gets another chance to provide evidence. This time, if 1 provides evidence $\{s_1\}$, 2 is punished and 1 is rewarded. More specifically, the outcome is $(\hat{a}_2, F_2, -F_2)$. If 1 provides no evidence again, the game ends with outcome $f(s_2)$. See Figure 1 for the game tree.

To see that this mechanism implements f, first consider state s_2. Since 1 cannot prove the event $\{s_1\}$ in state s_2, he has only one available move. At his information set, 2 chooses between saying s_1 and receiving payoff $v_2(\hat{a}_2, s_2) - F_1$ or saying s_2 and receiving $v_2(\hat{a}_2, s_2)$. Since $F_1 > 0$, 2 says s_2, so the outcome in s_2 is $f(s_2)$.

Now consider state s_1. At his last decision node, 1 can provide evidence $\{s_1\}$ and obtain payoff $v_1(\hat{a}_2, s_1) + F_2$ or not provide it and obtain $v_1(\hat{a}_2, s_1)$. Since $F_2 > 0$, he will provide the evidence.

Given this, consider 2's information set. If 2 chooses s_2, his payoff is $v_2(\hat{a}_2, s_1) - F_2$, while choosing s_1 instead yields $v_2(\hat{a}_2, s_1) - F_1$. Since $F_2 > F_1$, 2 will choose s_1 in state s_1.

Finally, then, consider 1's initial move. If 1 provides the evidence $\{s_1\}$, the game ends with outcome $f(s_1)$ and 1's payoff will be $v_1(\hat{a}_1, s_1)$. If he does not, then in the next stage, 2 will say s_1 and 1's payoff will be $v_1(\hat{a}_2, s_1) - F_1$. Since $F_1 > V(\Psi) \geq v_1(\hat{a}_2, s_1) - v_1(\hat{a}_1, s_1)$, 1 will provide the proof. Hence in either state s, the equilibrium outcome is $f(s)$.

In this game, budget balance does not obtain if player 2 accuses player 1 of lying by saying s_2. The fines charged to the two players cannot be given to either of them without interfering with their incentives. On the other hand, if there were a third player, he could receive the fines in this situation, allowing budget balance. This is why we obtain budget balance with at least three players but not, in general, with only two.

Note that the only thing SP needs to know about players 1 and 2's preferences in order to set up this mechanism is how large the fines need to be. As long as he can bound the v_i differences, he does not need to know anything else about the players' preferences. Similarly, these are the only facts players 1 and 2 need to know about each other's preferences. In Section 4, we formalize this notion of robustness and generalize this observation.

The mechanism of Theorem 1 does not require the planner to know much about the players or the players to know much about each other and has the appealing feature that it is a perfect information game. However, the mechanism relies on large monetary fines off the equilibrium path. The next result improves on this, though at the cost of moving to a mechanism which relies on an integer game (and ruling out the case of two players). In this case, the planner needs even less information about the preferences of the players as the boundedness assumption is no longer needed. Also, the mechanism

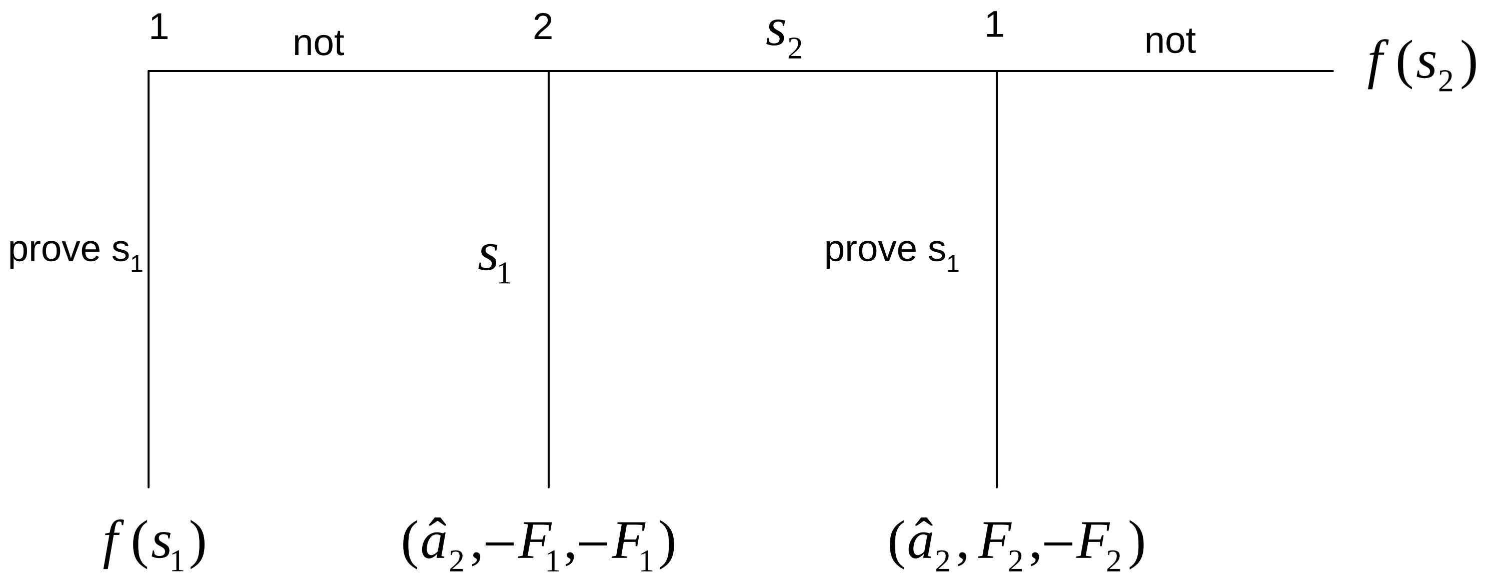
1
not
2
s_2
1
not
$f(s_2)$
prove s1
s_1
prove s1
$f(s_1)$
$(\hat{a}_2, -F_1, -F_1)$
$(\hat{a}_2, F_2, -F_2)$

Figure 1.

we use is a one–stage mechanism, so implementation is achieved in Nash equilibrium. Note that implementation in Nash equilibrium is more difficult to achieve in the sense that implementation in Nash equilibrium implies implementation in subgame perfect equilibrium but not conversely.[7]

Theorem 2 *Fix any $\varepsilon > 0$. Let Ψ be a social environment with monetary transfers and at least three players. Let f be an essential SCF for Ψ that satisfies measurability. Then there exists a budget–balanced one–stage mechanism with ε transfers Γ_f that implements f.*

Remark 3 As in the case of Theorem 1, we get a simple but striking corollary for social environments with monetary transfers and at least three agents such that for every s and s', with $s \neq s'$, there is some i with $M_i(s) \neq M_i(s')$. For such environments, *every* essential social choice function can be implemented by a budget–balanced mechanism with ε monetary transfers.

To see the intuition, consider the following variation on Example 1.

Example 2.

We now add a third state, s_3, and a third player. As before, we assume $M_1(s_1) = \{\{s_1\}, S\}$ and $M_1(s) = \{S\}$ for $s \neq s_1$. That is, player 1 can prove s_1 if it is true and nothing otherwise. We now assume $M_2(s_2) = \{\{s_2\}, S\}$ and $M_2(s) = \{S\}$ for $s \neq s_2$, so player 2 can prove s_2 if it is true and nothing otherwise. Finally, we assume $M_3(s) = \{S\}$ for all s, so player 3 never has any evidence.

Consider the following mechanism. The players move simultaneously, sending to the mechanism four things: evidence, a claim of a state, a requested outcome $\hat{a} \in \hat{A}$, and an integer. Let i's message be denoted $(E_i, c_i, \hat{a}_i, z_i) \in M_i \times S \times \hat{A} \times Z$ where $M_i = \cup_s M_i(s)$ and Z denotes the integers. (Thus $\hat{a}_i$ now refers to the $\hat{a}$ named by i, not the $\hat{a}$ in $f(s_i)$.) For simplicity, we constrain these messages so that if i proves some state s, then he must claim this s and conversely. That is, if player 1 provides evidence $\{s_1\}$, then his claim c_1 must equal s_1 and otherwise he cannot claim s_1 and similarly for player 2 regarding s_2. More precisely, player 1 has only two possible c_1's, either s_2 or s_3, and we interpret him as claiming s_1 when he proves it, and analogously for player 2.

The outcome is determined as a function of the claims as shown in the table below. In the matrix, $(f(s_j), i)$ is short for the outcome $f(s_j)$ with a transfer of ε to player i

[7]To see this, simply note that if a mechanism implements in Nash equilibrium, then, viewing the mechanism as a one–stage game, we see that it also implements in subgame perfect equilibrium.

paid for by the other two players. "n.a." is short for "not applicable" — note that it is impossible for player 1 to claim (and thus prove) s_1 at the same time as player 2 claims (and thus proves) s_2. Finally, "int" is short for "integers" — meaning that the outcome is determined by the integers in a manner explained after the table.

	s_1	s_2	s_3
s_1	$f(s_1)$	n.a.	$f(s_1)$
s_2	$f(s_1)$	$f(s_2)$	int
s_3	$f(s_1)$	int	$f(s_3)$

s_1

	s_1	s_2	s_3
s_1	$f(s_1)$	n.a.	int
s_2	$f(s_2)$	$f(s_2)$	$f(s_2)$
s_3	int	$f(s_2)$	$f(s_3)$

s_2

	s_1	s_2	s_3
s_1	$f(s_1)$	n.a.	$f(s_3), 1$
s_2	int	$f(s_2)$	$f(s_3)$
s_3	$f(s_3)$	$f(s_3), 2$	$f(s_3)$

s_3

In other words, if all players claim state s, then the outcome is $f(s)$ regardless of the outcomes or integers named by the players. If two out of three players claim s and the third does not provide proof that s is false, again, the outcome is $f(s)$. If two out of three claim s and the third does prove that s is false, the outcome is $f(s)$ but with a transfer of ε to the third player, paid by the first two. Finally, if the three players make conflicting claims, the outcome is determined by the player who chose the largest integer. Specifically, if i chose the largest integer (where we break ties in any fashion), then the outcome has $\hat{a} = \hat{a}_i$. In addition, i receives a transfer of ε and the other two players pay $-\varepsilon/2$.

To see that this mechanism implements f, we first show that for every state s, there is a Nash equilibrium with outcome $f(s)$. In this equilibrium, each player i sets $c_i = s$, providing the required proof as appropriate. It is easy to see that no feasible unilateral deviation by any player can change the outcome, so this is an equilibrium with outcome $f(s)$. Note, in particular, that if the state is s_3, player 1 cannot claim s_1 since he cannot prove it and analogously for player 2.

To see that there is no other Nash equilibrium outcome, first note that there is no equilibrium in which the outcome is determined by the integers chosen. If so, any player who did not choose the largest integer would deviate to a larger integer since he could change the $\hat{a}$ part of the outcome (if desired) and replace paying $\varepsilon/2$ with receiving a payment of ε. Similarly, there is no equilibrium in which a unilateral deviation by any player could make the outcome be determined by the integers. To see this, suppose such an equilibrium did exist. Then the player who could deviate to making the integers relevant would strictly gain by deviating to a claim which makes the integers determine the outcome, an integer larger than that selected by any player, and $\hat{a}_i$ equal to the $\hat{a}$ that would have obtained otherwise. It is easy to see by inspection of the table the deviating player would have either gotten no transfer or paid $\varepsilon/2$, while the deviation leaves $\hat{a}$ unchanged, but earns the deviator a "reward" of ε, paid for by the other two players. Obviously, then, this would be a profitable deviation.

By inspection of the matrix, these considerations rule out all strategy profiles other than those where all players claim the same state. Since i cannot claim s_i except in state

s_i for $i = 1, 2$, we cannot have an equilibrium where all players claim s_1 or s_2 unless the claim is true. Hence the only possibility for an equilibrium with the "wrong" outcome is if all players claim s_3 when the true state is either s_1 or s_2. It is easy to see that such strategies cannot form a Nash equilibrium since either 1 or 2 would be able to deviate to refuting this claim. By doing so, the deviator would not change $\hat{a}$ but would earn a reward of ε. Hence the deviation would be profitable, so the unique equilibrium outcome in state s is $f(s)$.

Note that the analysis of equilibria did not use any properties of the agent's preferences other than the fact that more money is better than less. Thus this is all the planner needs to know about the preferences of the players to set up this mechanism and know it will implement. Similarly, this is all the players need to know about each other. Theorem 4 in the next section formalizes and generalizes this observation.

Remark 4 While we assume that utility is linear in transfers, this is not necessary for Theorems 1 and 2. As the examples above suggest, the only properties we use in the proofs are the following. First, both theorems use the assumption utility is strictly increasing in one's own transfer and independent of any other agent's transfer. Second, Theorem 1 uses the assumption that there exists an i such that for any $\hat{a}$ and $\hat{a}'$, there are transfers t_i and $\hat{t}_i$ such that i strictly prefers $(\hat{a}, t_i)$ to $(\hat{a}', 0)$ to $(\hat{a}, \hat{t}_i)$ at every state s.

Next, we discuss a special but important case where implementation is achieved in a perfect information game with no transfers at all. Consider the problem of the allocation of a set of goods among a set of agents, one of the most basic problems in economics. It turns out that a simple modification of the proof of Theorem 1 establishes that in such a problem, every SCF that satisfies measurability and assigns each agent a positive amount of at least one divisible good can be implemented by a perfect information mechanism without monetary transfers. The basic intuition is simple: we can impose the equivalent of a large fine on agent i by giving the goods he would have received according to f to agent j instead.

More formally, we say that a social environment Ψ is an *allocation environment* if the following three statements hold. First, there exists an integer $K \geq 0$, nonempty sets $\bar{A}_k \subseteq \mathbf{R}$, $k = 1, \ldots, K$, and positive numbers $\bar{x}_k$, $k = 1, \ldots, K+1$, such that

$$A = \left\{ (x^1, \ldots, x^I) \mid x^i \in \left[\prod_{k=1}^{K} \bar{A}_k \right] \times \mathbf{R}_+ \text{ and } \sum_{i=1}^{I} x_k^i \leq \bar{x}_k, \ k = 1, \ldots, K+1 \right\}.$$

That is, A is a set of allocations of $K+1$ goods, at least one of which is divisible. Second, for every $i \in \mathcal{I}$ and $s \in S$, $(x^i, x^{-i}) \sim_{i,s} (x^i, \bar{x}^{-i})$. In other words, in any state, agent i is indifferent between allocations which give him the same goods. Finally, for every $i \in \mathcal{I}$

and every $s \in S$, $(x^i, x^{-i}) \succ_{i,s} (\bar{x}^i, x^{-i})$ if $x^i > \bar{x}^i$ (where $>$ denotes the vector ordering of weakly larger in every component, strictly larger in at least one component).

Given $\varepsilon > 0$, define

$$A^\varepsilon = \left\{ x \in A \mid x^i_{K+1} \geq \varepsilon, \ \forall i \in \mathcal{I} \right\}.$$

Theorem 3 *Fix any allocation environment with at least two players and any $\varepsilon > 0$. Let $f : S \to A^\varepsilon$ be a SCF that satisfies measurability. Then there exists a perfect information mechanism Γ_f that implements f.*

In a paper written after preliminary drafts of this paper were circulated, Kartik and Tercieux (2009) give results which extend Proposition 1 and Theorem 2.[8] They give a condition called evidence–monotonicity which combines Maskin monotonicity and our measurability condition and show this is necessary for Nash implementation. They then show that, given some additional conditions, evidence–monotonicity is also sufficient. Roughly speaking, evidence–monotonicity says that a pair of states should either satisfy our measurability condition or be related in the way described by Maskin monotonicity.[9]

The fact that their evidence–monotonicity condition refers both to evidence and to preferences allows them to obtain a condition weaker than the assumptions needed if preference variation alone or evidence variation alone is used to obtain implementation. On the other hand, the use of preference variation also implies that they do not obtain the robustness results mentioned earlier and discussed in detail in the next section.

We conclude this section with a comment on the relationship between the preference variation used in the standard literature to obtain implementation and our focus on evidence variation. For simplicity, consider an implementation problem where there are just two states,s_1 and s_2. In a standard model, such as Maskin (1977) or Moore and Repullo (1988), the implementation of the social choice function relies on preference reversals. For example, in Moore and Repullo, the implementation of the SCF relies on having an agent i and alternatives a and b such that i prefers a to b in state s_1 but has the opposite preference in state s_2. In this case, agent i can, in effect, "prove" the true state by choosing between a and b. If he chooses a, he effectively proves the state is s_1, while if he chooses b, he proves s_2.

Interestingly, with evidence, implementation is possible even when there is provability in only one direction. For example, in Example 1, 1 can prove the state is s_1 when this is true but neither agent can ever prove that the state is s_2.

[8]They also consider implementation with costly evidence fabrication.

[9]Intuitively, if a pair of states satisfy Maskin's monotonicity condition, then the variation in preferences across the two states can be exploited to implement different outcomes without evidence just as in Maskin's mechanism

Also, the social planner does not need to know as much about the preferences of agents if preference variation is not exploited. Further, the agents do not need to know as much about the preferences of other agents. We turn to formal statements of the robustness properties of our mechanisms in the next section.

4 Robustness

In this section, we define the notion of robustness discussed informally so far and demonstrate that our results are robust in this sense. In particular, we show that when the social choice function satisfies measurability, the planner can implement even if he does not know the agent's preferences in any state $s \in S$ (beyond that they satisfy the requirements of a monetary environment and, for one result, are bounded in a certain sense) or the agent's beliefs about other agents. Furthermore, we do not need the agents to know anything more than the planner about other agents' preferences or even to have a common prior. In addition, since measurability imposes very little on the structure of evidence, the planner requires little information about the evidence available to the agents. We obtain such robustness because we exploit variation in evidence across states rather than variation in preferences.

To demonstrate these facts, we make some changes in the model which apply only in this section. As in the rest of the paper, we let S denote the set of states, A the set of social alternatives, $f : S \to A$ the social choice function, $\mathcal{I}$ the set of players, and $M_i(s)$ the set of subsets of S that i can prove in state s. Our goal is to show robustness with respect to the planner's knowledge about the preferences and beliefs of the agents. Therefore, we now think of a state s as a specification of all the parameters that are relevant for the determination of the social alternative that SP wishes to implement and a specification of the evidence of each agent. However, s may not include enough information to identify the preferences of each agent or the beliefs of agents about the preferences of others.

Formally, we extend the model by adding *types* for each agent. For each i, let Θ_i denote i's set of types. A type $\theta_i \in \Theta_i$ determines the preferences of i and his beliefs over Θ_{-i}, both as a function of the state s which, as before, is assumed to be common knowledge among the agents. Of course, i's type is private information. For simplicity, we assume that each Θ_i is finite or countable. The set of all *full states* is Ω where $\Omega \subseteq S \times \prod_{i \in \mathcal{I}} \Theta_i$. Note that we do not require the set of full states to have a product structure. Thus we are not imposing any restrictions on whether types are correlated across players or are correlated with s. In particular, this formulation allows any relationship between the preferences of the agents and the planner's desired outcome. Given any s, we refer to a tuple $(s, \theta) \in \Omega$ as a full state consistent with s. We also say that θ_i is *consistent with s* if there exists θ_{-i} with $(s, \theta_i, \theta_{-i}) \in \Omega$. Let $\Theta_i(s)$ denote the set of θ_i consistent with s.

Agent i's utility function in full state $\omega = (s, \theta)$ is assumed to depend only on s and θ_i. That is, while we demonstrate robustness with respect to uncertainty about the preferences of other agents, we retain the assumption that each agent knows his own payoffs. Thus we write

$$u_i(\hat{a}, t_1, \ldots, t_I, s, \theta_1, \ldots, \theta_I) = v_i(\hat{a}, s, \theta_i) + t_i.$$

The belief of θ_i at s over the types of the other players is denoted by $\mu_i(s, \theta_i) \in \Delta(\Theta_{-i})$. (As usual, Θ_{-i} is the set of profiles of types of the agents other than i and $\Delta(\Theta_{-i})$ is the set of probability distributions over Θ_{-i}.) We do not require the agents to have a common prior. However, we do impose a common support on the agent's beliefs in the sense that

$$\text{supp}(\mu_i(s, \theta_i)) = \{\theta_{-i} \in \Theta_{-i} \mid (s, \theta_i, \theta_{-i}) \in \Omega\}$$

for all i, $s \in S$, and all $\theta_i \in \Theta_i$ consistent with s.

In this section only, we refer to a tuple

$$\Psi^* = \left\langle \mathcal{I}, A, S, (M_i(s))_{i \in I, s \in S}, (\Theta_i)_{i \in \mathcal{I}}, \Omega, (\mu_i(s, \theta_i), v_i(\cdot, s, \theta_i))_{i \in \mathcal{I}, s \in S, \theta_i \in \Theta_i(s)} \right\rangle$$

as a *monetary environment with partial information*. We use the term partial information instead of the more common term "incomplete information" to emphasize that we are considering an environment where the state s is common knowledge among the agents. An SCF for an environment with partial information is a function $f : S \to A$. The definition of a mechanism is unchanged. Given an environment with partial information Ψ^* and a state s, a mechanism Γ now induces a game of incomplete information among the agents. We use $\Gamma(s \mid \Psi^*)$ to denote this game. Note that a strategy for player i in this game is a function of θ_i, so an equilibrium outcome is a function of θ.

Since we use different equilibrium notions for our two results, we state the definition of implementation for an arbitrary equilibrium concept. We say that Γ implements f in environment Ψ^* if for every $s \in S$, for every equilibrium of $\Gamma(s \mid \Psi^*)$, the equilibrium outcome given θ is $f(s)$ for every θ such that $(s, \theta) \in \Omega$.

The robustness properties of our mechanisms differ slightly across results for two reasons. First, Theorem 1 requires that the environment be bounded, while Theorem 2 does not. Second, Theorem 2 is based on a one–stage mechanism, so when we introduce incomplete information among the players, Bayes–Nash equilibrium is the appropriate solution concept. On the other hand, Theorem 1 uses a game of perfect information, so the robust version will require use of perfect Bayesian equilibrium. Since Theorem 2 is simpler on both criteria, we begin with it.

Theorem 4 (Robust Version of Theorem 2) *Fix any $\varepsilon > 0$. Fix any monetary environment with partial information Ψ^* with at least three players. Let f be an essential SCF for Ψ^* that satisfies measurability. Then there exists a budget–balanced one–stage mechanism with ε transfers Γ_f that implements f in Bayes–Nash equilibrium.*

For any monetary environment with partial information Ψ^*, let

$$V(\Psi^*) = \sup_{i \in \mathcal{I}, s \in S, \theta_i \in \Theta_i(s), \hat{a}, \hat{a}' \in \hat{A}} v_i(\hat{a}, s, \theta_i) - v_i(\hat{a}', s, \theta_i).$$

We say that Ψ^* is bounded if $V(\Psi^*) < \infty$.

Since we analyze environments with partial information in this section, the game induced by the mechanism we used in Theorem 1 is no longer a game of perfect information. Hence we use the term *sequential mechanism* to refer to a mechanism for which each information set is either a singleton or contains nodes that differ only in Nature's moves. That is, all past actions are observable and there are no simultaneous moves. We emphasize that this is just a change of name — the mechanism is the same one we used earlier.

Theorem 5 (Robust Version of Theorem 1) *Let Ψ^* be a bounded monetary environment with partial information and at least two agents. Let f be an essential SCF for Ψ^* that satisfies measurability. Then there exists a sequential mechanism Γ_f that implements f in perfect Bayesian equilibrium. If there are at least three agents, there is such a mechanism satisfying budget–balance.*

Remark 5 Definitions of perfect Bayesian equilibrium in the literature impose a variety of conditions on beliefs off the equilibrium path. We put no restrictions on such beliefs.

Since Theorem 3 is essentially a generalization of Theorem 1, we omit the straightforward generalization of Theorem 5 which gives a robust version of it.

Summarizing, Theorems 4 and 5 formalize the robustness results we discussed in the previous sections. More specifically, subject to the boundedness assumption in the case of Theorem 5,

1. The social planner does not need to know anything about the preferences of the players — nothing about how their preferences relate to the evidence available, the social choice to be implemented, or each other.

2. No agent needs to know anything about the preferences of other agents.

In addition, we note that the only information the planner needs to have about evidence is what is summarized by the measurability condition. That is, he needs to know that if $f(s) \neq f(s')$, then *some* agent has different evidence in the two states, but he does not need to know anything about *which* agent does. In this sense, the planner requires very little information about the evidence. We do not provide a formal proof of this claim as it is a notationally complicated but straightforward extension of our other results.[10]

5 Discussion

In this section, we present a series of examples to show that stronger results cannot be obtained without additional hypotheses.[11] In particular, we demonstrate the following:

1. With one player, it may not be possible to implement a measurable f even with unbounded transfers. Thus Theorems 1 and 2 do not extend to the one agent case.
2. With two players, it may not be possible to implement a measurable f with ε transfers, whether or not we require budget–balance or restrict attention to perfect information games. Thus Theorem 1 does not extend to ε transfers and Theorem 2 does not extend to the case of two players.
3. Even with three players, it may not be possible to implement a measurable f with ε transfers in a perfect information game, even if we do not require budget–balance. Thus Theorem 1 does not extend to ε transfers even with more than two players and Theorem 2 does not extend to perfect information games even with more than two players.
4. For any number of players, it may not be possible to implement a measurable f without monetary transfers even when we allow for a general multistage mechanism. Thus Theorem 2 does not extend to the case where there are no transfers.

[10] More precisely, it is easy to obtain versions of Theorems 1, 2, 4, and 5 where the evidence available for each agent is common knowledge among the agents but is not determined by the state s. Hence the social planner does not know what evidence each agent has at each state s. He just knows that if $f(s)$ is different from $f(s')$, then some agent has different evidence across the two states.

[11] As we show in Section H of the Appendix, these claims hold whether or not we restrict attention to pure strategy equilibria.

5. Another question of interest is the "complexity" of the mechanisms needed to implement. In our proof of Theorem 1, we use a perfect information game with $I + 1$ stages, so one player gets two opportunities to send a message or present evidence, while all other players get only one such opportunity. While one can give examples where shorter mechanisms suffice, we provide an example which demonstrates that Theorem 1 does not hold if we require I or fewer stages.

All the examples we present are variations on Example 1 in Section 3.

We begin with the first point above. So consider Example 1 but now with only agent 1. Suppose his preference is $f(s_2) \succ_{1,s} f(s_1)$ for all s. Then it is impossible to implement f, robustly or otherwise, with large fines or small. To see this, suppose that f can be implemented. Hence there is a mechanism with outcome $f(s_2)$ in state s_2. But then player 1 can use the same strategy in state s_1 that he uses in s_2 to obtain $f(s_2)$ in state s_1. Since he prefers this to $f(s_1)$, it cannot be true that we implement f.[12]

Turning to the second point, recall that in our discussion of Example 1 in the previous section, we made no assumptions about preferences other than boundedness and showed that we could implement with a perfect information mechanism and large transfers. Now we demonstrate the second point by giving specific preferences for which we cannot implement with any perfect information mechanism restricted to "small" transfers. So consider the state independent preferences where $v_1(\hat{a}_2) = v_2(\hat{a}_1) = 1$ and $v_1(\hat{a}_1) = v_2(\hat{a}_2) = 0$. As in Example 1, the SCF is $f(s_k) = (\hat{a}_k, 0, 0)$. Also, player 1 can prove the state is s_1 in state s_1 and can prove nothing in s_2, while 2 cannot prove anything in either state. Clearly, if there are other feasible outcomes in $\hat{A}$ which serve the same role as transfers, then a bound on transfers is irrelevant. So we assume that $\hat{A} = \{\hat{a}_1, \hat{a}_2\}$. We assume transfers are small in the sense that we consider only mechanisms such that for every terminal history h, if $g(h) = (\hat{a}, t_1, t_2)$, then $|t_i| < 1/2$ for $i = 1, 2$.

In Section E of the Appendix, we prove the following.

Claim 1 *In this example, there is no mechanism with transfers bounded below 1/2 which implements f.*

For the third point, we show that we cannot combine the results of Theorems 1 and 2 to obtain implementation in a perfect information game with ε transfers, even with three or more agents. We show this by adding a third player to the previous example. Assume player 3 is exactly like player 2 in the sense that he has no evidence in any state and has the same utility function as player 2. By Theorem 1, we can implement f in a perfect

[12] See Glazer and Rubinstein (2004, 2006) for an interesting treatment of the one agent case.

information game without a bound on transfers. By Theorem 2, we can implement f in a game without perfect information but with only ε transfers. However, in Section F of the Appendix, we prove the following.

Claim 2 *In this example, there is no perfect information mechanism with transfers bounded below 1/2 which implements f.*

In both of these arguments, budget–balance plays no role, so dropping such a restriction does not overcome the negative results in the examples.

For the fourth point, we show that for any integer n, there exists an environment with n players and a measurable SCF that cannot be implemented in a multistage mechanism without transfers. We show this by generalizing the previous example so that there are $n-1$ players in the position of player 2 above. That is, players $2, \ldots, n$ have no evidence in either state and have the same state–independent utility function $v_i(\hat{a}_1) = 1$ and $v_i(\hat{a}_2) = 0$. Player 1, as above, can prove s_1 in state s_1 and nothing in s_2. He has the state–independent utility function $v_1(\hat{a}_1) = 0$ and $v_1(\hat{a}_2) = 1$. By Theorem 2, f can be implemented with a mechanism which involves only ε monetary transfers. However, in Section G of the Appendix, we prove the following.

Claim 3 *In this example, there is no multistage mechanism without transfers that implements f.*

Finally, to demonstrate the fifth point, consider the following example. Again, we have two players, 1 and 2. This time we have three states, s_1, s_2, and s_3. Player 1 can prove s_1 in s_1 and nothing in any other state, while player 2 can prove s_2 in s_2 and nothing in any other state. That is, $M_i(s_i) = \{S, \{s_i\}\}$, $i = 1, 2$, and $M_i(s_j) = \{S\}$ whenever $i \neq j$. Finally, assume preferences are state independent with $f(s_3) \succ_1 f(s_2) \succ_1 f(s_1)$ and $f(s_3) \succ_2 f(s_1) \succ_2 f(s_2)$. We show that f cannot be implemented in a perfect information game with two or fewer stages with these preferences.

Claim 4 *In this example, there is no perfect information mechanism with two or fewer stages that implements f.*

Proof: Suppose we have a perfect information mechanism with two or fewer stages and where player 2 moves first. Suppose there is an equilibrium in state s_3 with outcome $f(s_3)$. Thus there is a cheap talk message for player 2, say c^*, such that the continuation game leads to outcome $f(s_3)$. Suppose that if player 2 plays c^*, either the game ends or player 2 moves again, after which the game ends. Either way, this implies that player 2 could obtain outcome $f(s_3)$ in any state. Since this is player 2's most preferred outcome, we could not then implement f. Hence if we implement f, it must be true that player 1 gets a chance to speak after player 2 sends message c^*.

But then consider state s_2. Since player 1 has the same set of evidence available in states s_2 and s_3, player 1's best reply must be the same in the two states. Since the outcome is $f(s_3)$ in state s_3, this implies that if player 2 plays this message, the outcome is also $f(s_3)$ in state s_2. Since player 2 prefers this outcome to $f(s_2)$, again, f cannot be implemented. A symmetric argument covers the case where player 1 moves first. ∎

6 Conclusion

We have extended implementation theory in two ways. First, we allow the social choice function to depend on more than just the preferences of the agents. Second, we allow agents to support their statements with hard evidence.

We have shown that the measurability condition which is necessary for the implementation of a social choice function f when preferences are state independent is also a sufficient condition, with or without state independence, for the subgame perfect implementation of f when the social planner can perform monetary transfers. Furthermore, f can be implemented even if the planner knows very little about the agents' evidence, preferences, or their beliefs about each other's preferences. Theorem 1 establishes implementation with a perfect information mechanism when there are at least two players and when the social planner can perform "large" monetary transfers. In this case, the implementation is robust in the sense that the planner only needs to know an upper bound on the players' willingness to pay to change the outcome. Theorem 2 establishes that when there are at least three players, f can be implemented with a one–stage mechanism using only ε monetary transfers but which relies on an integer game. In this case, the planner does not need any information about the preferences, nor do the players need any information about each others' preferences. Finally, in the special but important case of allocation problems, Theorem 3 shows that we can implement under weak conditions using a perfect information game with no transfers. Again, this mechanism implements regardless of the preferences of the agents. In all cases, the only information the planner requires about evidence is that measurability holds.

There are many interesting directions for future research. First, the mechanisms we use in our results appear to have a variety of additional robustness properties which may be worth formalizing and exploring further. It may be of interest to characterize the mechanisms which require the least information on the part of the planner and/or the agents.

Second, clearly, implementation with a perfect information mechanism is more appealing than implementation by a mechanism which relies on integer games. It may be interesting to determine what can be implemented using perfect information mechanisms that allow the social planner to randomize but which do not rely on large monetary transfers.

Other general directions of interest are results without monetary transfers (or other "structural" assumptions which serve the same role), models with incomplete information among the agents, restrictions on or costs of evidence provision, and models where the social planner has less commitment power.

A Proof of Theorem 1

Fix a bounded environment Ψ and a social choice function f satisfying measurability. We write f as $f(s) = (\hat{a}(s), 0, \ldots, 0)$. Let F_1 and F_2 be numbers that satisfy $F_2 > F_1 > V(\Psi)$. (Recall that $V(\Psi) = \sup_{i \in \mathcal{I},\ s \in S,\ \hat{a}, \hat{a}' \in \hat{A}} v_i(\hat{a}, s) - v_i(\hat{a}', s)$. By boundedness, this is finite.) Thus, for any player i, a monetary fine or reward of F_1 or F_2 outweighs any utility gain from changing the alternative that is selected.

The mechanism Γ_f consists of $I + 1$ stages. In stage 1, player 1 sends a cheap talk message $c_1 \in S$. At each stage i for $2 \leq i \leq I$, player i sends a message $c_i \in \{0, 1\}$ and has a chance to present evidence. Finally, in stage $I + 1$, player 1 gets a chance to present evidence.[13] To define the outcome, let the sequence of cheap talk messages and events proved be $c_1, (c_2, E_2), \ldots, (c_I, E_I), E_1$. For any s and i, define $\bar{E}_i(s) = \cap_{E \in M_i(s)} E$. Intuitively, this is what i would prove in state s if he presented all the evidence he has. Recall that, by assumption, $\bar{E}_i(s) \in M_i(s)$ for all i and s.

To understand the mechanism, think of player 1 as making a cheap talk claim of the state when he reports c_1. Each other player i presents evidence and states whether he agrees ($c_i = 1$) or disagrees ($c_i = 0$) with 1's claim. Finally, player 1 has a chance to prove that he lied in the first stage.

There are three cases for defining the outcome of the mechanism.

First, if $c_i = 1$ for $i = 2, \ldots, I$, and $E_i = \bar{E}_i(c_1)$ for $i = 1, \ldots, I$, then the outcome is $f(c_1)$. Thus, in this first case, every player provides exactly the maximum evidence he would have if player 1's report were true and all players agree with 1. In this case, SP selects the outcome $f(c_1)$.

For the remaining cases, we must have some $i \in \{2, \ldots, I\}$ with $c_i = 0$ and/or some $i \in \{1, 2, \ldots, I\}$ with $E_i \neq \bar{E}_i(c_1)$. Let i^* be the first player in stages 2 through $I + 1$ with either $c_i = 0$ or $E_i \neq \bar{E}_i(c_1)$.

The second case is where $\bar{E}_{i^*}(c_1) \not\subseteq E_{i^*}$. In this case, the outcome is $\hat{a}(c_1)$ with a transfer of $(I - 1)F_2$ to player i^* and a transfer of $-F_2$ for every other player. To understand this case, note that $\bar{E}_{i^*}(c_1) \not\subseteq E_{i^*}$ implies that player i^* proves that player 1's initial statement is false (that is, the true state is not c_1).[14] In this case, SP selects

[13]Our definition of a mechanism has the set of cheap talk messages for a player independent of the subgame and allows players to send evidence each time they speak. Obviously, we can satisfy these requirements by letting 1 send evidence in the first stage and a cheap talk message in S at the last stage but having the mechanism disregard both messages.

[14]To see that $\bar{E}_{i^*}(c_1) \not\subseteq E_{i^*}$ implies that c_1 has been proven false, note that it implies that there is some $s \in \bar{E}_{i^*}(c_1)$ with $s \notin E_{i^*}$. But recall that $\bar{E}_{i^*}(c_1) = \cap_{E \in M_{i^*}(c_1)} E$, so $s \in E$ for every $E \in M_{i^*}(c_1)$.

$\hat{a}(c_1)$ but gives player i^*, who was the first to prove that player 1's initial statement is false, a reward of $(I-1)F_2$ which is paid by a fine of F_2 on each one of the other players.

The only remaining case, then, is where $\bar{E}_{i^*}(c_1) \subseteq E_{i^*}$. Hence it must be true that i^* either explicitly accuses 1 of lying by setting $c_{i^*} = 0$ or implicitly accuses 1 of lying by not proving as specific an event as he should be capable of if player 1's claim were true or both. However, unlike the previous case, i^* does not prove that 1 lied. In this case, SP selects the outcome $\hat{a}(c_1)$ with a transfer of $-F_1$ to i^*, $-F_1$ to player 1, and $2F_1/(I-2)$ to every other player if there are at least three players. Note that we can have $i^* = 1$ in which case player 1 is fined $2F_1$. If there are only two players, the transfers to 1 and i^* are still $-F_1$ and the mechanism is not budget–balanced unless $i^* = 1$.

To see that this mechanism implements f, consider the induced game in state s and the subgame where player 1 claims $c_1 = s$. Since $c_1 = s$, the second case is impossible — no one can refute the true state. No player can change the $\hat{a}$ and so only has a choice between paying a fine of F_1 or avoiding fines. It is easy to see that the unique subgame perfect equilibrium of this subgame has each player i proving the event $\bar{E}_i(s)$ and using the message $c_i = 1$. Thus, player 1 can achieve the outcome $f(s)$ by setting $c_1 = s$.

So consider a subgame where player 1 sets $c_1 = s'$ for some $s' \neq s$. If $M_i(s) = M_i(s')$ for all i , the same reasoning as above shows that the unique subgame perfect equilibrium outcome is $f(s')$. However, by measurability, $f(s') = f(s)$, so such choices still lead to outcome $f(s)$.

So suppose player 1 sets $c_1 = s' \neq s$ where there is some player i with $M_i(s) \neq M_i(s')$. First, suppose there is some $i \neq 1$ with $M_i(s) \not\subseteq M_i(s')$. In this case, player i can refute state s'. That is, there must be some $E \in M_i(s)$ with $E \notin M_i(s')$ and hence such that $s' \notin E$. Clearly, then, i's strategy in any subgame perfect equilibrium of this subgame must be to prove such an event if no such event has been proved before his turn since this gives him a reward of $(I-1)F_2$. In light of this, the best possible outcome for player 1 is if some other player before i accuses 1 of lying without proof. So player 1's payoff must be weakly less than $v_1(\hat{a}(s'), s) - F_1 < v_1(\hat{a}(s), s)$, so this cannot be an equilibrium. Hence it must be true that $M_i(s) \subseteq M_i(s')$ for all $i \neq 1$.

Second, suppose $M_1(s) \not\subseteq M_1(s')$. That is, player 1 is the only player who can refute s'. In this case, if each player $i \neq 1$ proves $\bar{E}_i(s')$ at his turn and sends message $c_i = 1$, player 1 will refute his own claim in the last stage since this does not change $\hat{a}$ but earns him a reward of $(I-1)F_2$. Given this, consider player I. If at his turn, every previous player i has sent evidence $\bar{E}_i(s')$ and message $c_i = 1$, then player I is better off with message $c_I = 0$ (or proving less than $\bar{E}_I(s')$ if this is possible[15]). If he uses this message,

Hence $s \notin E_{i^*}$ implies $E_{i^*} \notin M_{i^*}(c_1)$. Thus, in state c_1, player i^* cannot present the evidence E_{i^*} and therefore by consistency $c_1 \notin E_{i^*}$.

[15]It is possible that $\bar{E}_I(s') = \bar{E}_I(s) = S$ in which case it is impossible for I to trigger the smaller fines

he does not change the $\hat{a}$ but reduces his fine from F_2 to F_1 and so gains. Hence some player must claim that player 1 lied, since player I will do so if no other player does. It follows that player 1's payoff to claiming s' is $v_1(\hat{a}(s'), s) - F_1 < v_1(\hat{a}(s), s)$, so this cannot be an equilibrium. Hence we must have $M_1(s) \subseteq M_1(s')$.

Summarizing, we must have $M_i(s) \subseteq M_i(s')$ for all i, but $M_i(s) \neq M_i(s')$ for some i. But then there is some agent i for whom every event E that he can prove satisfies $\bar{E}_i(s') \subset E$ (where $\subset$ denotes strict inclusion). Hence if 1 claims s', we will necessarily end up in the third case and player 1's payoff will be $v_1(\hat{a}(s'), s) - F_1 < v_1(\hat{a}(s), s)$. Hence, again, this cannot be an equilibrium.

Hence we have shown that the only subgame perfect equilibrium outcome in state s is $f(s)$, so the mechanism implements f. ▌

B Proof of Theorem 3

Fix an allocation environment Ψ and SCF f satisfying the conditions of the theorem. Let Γ_f denote the mechanism that was defined in the proof of Theorem 1. We will show that a simple modification of Γ_f implements f.

For every player $j \neq 1$, let $\underline{a}_j$ be the allocation where both player 1 and player j receive a bundle with $\varepsilon/2$ units of good $K + 1$ and zero units of every other good. If there are any other players, the remaining goods are allocated to them in any fashion. For every player $j \in \mathcal{I}$, let $\bar{a}_j$ denote the allocation where player j gets all of each good and every other player gets zero.

Define the mechanism $\hat{\Gamma}_f$ as follows. Whenever Γ_f selects an outcome where player 1 and a player j pay a fine of F_1 in Γ_f, the outcome under $\hat{\Gamma}_f$ is the allocation $\underline{a}_j$. Whenever Γ_f selects an outcome where each player different than j pays F_2 to j, the outcome under $\hat{\Gamma}_f$ is $\bar{a}_j$. Otherwise, the outcome under $\hat{\Gamma}_f$ is the same as the outcome under Γ_f.

A careful reading of the proof of Theorem 1 shows that it relies only on the following assumptions about preferences for any player i and states s and s'. First, outcome $\hat{a}(s')$ and receiving a payment of $(I-1)F_2$ is strictly preferred by i at s to any point in the range of f. Second, any point in the range of f is strictly preferred by i at s to outcome $\hat{a}(s')$ and paying a fine of F_1. Finally, outcome $\hat{a}(s')$ and paying a fine of F_1 is strictly preferred by i at s to outcome $\hat{a}(s')$ and paying a fine of F_2.

It is easy to see that all three properties hold for the replacements of the fines used

by the evidence he provides.

by the mechanism $\hat{\Gamma}_f$. More specifically, since $f(s)$ gives every agent at least ε of the divisible good and $\bar{a}_i$ gives i everything, i must strictly prefer $\bar{a}_i$ to every point in the range of f at every state. On the other hand, since $\underline{a}_i$ leaves 1 and i with only $\varepsilon/2$ of the divisible good and nothing else, this must be strictly worse for 1 and i than anything in the range of f at any state. Finally, since $\underline{a}_i$ gives 1 and i $\varepsilon/2$ of the indivisible good while $\bar{a}_k$, $k \neq i, 1$, leaves them with nothing, each prefers the former in every state. ▌

C Proof of Theorem 2

The proof is by construction of a one–stage mechanism Γ_f which implements f. In the mechanism we construct, every agent chooses a piece of evidence and a cheap talk message in $S \times \hat{A} \times Z$ where Z denotes the positive integers. We write a typical choice of cheap talk message for i as $(s_i, \hat{a}_i, z_i)$.

To define the outcome as a function of the profile of evidence and cheap talk reports, we distinguish between several cases. As in the proof of Theorem 1, let $\bar{E}_i(s)$ denote what i proves in state s if he presents all his evidence. That is, $\bar{E}_i(s) = \cap_{E \in M_i(s)} E$.

First, suppose there is a state s such that $(s_i, E_i) = (s, \bar{E}_i(s))$ for all i. In this case, the outcome is $f(s)$.

Second, suppose there is a state s and an agent j such that $(s_i, E_i) = (s, \bar{E}_i(s))$ for all $i \neq j$ but $(s_j, E_j) \neq (s, \bar{E}_j(s))$. There are two subcases. First, suppose $s \notin E_j$. In this case, the outcome is $f(s)$ with a transfer of $\varepsilon/2$ to j, $\varepsilon/2$ to that $i \neq j$ who chooses the largest z_i (breaking ties by choosing the largest i who names the largest z_i), and transfers of $-\varepsilon/(I-2)$ to the other agents. Second, suppose $s \in E_j$. In this case, the outcome is $f(s)$ with a transfer of ε to that $i \neq j$ who chooses the largest z_i (with ties broken as above) and $-\varepsilon/(I-1)$ to the other agents.

Finally, for any other profile of messages, the alternative that is selected is determined by the integers chosen. Specifically, let i be the player who chose the highest integer z_i (breaking ties as above). The alternative that is selected is $\hat{a}_i$ with a transfer of ε to player i and $-\varepsilon/(I-1)$ to every other player.

It is easy to see that for each s, there is an equilibrium with outcome $f(s)$. Specifically, the strategies $((s, f(s), 0), \bar{E}_1(s)), \ldots, ((s, f(s), 0), \bar{E}_I(s))$ form a Nash equilibrium with outcome $f(s)$ as no feasible unilateral deviation by any player can improve the outcome for him.

We now show that there is no (pure or mixed) Nash equilibrium in state s with an outcome different from $f(s)$. Fix any mixed strategies σ that form an equilibrium of $\Gamma(s)$. Let H^* denote the set of pure strategy profiles that fall into any case where the outcome is determined by the integers. It is easy to see that if $h \in H^*$, then $\sigma(h) = 0$. Otherwise, there must be some player who could increase his integer and improve his expected payoff conditional on $h \in H^*$ and hence improve his unconditional expected payoff.

Hence for any h with positive probability, we have $(s_i, E_i) = (s', \bar{E}_i(s'))$ for all i. Clearly, this implies that every i has $(s_i, E_i) = (s', \bar{E}_i(s'))$ with probability 1. Otherwise, there is a positive probability of a realization $h \in H^*$.

The outcome under these strategies is $f(s')$. Suppose that the outcome is not $f(s)$, so $f(s) \neq f(s')$. By measurability, there is some i with $M_i(s) \neq M_i(s')$. Since every i is presenting evidence $\bar{E}_i(s')$, it must be true that $\bar{E}_i(s') \in M_i(s)$ for all i. This implies $M_i(s') \subseteq M_i(s)$. To see this, suppose to the contrary that there is $E \in M_i(s')$ with $E \notin M_i(s)$. By consistency, $s \notin E$. But then $s \notin \bar{E}_i(s')$ so we cannot have $\bar{E}_i(s') \in M_i(s)$, a contradiction.

So $M_i(s') \subseteq M_i(s)$ for all i. Hence measurability implies that there must be some i for whom this inclusion is strict. Obviously, then, this agent could deviate to the same c_i but to evidence $E_i \in M_i(s) \setminus M_i(s')$. Since $E_i \notin M_i(s')$, we have $s' \notin E_i$. This would not change $\hat{a}$, but would yield i a transfer of $\varepsilon/2$ and hence make him better off, a contradiction. ▌

D Proofs of Robustness Results

D.1 Proof of Theorem 4

The proof parallels that of Theorem 2. The mechanism is exactly the same as the one used there.

Fix an environment with partial information Ψ^*. Fix any $s \in S$. As in the proof of Theorem 2, it is easy to see that there is a Bayes–Nash equilibrium with outcome $f(s)$ for all θ such that $(s, \theta) \in \Omega$. Specifically, take the strategy for every agent i to be $\sigma_i(\theta_i) = ((s, f(s), 0, 0), \bar{E}_i(s))$ for every $\theta_i \in \Theta_i$. Since no feasible unilateral deviation can change the outcome, these strategies form a Bayes–Nash equilibrium with outcome $f(s)$ in every full state $(s, \theta) \in \Omega$.

The proof of Theorem 2 showed that the original mechanism had no pure or mixed Nash equilibria with an outcome different from $f(s)$. Since no agent's utility is affected by any other agent's type, the effect of uncertainty about other agents' types is the same as allowing mixing. Because of this, it is easy to adapt that proof to show that there is no pure or mixed Bayes–Nash equilibrium whose outcome differs from $f(s)$ in any full state $(s, \theta) \in \Omega$. In particular, we replace H^* in that proof with the set of pure strategy profiles played with positive probability given some $(s, \theta) \in \Omega$ such that the outcome is determined by the integers. Then the proof only has to be modified to clarify which type of a given player deviates to prevent alternative strategy profiles from forming an equilibrium. In all cases, any type who gives positive probability to the (s, θ) in question suffices.

For example, consider the part of the proof of Theorem 2 which establishes that H^* has zero probability in equilibrium. To extend this, fix some $(s^*, \theta^*) \in \Omega$ for which there is positive probability of a profile of pure strategies for which the integers determine the outcome. Let i be any agent who does not say the highest integer given this realization of the pure strategy profile and given (s^*, θ^*). Then i must give probability strictly less than 1 to the event that $z_i > z_j$ for all $j \neq i$ given s^* and θ_i^* and given that the integers determine the outcome.[16] Clearly, then, i would gain by changing his strategy when he is type θ_i^* to a larger integer (possibly also changing $\hat{a}$). Hence there is no such equilibrium. The other parts are extended analogously. ▌

D.2 Proof of Theorem 5

Let Ψ^* be a monetary environment with partial information and fix any $s \in S$. Let Γ denote the mechanism defined in the proof of Theorem 1 and let $\Gamma^*(s)$ denote the sequential game of incomplete information induced by Γ and Ψ^* at state s.

Fix any profile of types $\bar{\theta}$ such that $(s, \bar{\theta}) \in \Omega$ and let $\bar{\Gamma}(s)$ denote the complete information game defined by Γ in the environment where it is common knowledge that the utility functions are given by $v_i(\cdot, s, \bar{\theta}_i)$ for each i. We now show that for every θ with $(s, \theta) \in \Omega$, the equilibrium outcome of $\Gamma^*(s)$ is the same as the equilibrium outcome of $\bar{\Gamma}(s)$.

To be more precise, note that any history of actions in the mechanism corresponds to a collection of information sets in $\Gamma^*(s)$, one information set for each type of the player whose turn it is to move at this history. Thus, in general, there is a set of outcomes following a given history, one for each profile of types. We establish that for any equilibrium, for any history, there is a unique equilibrium outcome conditional on reaching that history, so the outcome is independent of the types of the players. In

[16]This statement uses both our common support assumption and the assumption that each Θ_i is at most countable. Together, this ensures that $\mu_i(s^*, \theta_i^*)(\theta_{-i}^*) > 0$.

particular, this unique outcome is the same as the equilibrium outcome conditional on that sequence of actions in the game $\bar{\Gamma}(s)$.

The proof is by induction on the stages of the mechanism. So first consider any history of actions which puts us in stage $I+1$ where it is agent 1's turn to present evidence. If some previous agent either proved 1 lied or claimed that 1 did so, then 1's action at this point has no effect on the outcome. In this case, $\Gamma^*(s)$ has a unique equilibrium outcome conditional on this history and the outcome is the same as the unique equilibrium outcome conditional on this history in $\bar{\Gamma}(s)$.

So suppose 1's action can have an effect. In this case, 1's action can affect his transfer but cannot affect the $\hat{a}$ which is chosen. Furthermore, in $\Gamma^*(s)$, his action affects his transfer, but his type and the types of the other players do not. Note that the types of the players do not affect the way transfers affect 1's utility or his set of feasible actions. Furthermore, he cannot be indifferent between two distinct outcomes he can induce (since they must have different transfers for him), so he cannot break ties in a manner which depends on his type or the particular equilibrium being considered. Hence 1's sequentially rational choice, and thus the outcome conditional on this history, is unique and independent of the types of the players. It is easy to see that this unique choice is also the unique sequentially rational choice on this history in $\bar{\Gamma}(s)$.

So consider any history of actions h putting us in stage k for some $k = 2, \ldots, I$ and suppose we have proved the claim for all later stages. Consider player k who makes a claim and sends evidence at this stage. Again, for some histories, k's action has no effect and so the outcome in $\Gamma^*(s)$ conditional on reaching this history is independent of k's type and must be the same as the outcome in $\bar{\Gamma}(s)$ conditional on reaching this history. For any history where k's claim and evidence can have an effect, k can only affect the transfers, not the $\hat{a}$ choice. By the induction hypothesis, the types of players moving in later stages do not affect the outcome, so k's beliefs about the types of others are irrelevant. Thus, just as above, k's sequentially rational choice in $\Gamma^*(s)$ is unique, does not depend on his type, and is equal to his choice in $\bar{\Gamma}(s)$ on this history. Hence the outcome conditional on any history putting us in any stage from stage 2 onward is the same in $\Gamma^*(s)$ and $\bar{\Gamma}(s)$.

Finally, consider player 1 in stage 1 of $\Gamma^*(s)$. From the induction hypothesis, the outcome conditional on any choice by player 1 is unique, does not depend on any player's type, and is the same as the outcome conditional on that choice in $\bar{\Gamma}(s)$. Hence 1's optimal choice does not depend on his beliefs about the types of the other players. It is then easy to see that the argument used in the proof of Theorem 1 applies to show that 1 will claim the true state and the outcome will be $f(s)$ in every equilibrium for every vector of types, just as in $\bar{\Gamma}(s)$. ▮

E Proof of Claim 1

Suppose by contradiction that there exists a mechanism Γ that implements f. Let $\Gamma(s_k)$ and $\Sigma_i(s_k)$, $k = 1, 2$, $i = 1, 2$ denote respectively the game that is induced by Γ at s_k and the set of strategies of player i in $\Gamma(s_k)$. Let $U_i : \Sigma_1(s) \times \Sigma_2(s) \to \mathbf{R}$ denote the payoff functions for the normal form of $\Gamma(s)$. Since Γ implements f, there is a subgame perfect equilibrium and hence a Nash equilibrium, $(\hat{\sigma}_1, \hat{\sigma}_2)$, in $\Gamma(s_2)$ with outcome $f(s_2)$. Thus $U_1(\hat{\sigma}_1, \hat{\sigma}_2) = 1$ and $U_2(\hat{\sigma}_1, \hat{\sigma}_2) = 0$. Since $(\hat{\sigma}_1, \hat{\sigma}_2)$ is a Nash equilibrium and since monetary transfers must be strictly less than 1/2, we have that for *every* $\sigma_2 \in \Sigma_2(s_2)$, the probability that $\hat{a}_1$ is selected when the profile $(\hat{\sigma}_1, \sigma_2)$ is played is strictly less than 1/2. To see this, suppose that it is not true. Then there is a strategy for player 2, say σ_2', such that when $(\hat{\sigma}_1, \sigma_2')$ is played, the probability of $\hat{a}_1$ is greater than or equal to 1/2. If player 2 deviates to this strategy, his payoff is strictly positive since he pays a fine of strictly less than 1/2. Hence $(\hat{\sigma}_1, \hat{\sigma}_2)$ is not a Nash equilibrium in s_2, a contradiction. Given this, we must have $U_1(\hat{\sigma}_1, \sigma_2) > 0$ for every $\sigma_2 \in \Sigma_2(s_2)$ since the alternative $\hat{a}_2$ is selected with probability at least 1/2 and player 1 pays a fine strictly less than 1/2.

Consider now the game $\Gamma(s_1)$. Since $\hat{\sigma}_1 \in \Sigma_1(s_1)$ and since $\Sigma_2(s_1) = \Sigma_2(s_2)$, player 1 can guarantee a strictly positive payoff by playing $\hat{\sigma}_1$. It follows that in every Nash equilibrium in $\Gamma(s_1)$, player 1 obtains a strictly positive payoff. Hence there is no Nash equilibrium (let alone a subgame perfect equilibrium) in $\Gamma(s_1)$ with outcome $f(s_1)$. Hence Γ cannot implement f, a contradiction. ▌

F Proof of Claim 2

Fix any perfect information mechanism Γ such that if $g(h) = (\hat{a}, t_1, t_2, t_3)$ for some terminal history h, then $|t_i| < 1/2$, $i = 1, 2, 3$. We show that this game cannot implement f. The proof works by showing that if there is a subgame perfect equilibrium of $\Gamma(s_2)$ with outcome $\hat{a}_2$ and some vector of transfers t, then every subgame perfect equilibrium of $\Gamma(s_1)$ has outcome $\hat{a}_2$ plus some transfers $\hat{t}$. The proof is by induction on the depth of $\Gamma(s_2)$.

So suppose $\Gamma(s_2)$ has a subgame perfect equilibrium with outcome $(\hat{a}_2, t_1, t_2, t_3)$.

First, suppose the depth of $\Gamma(s_2)$ is 1. Suppose player 2 is the only one who moves. Then it must be true that given any choice he makes, the outcome is $\hat{a}_2$ and some vector of transfers $\hat{t}$. To see this, suppose that there is some choice which leads to outcome $\hat{a}_1$ and some vector of transfers $\bar{t}$. Player 2's payoff from this choice is $1 + \bar{t}_2$. Since the equilibrium outcome is $(\hat{a}_2, t_1, t_2, t_3)$, we must have

$$1 + \bar{t}_2 \leq t_2$$

or $t_2 - \bar{t}_2 \geq 1$. Hence

$$|t_2| + |\bar{t}_2| \geq |t_2 - \bar{t}_2| \geq 1,$$

so

$$\max\{|t_2|, |\bar{t}_2|\} \geq \frac{1}{2},$$

contradicting our bound on transfers. Hence, as asserted, every choice available to 2 in $\Gamma(s_2)$ leads to $\hat{a}_2$ and some vector of transfers.

Given this, consider state s_1. Since player 2 has no proof available, his strategy set in $\Gamma(s_1)$ is the same as his strategy set in $\Gamma(s_2)$. So in state s_1, the equilibrium outcome must be $\hat{a}_2$ with some vector of transfers. Obviously, a symmetric argument applies to player 3.

So now suppose it is player 1 who moves in this game. Then there must be some cheap talk message available to player 1 which yields outcome $\hat{a}_2$ and some vector of transfers t. So consider the game $\Gamma(s_1)$. Since player 1 has proof, his set of strategies is larger in this game. Also, it is possible that the game now has larger depth, since presentation of proof might lead to a subgame. However, it must still be true that player 1 has available a message which leads to outcome $\hat{a}_2$ and transfers t. Suppose that in $\Gamma(s_1)$, there is an equilibrium with outcome $\hat{a}_1$ and some vector of transfers $\bar{t}$. Then player 1 must prefer this outcome to $\hat{a}_2$ with transfers t. Similar reasoning to the above shows that this implies

$$\max\{|t_1|, |\bar{t}_1|\} \geq \frac{1}{2},$$

again violating our bound on transfers. Hence any subgame perfect equilibrium in state s_1 must have outcome $\hat{a}_2$ with some vector of transfers, as asserted.

Now consider the induction step. So suppose we have proved the claim for all mechanisms such that the depth of $\Gamma(s_2)$ is less than or equal to $k-1$ and consider a mechanism such that $\Gamma(s_2)$ has depth k. Suppose player 2 moves first. An argument similar to the one above for the depth 1 case shows that if a subgame perfect equilibrium has outcome $\hat{a}_2$ and some vector of transfers t, then it must be true that each choice available to player 2 leads to a subgame in which every subgame perfect equilibrium has outcome $\hat{a}_2$ and some vector of transfers. By the induction hypothesis, this remains true in state s_1, so every initial choice available to player 2 in $\Gamma(s_1)$ leads to a subgame in which every subgame perfect equilibrium has an outcome of $\hat{a}_2$ and some vector of transfers. It readily follows that every subgame perfect equilibrium of $\Gamma(s_1)$ has such an outcome. Thus we have proved the claim for this mechanism if player 2 moves first. A similar argument applies if player 3 moves first.

So suppose player 1 moves first. Since there is a subgame perfect equilibrium with outcome $\hat{a}_2$ and transfers t in $\Gamma(s_2)$, player 1 has a cheap talk message available that leads to a subgame with a subgame perfect equilibrium that has this outcome in state s_2. By the induction hypothesis, every subgame perfect equilibrium of this subgame will have an outcome of $\hat{a}_2$ and some vector of transfers $\hat{t}$ in state s_1. Using the same reasoning as in the depth 1 case, the bound on transfers implies that this is better for player 1 than any feasible outcome with $\hat{a} = \hat{a}_1$. Hence in state s_1, 1's optimal choice must lead to an outcome with $\hat{a} = \hat{a}_2$. ▌

G Proof of Claim 3

Fix any multistage mechanism Γ such that for *every* history h, $g(h) = (\hat{a}, 0, \ldots, 0)$ for some $\hat{a} \in \hat{A}$. That is, there are no transfers on any history. Let $\Sigma_i(s_k)$ denote the set of strategies for player i in $\Gamma(s_k)$. We show that Γ cannot implement f. The proof works by showing that if $\Gamma(s_2)$ has a subgame perfect equilibrium with outcome $\hat{a}_2$, then $\Gamma(s_1)$ has a subgame perfect equilibrium with outcome $\hat{a}_2$ as well. The proof is by induction on the depth of the game $\Gamma(s_2)$.

First, suppose that the depth of $\Gamma(s_2)$ is 1. Let $\sigma = (\sigma_1, \ldots, \sigma_n)$ be a subgame perfect equilibrium of $\Gamma(s_2)$ with outcome $\hat{a}_2$. We claim that $\hat{a}_2$ is a subgame perfect equilibrium outcome of $\Gamma(s_1)$ as well. To see this, define a profile of strategies $\hat{\sigma}$ in $\Gamma(s_1)$ as follows. At the first stage of the game, each player i plays the action σ_i. (Note that $\Sigma_i(s_2) \subseteq \Sigma_i(s_1)$ for all i, so this is feasible.) Let $b = (b_1, .., b_n)$ be some profile of actions that is played in the first stage. If b is a profile of actions which is feasible in $\Gamma(s_2)$, the game must terminate when b is played since the depth of $\Gamma(s_2)$ is 1. So suppose $\Gamma(s_1)$ does not terminate when b is played. Then b must not be feasible in s_2, implying that player 1's action, b_1, involved presentation of a proof that the state is s_1. Let $\Gamma^b(s_1)$ denote the subgame that is the continuation of $\Gamma(s_1)$ after the play of b and let τ_b be any subgame perfect equilibrium of $\Gamma^b(s_1)$. For any history h within the subgame $\Gamma^b(s_1)$, let $\hat{\sigma}(b, h) = \tau(h)$.

We now show that $\hat{\sigma}$ is a subgame perfect equilibrium of $\Gamma(s_1)$. Obviously, the outcome when $\hat{\sigma}$ is played is $\hat{a}_2$, so this will establish our claim for depth 1. Since the restriction of $\hat{\sigma}$ to every subgame of $\Gamma(s_1)$ that starts at the second stage of the game is a subgame perfect equilibrium of the subgame, the only thing we need to show is that no player i can gain from a unilateral deviation in the first stage of the game.

To see that this holds, note that player 1 gets his best possible outcome, $\hat{a}_2$, when $\hat{\sigma}$ is played so clearly he cannot gain from deviating. For any player $i \neq 1$, any action $b_i \neq \sigma_i$ that he can play at the first stage of $\Gamma(s_1)$ is an action that he can play in the game

$\Gamma(s_2)$ as well. Hence the fact that $\Gamma(s_2)$ has depth 1 implies that any such action leads to the termination of the game $\Gamma(s_1)$ with an outcome that is identical to the outcome in $\Gamma(s_2)$ when i deviates to b_i. Since σ is an equilibrium in $\Gamma(s_2)$, we conclude that player i cannot gain from deviating. Hence $\hat{\sigma}$ is a subgame perfect equilibrium of $\Gamma(s_1)$ with outcome $\hat{a}_2$.

So suppose that the claim has been proved for every mechanism Γ' such that the depth of the game $\Gamma'(s_2)$ is smaller than K. Let Γ be a mechanism such that the depth of $\Gamma(s_2)$ is K and let σ be a subgame perfect equilibrium of $\Gamma(s_2)$ with outcome $\hat{a}_2$. We will show that $\Gamma(s_1)$ has a subgame perfect equilibrium $\hat{\sigma}$ with outcome $\hat{a}_2$.

For each profile of actions b which can be played in the first stage of $\Gamma(s_2)$ and does not terminate the game, let Γ^b denote the game form that is the continuation of Γ following a play of b in the first stage. Γ^b is a mechanism such that $\Gamma^b(s_2)$ has depth less than K. Let σ^b denote the profile of strategies in $\Gamma^b(s_2)$ that is induced by σ. Since σ is a subgame perfect equilibrium of $\Gamma(s_2)$, σ^b is a subgame perfect equilibrium of $\Gamma^b(s_2)$. If the outcome under σ^b is $\hat{a}_2$, then the induction hypothesis implies that $\Gamma^b(s_1)$ has a subgame perfect equilibrium, say $\tilde{\sigma}^b$, with outcome $\hat{a}_2$.

With this in mind, we construct a subgame perfect equilibrium $\hat{\sigma}$ of $\Gamma(s_1)$ with outcome $\hat{a}_2$ as follows. In the first stage of the game, each player i chooses his action according to σ_i. Let b be a profile of first stage actions which does not terminate the game. If b is feasible in state s_2 and if the outcome of $\Gamma^b(s_2)$ under σ^b is $\hat{a}_2$, then let $\hat{\sigma}$ on $\Gamma^b(s_1)$ be the subgame perfect equilibrium $\tilde{\sigma}^b$. For any other b, choose any subgame perfect continuation strategies.

To see that the outcome under $\hat{\sigma}$ is $\hat{a}_2$, note that the first stage strategies under $\hat{\sigma}$ are the same as those under σ. Hence the resulting profile of actions b must either terminate the game with outcome $\hat{a}_2$ or lead to a subgame where the continuation equilibrium has outcome $\hat{a}_2$. To see that $\hat{\sigma}$ is a subgame perfect equilibrium of $\Gamma(s_1)$, note that the restriction of $\hat{\sigma}$ to every proper subgame in $\Gamma(s_1)$ is subgame perfect so we just need to show that no player i can gain from a deviation at the first stage of the game. The outcome $\hat{a}_2$ is the best outcome for player 1 so, clearly, he will not gain by deviating.

So consider a deviation in the first stage by any player $i \neq 1$ to an action b_i which has zero probability under $\hat{\sigma}_i$. Suppose this deviation leads to outcome $\hat{a}_1$ with positive probability. Note that i's deviation is also feasible in s_2 and that the players other than i are playing the same actions as they played in s_2 under σ. Hence i's deviation would lead to a positive probability of outcome $\hat{a}_1$ under σ in $\Gamma(s_2)$ as well. Since i prefers $\hat{a}_1$ to $\hat{a}_2$, this contradicts σ being a subgame perfect equilibrium of $\Gamma(s_2)$. ▌

H Mixed Strategies

Our discussion has focused on implementation in pure strategy subgame perfect equilibrium. In this section, we show that all our results carry over for the case where the solution concept is mixed strategy subgame perfect equilibrium.[17] First, it is easy to see that Proposition 1 holds for mixed equilibria as well as pure. It is also easy to see that the proofs given earlier for Theorems 2 and 4 and Claims 1, 3, and 4 apply to implementation both in pure and in mixed strategies.

Finally, we turn to the extension for the results that refer to implementation in perfect information games, namely, Theorems 1, 3, and 5 and Claim 2. For Theorems 1, 3, and 5, we claim that the mechanisms which were defined in Parts A, B, and D.2 of the Appendix implement the SCF f in mixed strategies as well. To see this, consider the mechanism Γ_f which is defined in the proof of Theorem 1 (the arguments for Theorems 3 and 5 are identical). We first note that the proof of Theorem 1 establishes that for every state s, there exists a subgame perfect equilibrium in the game $\Gamma_f(s)$ with outcome $f(s)$. It is easy to see that the argument which establishes that there is no pure subgame perfect equilibrium with a different outcome than $f(s)$ also implies that there is no mixed subgame perfect equilibrium which gives positive probability to an outcome different from $f(s)$. Specifically, if player 1 sets $c_1 = s'$ where $M_i(s) = M_i(s')$ for all i, then, in equilibrium, it must be the case that $E_i = \bar{E}_i(s')$ for $i = 1, \ldots, I$ and no player accuses 1 of lying. Hence the outcome is $f(s')$ which, by measurability, equals $f(s)$. In particular, player 1 can get the outcome $f(s)$ by setting $c_1 = s$. On the other hand, if player 1 sets $c_1 = s' \neq s$ and there is some player i with $M_i(s) \neq M_i(s')$, then the argument given in the proof shows that player 1 will end up with an outcome that is inferior to $f(s)$. Hence in equilibrium, player 1 will not announce such a state s' with positive probability.

For Claim 2, it is easy to see that the proof of the claim in Part F of the Appendix establishes that if Γ is a perfect information mechanism such that there exists a *pure* subgame perfect equilibrium of $\Gamma(s_2)$ with outcome $\hat{a}_2$ plus some transfers t, then every *pure* subgame perfect equilibrium of $\Gamma(s_1)$ has outcome $\hat{a}_2$ plus some transfers $\hat{t}$. Since $\Gamma(s_2)$ is a perfect information game, it has a pure strategy equilibrium. Therefore, if Γ implements f in a subgame perfect equilibrium in mixed strategies, then there is a pure subgame perfect equilibrium in $\Gamma(s_2)$ with an outcome $\hat{a}_2$ and some vector of transfers t. Hence there is a (pure) subgame perfect equilibrium with outcome $\hat{a}_2$ and a vector of transfers $\hat{t}$ in $\Gamma(s_1)$, so f cannot be implemented.

[17]The fact that the mixed equilibria include the pure equilibria can make implementation with mixed strategies easier or more difficult. More specifically, let f be a SCF and Γ a mechanism. It is possible that Γ implements f in mixed strategies but not in pure strategies (because there may exist only a non–pure equilibrium in $\Gamma(s)$ with outcome $f(s)$) and it is possible that Γ implements f in pure strategies but not in mixed strategies (because there may exist a non–pure equilibrium which induces an outcome different than $f(s)$ at state s while all the pure equilibria have outcome $f(s)$.)

References

[1] Abreu, D., and H. Matsushima, "Virtual Implementation in Iteratively Undominated Strategies: Complete Information," *Econometrica*, **60**, September 1992, 993–1008.

[2] Alger, I., and A. Ma, "Moral Hazard, Insurance, and Some Collusion," *Journal of Economic Behavior and Organization*, **50**, February 2003, 225–247.

[3] Bull, J., and J. Watson, "Evidence Disclosure and Verifiability," *Journal of Economic Theory*, **118**, September 2004, 1–31.

[4] Bull, J., and J. Watson, "Hard Evidence and Mechanism Design," *Games and Economic Behavior*, **58**, January 2007, 75–93.

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

[6] Fishman, M. and K. Hagerty, "The Optimal Amount of Discretion to Allow in Disclosure," *Quarterly Journal of Economics*, **105**, May 1990, 427–444.

[7] Forges, F. and F. Koessler, "Communication Equilibria with Partially Verifiable Types," *Journal of Mathematical Economics*, **41**, November 2005, 793–811.

[8] Forges, F. and F. Koessler, "Long Persuasion Games," *Journal of Economic Theory*, forthcoming.

[9] Glazer, J., and A. Rubinstein, "Debates and Decisions: On a Rationale of Argumentation Rules," *Games and Economic Behavior*, **36**, August 2001, 158–173.

[10] Glazer, J., and A. Rubinstein, "On Optimal Rules of Persuasion," *Econometrica*, **72**, November 2004, 1715–1736.

[11] Glazer, J., and A. Rubinstein, "A Study in the Pragmatics of Persuasion: A Game Theoretical Approach," *Theoretical Economics*, **1**, December 2006, 395–410.

[12] Green, J., and J.-J. Laffont, "Partially Verifiable Information and Mechanism Design," *Review of Economic Studies*, **53**, 1986, 447–456.

[13] Grossman, S., "The Informational Role of Warranties and Private Disclosure about Product Quality," *Journal of Law and Economics*, **24**, 1981, 461–483.

[14] Grossman, S., and O. Hart, "Disclosure Laws and Takeover Bids," *Journal of Finance*, **35**, May 1980, 323–334.

[15] Hurwicz, L., "On Infomationally Decentralized Systems," in Radner, R., and C. B. McGuire, eds., *Decision and Organization*, Amsterdam: North Holland, 1972.

[16] Hurwicz, L., E. Maskin, and A. Postlewaite, "Feasible Implementation of Social Choice Rules when the Designer does not Know Endowments or Production Sets," in Ledyard, J., ed., *The Economics of Informational Decentralization: Complexity, Efficiency, and Stability*, Boston: Kluwer Academic Publishers, 1995, 367–433.

[17] Kartik, N., and O. Tercieux, "Implementation with Evidence: Complete Information," Columbia University working paper, 2009.

[18] Lipman, B., and D. Seppi, "Robust Inference in Communication Games with Partial Provability," *Journal of Economic Theory*, **66**, August 1995, 370–405.

[19] Maskin, E., "Nash Equilibrium and Welfare Optimality," MIT working paper, 1977.

[20] Maskin, E., "Nash Equilibrium and Welfare Optimality," *Review of Economic Studies*, **66**, January 1999, 23–38.

[21] Milgrom, P., "Good News and Bad News: Representation Theorems and Applications," *Bell Journal of Economics*, **12**, Autumn 1981, 380–391.

[22] Milgrom, P., and J. Roberts, "Relying on the Information of Interested Parties," *Rand Journal of Economics*, **17**, Spring 1986, 18–32.

[23] Moore, J., and R. Repullo, "Subgame Perfect Implementation," *Econometrica*, **56**, September 1988, 1191–1220.

[24] Okuno-Fujiwara, M., A. Postlewaite, and K. Suzumura, "Strategic Information Revelation," *Review of Economic Studies*, **57**, January 1990, 25–47.

[25] Osborne, M., and A. Rubinstein, *A Course in Game Theory*, MIT Press, 1994.

[26] Palfrey, T., and S. Srivastava, "Nash Implementation Using Undominated Strategies," *Econometrica*, **59**, March 1991, 479–501.

[27] Postlewaite, A., and D. Wettstein, "Feasible and Continuous Implementation," *Review of Economic Studies*, **56**, October 1989, 603–611.

[28] Seidmann, D., and E. Winter, "Strategic Information Transmission with Verifiable Messages," *Econometrica*, **65**, January 1997, 163–169.

[29] Sher, I., "Persuasion and Limited Communication," University of Minnesota working paper, April 2008.

[30] Shin, H. S., "The Burden of Proof in a Game of Persuasion," *Journal of Economic Theory*, **64**, October 1994, 253–264.