JOURNAL OF ECONOMIC THEORY 66, 370–405 (1995)

Robust Inference in Communication Games with Partial Provability*

BARTON L. LIPMAN

Queen's University, Department of Economics, Kingston, Ontario K7L 3N6

AND

DUANE J. SEPPI

Carnegie Mellon University, Graduate School of Industrial Administration, Pittsburgh, Pennsylvania 15213

Received October 28, 1993; revised June 28, 1994

We study sequential message-sending games with an uninformed decision maker and multiple self-interested informed agents in which the ability to prove claims is limited. We give necessary and sufficient conditions for the existence of robust inference rules—that is, rules which lead to full, correct inferences even if the decision maker has very little information about speakers' preferences or strategies. Surprisingly little provability is needed when the decision maker only knows that the speakers have conflicting preferences over his actions. Conflicting preferences guarantees that someone will have an incentive to "correct" any mistaken inference. *Journal of Economic Literature* Classification Numbers: C72, D82.

I. INTRODUCTION

Numerous economic, political, and legal activities involve efforts by one or more self-interested parties to persuade uncommitted decision makers to take certain actions. Managers issue audited earnings reports to shareholders. Lobbyists brief legislators about pending bills. Litigants hire experts

* We thank Ed Green, Debra Holt, Steve Matthews, Georg Nöldeke, Tom Palfrey, Mike Peters, John Roberts, an anonymous associate editor and referee, and seminar participants at Simon Fraser, Toronto, Western Ontario, the 1990 Carnegie Mellon Conference on Political Economy, the 1993 Canadian Economic Theory meetings, and the 1993 Winter Econometric Society meetings for helpful comments. Support for this research was provided by a Sloan Foundation grant to the Center for the Study of Public Policy at Carnegie Mellon University and by the Social Sciences and Humanities Research Council of Canada.

0022-0531/95 $12.00

to testify on the facts of a case. Conversely, decision makers typically try to maximize the information elicited from such communication, recognizing, of course, the possibility of bias due to the speakers' self-interest. Indeed, what an interested party wants the decision maker to do (e.g., retain incumbent management, vote for a bill) is often independent of the variables of interest to the decision maker (e.g., managerial quality, the bill's merits). In some situations the decision maker may not even know a priori exactly what the speakers' preferences are.

Certainly, statements which include irrefutable proof of some fact have information content regardless of the speaker's preferences. Unfortunately, how much is explicitly provable is often limited. That is, an interested party may be able to prove only some—but not all—of what he knows. We call this *partial provability*. Limitations on the ability to prove claims arise from at least two sources. First, there may be limitations on the number of facts speakers can disclose. In political debates, for example, time constraints may limit information transmission if voters are unable to absorb more than a certain amount of information in the time available. Second, definitive proof of some true facts simply may not exist or speakers may not be allowed to provide it. For example, while a pianist can easily demonstrate that he can play the piano, it is hard to imagine how a non-musician could prove that he cannot.[1] Alternatively, the exclusionary rule prohibits the introduction in trials of evidence (factual "proofs" which prosecutors do have) obtained through illegal searches.

Economic theory, in view of the perceived pervasiveness of limits on provability, has largely focused on information revelation through observation of actions and/or outcomes as opposed to written or verbal communication.[2] In contrast, we are interested solely in such direct communication. We show that full revelation of all information of interest to the decision maker does not always require speakers to prove all facts known to them. Rather, the "burden of proof" may be much weaker in that it involves proving only certain key facts. Of course, which facts are key depends on what the speakers' preferences are—that is, on which "lies" speakers would like to tell.

We focus here on the important case of *conflicting preferences* where the interested parties disagree among themselves about the relative desirability of possible actions by the decision maker. Examples include corporate

[1] We thank Mike Peters for this example. In terms of the analysis that follows, this limitation is a consequence of the fact that everything a nonmusician can do at the keyboard is a subset of what a pianist can do.

[2] Two important exceptions are the literature on the role of cheap talk (Farrell [5]) and the literature on political campaigns and lobbying (for example, Austin-Smith and Wright [1], Banks [2], Harrington [10]). These papers focus on issues other than the role of provability.

proxy battles, rival advertising campaigns, dispute mediation, and congressional hearings. In such situations we show that the associated burden of proof for full revelation is surprisingly weak. Moreover, the decision maker's inferences are robust in the sense that they require very little knowledge about the speakers' preferences or strategies.

Thus it is important to include direct communication in economic models because even minimal provability may radically affect predicted outcomes. The following example illustrates this point.

EXAMPLE 1. Suppose there are a large number of possible "states of nature"—say 1000—and that in each state the decision maker (if he knew the true state) would take a different action. Initially the decision maker is uninformed, but he receives advice from two lobbyists who each know the true state. He knows that the lobbyists disagree about the relative desirability of any pair of possible actions, but nothing else. Lobbyist 1 speaks first followed by lobbyist 2 who speaks after seeing 1's message. The available messages allow a lobbyist (a) to assert unverifiably that a particular state s is true (even if it is not) and then (b) to submit a single piece of evidence ruling out any single untrue state s'. A state cannot be ruled out if it is actually true. We call these "not" messages since they unambiguously prove only that a single s' is not the true state. Thus, a not message is the minimally informative message in terms of what it explicitly proves.

Taken together the lobbyists' two messages rule out at most only two of the thousand states directly. However, despite the limited informativeness of not messages, there is an inference rule which supports a perfectly revealing equilibrium! It is simply to believe lobbyist 1's assertion unless it is disproven by lobbyist 2. In this case, believe lobbyist 2's assertion, so long as it is not disproven by either of the two messages. Clearly lobbyist 2, if he can disprove 1's assertion, will do so. Furthermore, since he will make the best possible assertion for himself, the conflict between his preferences and those of lobbyist 1 guarantees that this is the worst possible assertion for lobbyist 1. Hence in equilibrium lobbyist 1 will tell the truth (except possibly in his least preferred state where lying and getting caught does not affect the final outcome).

Intuitively, complete revelation is possible because 2's message conveys more in equilibrium than simply ruling out a single state. The failure of 2 to "refute" 1's claim is taken as evidence of the inability to refute. More generally, an interested party's failure to prove certain facts may be construed to mean that these facts are untrue. While this also occurs in models with only one speaker, this example suggests that the potential informational gain with multiple speakers is substantial. This is because the interpretation of a message from one speaker can now depend on what other speakers say.[3] Thus the equilibrium interpretation of a "not state s'" message from lobbyist 2 changes dramatically depending on whether 1 initially claimed state s'.

Games with communication and no provability have been widely studied in the signaling and mechanism design literature. In this literature, agents have different preferences depending on their private information. In signaling models, these preference differences may take the form of differential "signaling costs" as in Spence [19]. Alternatively, the speaker's preferences may vary with his information in manner similar to that of the hearer as in Crawford and Sobel [3]. In the mechanism design or implementation literature (see Harris and Townsend [11], Maskin [12], Moore and Repullo [15], and Palfrey and Srivastava [17]), a "social planner" uses differences in preferences across states to get informed agents to reveal their information truthfully. As discussed above, part of our interest is in settings where the speaker's preferences are independent of his private information and/or unknown to the decision maker.

The study of games with provability[4] is relatively new. Grossman [9], Milgrom [13], and Milgrom and Roberts [14] study signaling in the special case of *complete provability*—that is, when an interested party can prove any true claim. Their insight is that a decision maker, by adopting an attitude of "scepticism in the face of vagueness," can force complete disclosure of all information in equilibrium. As a practical matter, however, this approach has two weaknesses. First, it makes strong assumptions about the set of available messages. In particular, to prove the full truth unambiguously may require exhaustive specificity in some proofs. Second, it cannot explain the prevalence of adversarial debate among multiple parties in many real-world decision-making processes (e.g., labor mediation, trials, congressional hearings). In particular, with complete provability and symmetrically informed speakers, all information can be elicited from a single speaker. Thus there is no informational gain from competition between multiple interested parties.

The idea of partial provability first appeared in Milgrom [13]. The attraction of this idea is that, while interested parties may be unable to prove much, it is often unrealistic to assume that they can prove *nothing*. The complication partial provability introduces is that some vagueness (i.e., incompleteness in proof) is unavoidable. The key insight is that these limitations can often be overcome because in equilibrium the failure to

[3] Matthews and Fertig [4] demonstrate a similar phenomenon in a Spence [19]-type signaling model.

[4] We favor the term "provability" over the more widely used "verifiability" because the former suggests that the onus is on the speaker to prove the truth of his statement, rather than on the listener to confirm it.

prove certain key supporting facts may be construed as evidence that certain claims are untrue, as in Example 1. The subsequent literature on partial provability is still small. Fishman and Hagerty [6] and Shin [18], like Milgrom [13], consider the case of a single interested party in the context of particular message structures. Okuno–Fujiwara, Postlewaite, and Suzumura [16] considers multiple asymmetrically informed interested parties with a specialized message structure. However, their assumptions effectively decompose the game into a collection of parallel single interested party problems. Finally, in quite a different vein, Green and Laffont [8] considers the role of partial provability in a principal–agent context.

In this article, we study communication in a large class of games in which interested parties speak sequentially. Our focus is on the following question. How much provability is required to achieve full and robust revelation in such games? We say that the decision maker's inference rule is robust for a set $\mathscr{P}$ of possible speaker preferences if it leads to correct inference given any equilibrium responses by the speakers to this rule and any profile of speaker preferences in $\mathscr{P}$. We find that a simple condition on the structure of provability, which we call *refutability*, is sufficient for robust full revelation when the speakers have conflicting preferences over the decision maker's action. We also provide a necessary and sufficient condition called *weak refutability* for robust full revelation in *open forums*—games where each speaker only speaks once. As in Example 1, a simple inference rule supports robust revelation. It is to provisionally believe claims unless they are subsequently refuted. However, in general the burden of proof for claims to be put "on the table" is more subtle than in Example 1. In particular, claims must be made in such a way that they can later be disproven if they are false. Refutability or weak refutability simply ensure enough provability to meet this burden in each state. Either refutability condition is weaker than complete provability or any of the particular forms of partial provability considered in the previous literature.

This article is organized as follows. Section II presents the basic model. Section III presents sufficient conditions for robust revelation in the special case of an open forum. Section IV extends this analysis to general sequential games. In Section V, we discuss some related issues. Concluding remarks are offered in Section VI.

II. The Model

We have $n+1$ players, n "senders" of information, and one "receiver." Let $N=\{1, \ldots, n\}$ denote the set of senders. The senders send messages to the receiver, who then chooses an action affecting both his own utility and the utility of each sender. The senders are symmetrically informed with information which is not known to the receiver, but which affects the receiver's payoff and possibly their own. Messages have no effect on any player's utility except through any influence they exert on the decision maker's action.[5] Thus the sending of costless messages can be interpreted as the senders' attempts to persuade the receiver to choose actions they like.

Let S denote the set of possible states of the world. A given state $s \in S$ specifies all facts known to the senders which affect the receiver's payoff.[6] For example, a state might specify the circumstances of a crime, the relative merits of various brands of a good, the talent of a firm's incumbent management, or the costs and benefits of various health care reform proposals. For simplicity, S is taken to be finite. The number of states is L.

Messages sent to the receiver can include evidence such as documents or physical items. Thus a state s also specifies the availability of such evidence.[7] Letting M denote the set of all possible messages, we have a function $M(s)$ which for each s gives that (nonempty) subset of M consisting of those messages which are available or *feasible* in state s. We also define the inverse function $F(m) = \{s \mid m \in M(s)\}$ giving the set of states in which a message m is feasible.

Our modeling approach captures at least two distinct forms of provability. First, as noted, a message may include a presentation of documents or other "hard" evidence substantiating some set of facts. Second, a message can include a logical proof that known facts imply a particular conclusion. Such a proof cannot be produced if the facts are inconsistent with this claim.[8] Either way, it is precisely the fact that a message m is available in some states and not in others that makes it useful as evidence. Thus, a message m proves that the true state is in $F(m)$ and rules out states in $S \backslash F(m)$. We refer to $F(m)$ as the *pure information content* of m. As Example 1

[5] An alternative interpretation of our model is that the cost of sending a given message is zero in a state in which that proof is available and infinite otherwise.

[6] In other words, we do not distinguish between underlying states which senders cannot distinguish. Also, note that the receiver's information is incomplete in two respects. First, he does not know some facts which directly affect his payoffs. Second, he may not know the senders' preferences. A "state" here only refers to the former, but our results all have natural analogues if we redefine "states" to include both types of uncertainty.

[7] For example, if a murder victim was stabbed and not shot, then there cannot be a bullet in the corpse. Less dramatically, a (legitimate) house deed with a given individual's name on it exists only if that person is a homeowner. Note that what a given piece of evidence proves depends on the state set S. For example, if deeds can be forged, then showing a house deed rules out all states except those in which the named individual is a homeowner and those in which the deed was forged.

[8] Implicit in this interpretation is the view that the receiver may be unaware of all implications of his prior knowledge.

illustrates, however, in equilibrium messages can convey information beyond what they prove. Thus, even a message which is in every $M(s)$—and which hence has no pure information content whatsoever—can still have significant equilibrium information content.

To avoid trivial barriers to communication (e.g., fewer messages than states), we assume throughout a *rich language condition*. Intuitively, the language is sufficiently rich so that a sender can always include in his message a "cheap talk" claim of any state that the message does not disprove. More formally, for every message m, the number of messages with the same pure information content $\#\{m' \mid F(m') = F(m)\}$ is no less than $\# F(m)$, the number of states in which m is feasible.

The receiver's payoff depends on both his action and the true state $s \in S$. Given the messages sent, the receiver updates his beliefs and then chooses an action to maximize his expected utility.[9] The receiver's posterior beliefs are an inference $\delta \in \Delta$ where Δ is the set of probability distributions over S. His prior is some $\delta^0 \in \Delta$, where $\delta^0(s) > 0$ for all $s \in S$. When there is little risk of confusion, we will also use s to denote the probability distribution which puts probability 1 on state s. Such a probability distribution is called a *degenerate inference*.

In each state s, each sender i has preferences over the possible actions by the receiver. These in turn induce preferences over possible inferences by the receiver. In other words, sender i prefers the inference δ to δ' in state s if, knowing s is the true state, he prefers the action the receiver takes given inference δ to the action taken given inference δ'.[10] Formally, the preference ordering of sender i in state s is a complete, reflexive, and transitive ordering $\succcurlyeq_{i,s}$ over the set Δ. Let $\succcurlyeq_i = (\succcurlyeq_{i,s_1}, ..., \succcurlyeq_{i,s_L})$ denote i's preferences in different states and let $\succcurlyeq = (\succcurlyeq_1, ..., \succcurlyeq_n)$ denote a *preference profile* across all n senders. It is important to note though that we do not exclude the possibility that a sender's preferences are independent of the true state, s.

We study a large class of finite extensive form games, which we call *sequential games*, in which senders never speak simultaneously. That is, every information set for every sender is a singleton. Whenever it is his turn to speak, a sender first observes the sequence of messages from previous senders and then chooses any one (feasible) message to send to the receiver. The sequence of senders may be either fixed or endogenously determined by the messages sent. Of particular interest are two special

[9] An alternative interpretation is that there are many symmetrically informed receivers who choose actions in some game after observing the senders' messages. Of course, with many receivers, more information (our focus here) could make all receivers worse off depending on the game they play.

[10] This implicitly assumes that the receiver's optimal action is unique or that his "tie-breaking" rule is common knowledge.

cases. The first is a type of sequential game we call an open forum in which each sender has exactly one turn to speak. The second, a *balanced sequential game*, generalizes the open forum to games in which each sender has at least one turn to speak.

The assumption of only one message per turn is typically restrictive given the message sets and the number of rounds of message sending in the sense that it limits how much can be proven. Indeed, the inability to communicate all feasible messages is an important form of partial provability in practice. However, the one-message-per-turn assumption in no way restricts the *class* of sequential games studied here. For example, a sequential game in which, say, two messages are sent per turn is equivalent to one in which each speaker has two successive turns to send one message. Even holding fixed the number of rounds, the two-messages-per-turn game can be recast as a one-message-per-turn game by redefining the message sets and letting each speaker send one message from redefined message sets $M'(s) = [M(s)]^2$.

We now define sequential games more precisely. For simplicity, we assume the number of rounds of message sending is some fixed K independent of the messages sent. Let $H_k(s) = [M(s)]^k$ be the set of sequences of exactly k messages feasible in state s. Let h_0 denote the initial (empty) history when the game begins (and no messages have been sent) and let $H_0(s) = \{h_0\}$ for all s. Let

$$H^K(s) = \bigcup_{k=0}^{K} H_k(s)$$

be the set of possible histories of up to K feasible messages in state s and let

$$H^K = \bigcup_{s \in S} H^K(s)$$

denote the set of all possible feasible histories. Given two histories $h, h' \in H^K$, let $h \cdot h'$ denote the sequence of messages in h followed by those in h'. Given a history $h \in H^K$, we say that $h' \in H^K$ is a *subhistory* of h if there is some $h'' \in H^K$ such that $h' \cdot h'' = h$. Also, given any $h = (m_1, ..., m_k)$, let $F(h) = \bigcap_{j=1}^{k} F(m_j)$—that is, $F(h)$ is the set of states in which the history h is possible. Using this notation, we have

DEFINITION. A *sequential game* is a pair $(K, \mathscr{I})$ where K is the number of rounds and $\mathscr{I}: H^{K-1} \rightarrow N$ specifies which sender speaks on each possible history.[11]

[11] More precisely, these are the "rules of play" for a game since payoffs and the receiver's actions are not included.

A history for sender i in state s in sequential game $(K, \mathcal{I})$ is a history $h \in H^{K-1}(s)$ such that $\mathcal{I}(h) = i$. In other words, it is simply a feasible history after which it is sender i's turn to speak. Let $H_i(s)$ denote the set of such histories for sender i in state s and let Σ_i^s denote the set of functions $\sigma_i^s: H_i(s) \to M(s)$. A *strategy* for sender i is a collection of functions $\sigma_i = \{\sigma_i^s\}_{s \in S}$, where $\sigma_i^s \in \Sigma_i^s$ for each s. We define $\sigma^s = (\sigma_1^s, ..., \sigma_n^s)$ and $\sigma = (\sigma_1, ..., \sigma_n)$. Since we focus on the existence of separating equilibria, we do not define mixed strategies.

Unlike the senders, the receiver does not observe the state itself, but only the messages sent. The set of possible histories he can observe in a game with K rounds of messages is

$$H_R = \bigcup_{s \in S} [M(s)]^K.$$

An *inference rule* for the receiver is a function $\delta: H_R \to \Delta$ giving the receiver's beliefs as a function of the observed messages. The senders' strategies for state s, σ^s, determine the particular history of messages $h_R(\sigma^s)$ he observes in s. Also, given any state s, prior history $h \in H_k(s)$, and strategies σ^s, let $h_R(h, \sigma^s)$ denote the final history the receiver would eventually observe if the play of the game began with the sequence of messages h.

Our equilibrium notion is essentially perfect Bayesian equilibrium (Fudenberg and Tirole [7]). An equilibrium of $(K, \mathcal{I})$ given a preference profile $\succcurlyeq$ is a pair (σ, δ) consisting of a vector of sender strategies σ and an inference rule δ for the receiver satisfying the following. First, for each state s and history h, every sender i's strategy is optimal given the other senders' strategies and the receiver's inference rule:

$$\delta[h_R(h, \sigma_i^s, \sigma_{-i}^s)] \succcurlyeq_{i,s} \delta[h_R(h, \hat{\sigma}_i^s, \sigma_{-i}^s)], \quad \forall \hat{\sigma}_i^s \in \Sigma_i^s,\ h \in H_i(s),\ s \in S. \tag{1}$$

Second, the receiver updates his beliefs using Bayes' Rule whenever the final history has nonzero probability under the strategies σ. Thus, if a history $h \in H_R$ is on the equilibrium path in the sense that $S^*(h, \sigma) = \{s \in S \mid h = h_R(\sigma^s)\}$ is nonempty, then

$$\delta(h)(s) = \begin{cases} \delta^0(s) / \sum_{s' \in S^*(h, \sigma)} \delta^0(s') & \text{if } s \in S^*(h, \sigma); \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

Third, off the equilibrium path (i.e., when $S^*(h, \sigma)$ is empty), inferences only need to be *consistent with feasibility* in that $\delta(h)(s) = 0$ for any state s not in $F(h)$ (i.e., for states explicitly disproven by some message m in h).

Our focus here is on separating equilibria.

DEFINITION. A *separating equilibrium* of a game $(K, \mathscr{I})$ given $\succcurlyeq$ is an equilibrium (σ, δ) in which $\delta(h_R(\sigma^s)) = s$ for all s so that the receiver always learns the true state.

As in Milgrom and Roberts [14], we are interested in situations where the receiver has very little information about the senders. In particular, we seek inference rules which are "robust" in the sense that the receiver can use them and be sure, given only minimal information about the preference profile and equilibrium strategies of the senders, that his inference is correct.

To see the intuition, fix a sequential game $(K, \mathscr{I})$, a preference profile $\succcurlyeq$, and an equilibrium (σ, δ). In each state s, the senders in effect play a game among themselves. Since (σ, δ) is an equilibrium, from (1) the senders' strategies σ^s clearly are one equilibrium of this "subgame." More formally, given δ and $\succcurlyeq$, the *induced game* for state s is the game of perfect information where sender i's strategy set is Σ_i^s and where his preferences over final histories of play $h \in [M(s)]^K$ are those induced by δ and $\succcurlyeq_{i,s}$. That is, sender i weakly prefers the history h to h' in the induced game iff $\delta(h) \succcurlyeq_{i,s} \delta(h')$. Let $E(\delta, s, K, \mathscr{I}, \succcurlyeq)$ denote the set of sender strategies which are (pure strategy subgame perfect) equilibria in the induced game.

Suppose that for some s in some equilibrium of the induced game $\hat{\sigma}^s \neq \sigma^s$, the receiver infers incorrectly given δ. For the receiver to know whether his inferences are correct, then he must know which strategies in $E(\delta, s, K, \mathscr{I}, \succcurlyeq)$ the senders are playing. However, if his inferences are correct for all equilibria in the induced game for each s, then he does not have this problem. Moreover, suppose that, given a different preference profile, there would then be an equilibrium $\hat{\sigma}^s$ in the induced game leading him to infer incorrectly. Then the receiver must be sure of the preference profile to know that his inferences are correct. A robust inference rule is precisely one which avoids these two problems.

DEFINITION. Given a sequential game $(K, \mathscr{I})$ and a set of preference profiles $\mathscr{P}$, δ is a robust inference rule for $\mathscr{P}$ if for every $\succcurlyeq \in \mathscr{P}$, there is a σ such that (σ, δ) is an equilibrium and if for every s and $\sigma^s \in E(\delta, s, K, \mathscr{I}, \succcurlyeq)$, we have $\delta(h_R(\sigma^s)) = s$.

More intuitively, if the receiver has a robust inference rule for $\mathscr{P}$, then he can know that, whatever the actual preference profile is—as long as it is in $\mathscr{P}$—and whatever equilibrium is played in the induced games among the senders, his inferences will be correct.

An obvious but useful fact about robust inference rules is the following. Suppose $\mathscr{P} \subset \mathscr{P}'$. Then if δ is robust for $\mathscr{P}'$, it must be robust for $\mathscr{P}$. Hence any condition which is *sufficient* for the existence of a robust inference rule for $\mathscr{P}'$ is also sufficient for $\mathscr{P}$, while any condition which is *necessary* for the existence of a robust rule for $\mathscr{P}$ is also necessary for $\mathscr{P}'$. For this reason, we state our necessary conditions for "small" $\mathscr{P}$ sets and our sufficient conditions for "large" sets.

Robust inference for the set of *all* preference profiles is not possible in general, even with complete provability. Consequently, the receiver must have some prior information about the senders' preferences for a robust inference rule to exist. We focus here on the case of conflicting preferences since, as noted above, adversarial debate among competing interested parties is an important feature of many real-world decision processes.

DEFINITION. Preference profile $\succcurlyeq$ satisfies conflicting preferences if for every $s' \neq s$, there exists i with $s \succ_{i,s} s'$.

Unlike the preference assumption in Example 1, this definition does not require disagreement among senders over *every* pair of inferences. Instead, the definition simply says that in state s, for any given false inference s', there is at least one sender who prefers the true inference s over s'. It is perhaps more easily interpreted when preferences also satisfy a property we call *state independence*—that is, where $\succ_{i,s} = \succ_{i,s'}$ for all s and s'.[12] With state independence and two senders, conflicting preferences implies that their preferences over degenerate beliefs disagree on every comparison. With state independence and more than two senders, conflicting preferences only says given any pair of degenerate beliefs, there is some pair of senders who disagree. Last, with or without state independence, this assumption says nothing about comparisons between nondegenerate beliefs.

Let $\mathscr{P}^*$ denote the set of conflicting preference profiles and let $\mathscr{P}_I^*$ denote the set of state independent conflicting preference profiles. Robust inference for classes of preferences as large as $\mathscr{P}^*$ or $\mathscr{P}_I^*$ is clearly still a strong property. It cannot exploit differences in preferences across particular individuals (since they are not known) or across states (since there may be none). In particular (and unlike the analysis of complete provability), the trick of "punishing" a sender for sending unhelpful (i.e., vague or otherwise incomplete) messages is not available (i.e., since the decision maker does not know which of the feasible states the sender likes least). We show in the next two sections that unexpectedly weak conditions are nonetheless sufficient for robust inference in sequential games.

Readers familiar with the literature on full implementation may find a comparison useful. In this literature (see Maskin [12], Moore and Repullo [15], Palfrey and Srivastava [17]), one seeks a game form, say (M, O), which fully implements a social choice correspondence. More precisely, $M = \prod_i M_i$ where M_i is the message set for agent i. O is an outcome function mapping M

[12] Milgrom and Roberts [14] use a similar assumption in showing that the equilibrium outcome is unique.

into some outcome space. The agents all have preferences over outcomes where these preferences depend on the state of the world. A social choice correspondence is fully implemented by this game form if, for each state, the set of equilibrium outcomes equals the social choice set for the state. By contrast, in our model the set of feasible messages (but not necessarily preferences) varies with the state. Our notion of robust inference is similar to the notion of full implementation. With robust inference, given a state and the inference rule, all equilibria for any preference profile in the class have the same outcome. With full implementation, all equilibria given any preference profile are "acceptable" in the sense that their outcomes satisfy the social choice rule. In our model, however, the receiver is also a player in the game and cannot commit himself to choosing suboptimally (e.g., taking an action which is optimal only in a state which was explicitly ruled out in the course of play). In the full implementation literature, the only receiver is the mythical social planner who can commit to any (feasible) outcome out of equilibrium.

III. Sufficient Conditions for Robust Inference in an Open Forum

In this section we give a simple condition on the structure of provability —as represented by the message sets $M(s)$—that is sufficient for robust inference for $\mathscr{P}^*$ in an open forum. This condition, in particular, is weaker than either complete provability or the forms of partial provability used in the previous literature.

While this analysis is a special case of results for sequential games in the next section, the simpler structure of an open forum allows us to present the underlying intuition more directly. An open forum is a sequential game $(n, \mathscr{I})$ in which each sender has one and only one turn to send a single message. In other words, every final history $h \in H_R$ has exactly one subhistory h' such that $\mathscr{I}(h') = i$ for each $i \in N$.

A key to our results is the robustness of a very simple and plausible type of inference rule which we call a *believe-unless-refuted* (or BUR) *rule*. Intuitively, with a BUR rule the receiver provisionally believes any claim satisfying a certain burden of proof unless it is explicitly refuted by a subsequent sender.

Definition. The inference rule δ in a K round game is a believe-unless-refuted rule if for every s and every history $h \in H_k(s)$ with $k < K$, there is a message $m_{s,h} \in M(s)$ such that $\delta(h \cdot m_{s,h} \cdot h') = s$ for every $h' \in H_{K-k-1}(s)$.

With a BUR rule, it may be that only certain messages $m_{s,h} \in M(s)$ can be used to claim s—that is, lead to $\delta(h \cdot m_{s,h} \cdot h') = s$ if h' does not refute s.

In other words, messages must meet a *burden of proof* for them to be used to claim s. In particular, as Example 1 clearly shows, any message m used to make a believe-unless-refuted claim of s must at a minimum refute any prior claim s' with BUR status—that is, we must require $m \in M(s) \setminus M(s')$. Otherwise, the receiver could have two different inferences, s' and s, on the table each with BUR status. However, simply refuting prior outstanding claims may not be enough. Determining the appropriate burden of proof for a BUR rule in general is a key issue of the article.

As this discussion suggests, a BUR rule might not exist. A BUR rule does not simply say that the receiver believes any claim until disproven; it must also be true that any false claim can be refuted and replaced by claim of the true state. For example, if *nothing* is provable—that is, if all states have the same message set—then clearly no BUR rule exists. Thus, the existence of a BUR rule depends on the structure of provability.

DEFINITION. An inference rule δ is *degenerate* if for every h, $\delta(h)$ is a degenerate inference.

PROPOSITION 1. *If δ is a degenerate BUR rule, then it is robust for $\mathscr{P}^*$ for any open forum.*

The intuition is simple. Fix any conflicting preferences $\succcurlyeq$ and any state s. Suppose that there is an equilibrium σ^s in the induced game such that the receiver does not infer s. Because δ is degenerate, it must be some single state s' that is inferred instead. By conflicting preferences, some sender i strictly prefers the true inference of s. Furthermore, because the inference rule is a BUR rule, this sender i could, at his turn, disprove whatever claim is then on the table (not necessarily s') and claim s. Since the truth cannot be refuted, s would be inferred, making i strictly better off. This, however, contradicts σ^s being an equilibrium. Thus, no equilibrium in the induced game leads to a false inference. However, the induced game, being a finite game of perfect information, does have a pure strategy equilibrium.[13] Since the receiver must infer correctly in every such equilibrium, the inference rule is robust.

How much provability is needed for a BUR rule to exist? As noted above, they do not exist in games in which nothing is provable. At the other extreme, they exist trivially if everything is provable—in which case, for each s there is a message which explicity disproves every other state and so can play the role of $m_{s,h}$ in a BUR rule. The interesting question then is just how little provability is needed.

[13] It is precisely this step which fails when trying to derive an analogous result for simultaneous message-sending games since games with simultaneous moves may, of course, not have a pure strategy equilibrium.

We begin by first defining complete provability. Strictly speaking, this could mean that literally every true statement is provable. That is, for every set $\hat{S} \subseteq S$, there is a message $m_{\hat{S}}$ such that $F(m_{\hat{S}}) = \hat{S}$. However, we will define complete provability as a weaker property; namely that for every $s \in S$, there is a message m_s which proves s—that is, $F(m_s) = \{s\}$. We use this weaker definition because it is the key to the separation proofs in Grossman [9] and Milgrom [13]. Complete provability, as defined here, is shown below to consist of two components. First, it requires *two-way disprovability*—that is,

$$s \neq s' \Rightarrow M(s) \not\subseteq M(s').$$

This means that, whenever s is true, a message $m \in M(s) \backslash M(s')$ is available disproving s' and vice versa. Thus two-way disprovability is a natural generalization of the not messages considered in Example 1. When relaxing this condition, we will generally wish to maintain a weaker condition called *one-way disprovability*:

$$s \neq s' \Rightarrow M(s) \neq M(s').$$

One-way disprovability means *either* that in state s there is a message which disproves state s' *or* vice versa (or both). Second, complete provability requires that a single summary message is available in state s which proves by itself what all messages in $M(s)$ prove jointly. We call this summary message the *full report* for state s. Formally, the *full reports condition* says that for every s, there is a message m_s^* such that

$$F(m_s^*) = \bigcap_{m \in M(s)} F(m).$$

A full report m_s^* need not prove s since if $M(s) \subset M(s')$, then s' is also in $F(m_s^*)$.

Clearly, two-way disprovability and the full reports condition are each quite strong. As in the piano player example in the Introduction, there are many situations where it is not possible to prove some fact which is true. As a more economic example, it may be difficult for an agent to prove that he has no private information of use to the receiver, as in Shin [18]. In such situations, two-way disprovability does not hold. If complete proof requires more time or space than is available, then the full reports condition does not hold. Candidate debates are a natural example. Previous models of partial provability relax one but not both of these assumptions in very specific ways. Thus the "any-k-signals" structure of Fishman and

Hagerty [6] and Milgrom [13][14] satisfies two-way disprovability but not the full reports condition. The "not-less-than" message structure of Okuno-Fujiwara *et al.* [16][15] satisfies the full reports condition and one-way disprovability but not two-way disprovability, as does the message structure studied by Shin [18].[16]

We now show that a much weaker condition than either two-way disprovability or full reports plus one-way disprovability is sufficient for existence of a BUR rule. To state it, we first define a set $S^*(s) = \{s' \neq s \mid M(s') \subseteq M(s)\}$ which we use to define

$$T(s) = M(s) \setminus \left[\bigcup_{s' \in S^*(s)} M(s') \right].$$

DEFINITION. The message sets satisfy *refutability* if for every s' and every $s \notin S^*(s')$, we have $T(s) \not\subseteq M(s')$.

To understand refutability, recall that the rich language condition allows us to interpret messages as including a cheap-talk (i.e., "proof-less") claim of a state. However, the interpretation of messages is part of the receiver's inference rule, so he is free to choose how he associates claims with messages. Hence he need not interpret every message in $M(s)$ as a claim of s, but rather only those which meet some burden of proof in that they explicitly rule out (i.e., disprove) certain "problem states."

For a BUR rule, the problematic states for a claim s are those in the set $S^*(s)$. To see why, note that a claim of s is *irrefutable* at s' if $M(s') \subset M(s)$—that is, it is impossible to disprove s when s' is true. $T(s)$, then, is the set of messages which are feasible in state s and which disprove all of the problem states $s' \in S^*(s)$. In light of the preceding discussion, such messages can be seen as *trustworthy* ways to claim state s in the sense that the receiver knows that such a claim either is true or can be disproved. Refutability then simply guarantees that a message $m \in T(s) \setminus M(s')$ is available to claim the true state s in a trustworthy way while *simultaneously* refuting any outstanding claim s', as long as s' is refutable at s.

Figure 1 gives a simple example in which refutability holds, but neither two-way disprovability (since $M(s_4) \subset M(s_1)$) nor full reports (since

[14] In an example, Milgrom [13] assumes that the sender observes N signals. The set of feasible messages is the set of truthful disclosures of at most k of these signals. In Fishman and Hagerty [6], each signal takes on one of two values, "high" or "low," and $k = 1$.

[15] They assume that a sender observes a signal from a finite, ordered set. The set of messages is the set of truthful lower bounds for the signal.

[16] Shin [18] assumes that the sender observes some randomly chosen set of signals. He can prove what values the observed signals took on, but cannot prove which signals were *not* observed.

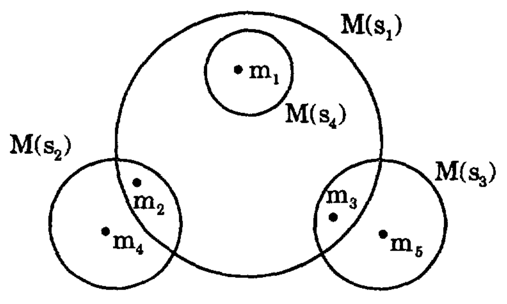

FIG. 1. Refutability without two-way disprovability or full reports.

sending all messages in $M(s_1)$ proves s_1, but no one message does) holds. More generally, the message properties introduced thus far are related by

PROPOSITION 2. *Complete provability holds iff two-way disprovability and the full reports condition hold. Two-way disprovability implies refutability. One-way disprovability and the full reports condition imply refutability.*

Our next result confirms that refutablity does indeed ensure enough provability for BUR rules to exist.

PROPOSITION 3. *A degenerate BUR rule exists if the message sets satisfy refutability.*

Intuitively, suppose that the first sender puts a state s on the table using a message in $T(s)$. His claim then stays on the table until refuted by some subsequent sender using a message $m \in T(s')\backslash M(s)$ to make a new trustworthy claim s'. The game then continues in this fashion with whatever claim is left on the table after the last speaker's turn being the receiver's final inference. Refutability guarantees that senders can always replace any false trustworthy claim on the table at their turn with a trustworthy claim of the true state and know that their claim cannot itself be subsequently replaced by yet another claim.

Summarizing then, refutability is sufficient for the existence of a degenerate BUR rule and hence for robust inference for $\mathscr{P}^*$ (or for any subset like $\mathscr{P}_\Gamma^*$) in an open forum. In equilibrium, the receiver is able to make an inference of s and be confident that it is true—even though typically he will not have seen conclusive proof to this effect (i.e., $F(h) \neq \{s\}$). He can do this because refutability allows him to establish trustworthiness as a burden of proof for making claims. This in turn guarantees that if his inference were to be incorrect, each sender would be able to correct him. Finally, conflicting preferences guarantees that at least one sender has an incentive to do so.

EXAMPLE 2. Trials provide a particularly nice illustration of our model. In Fig. 2, the accused is guilty in state s_2 (in which he was videotaped

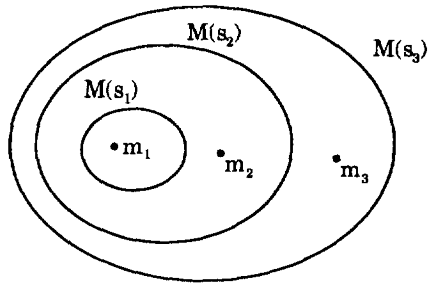

FIG. 2. Message space for Example 2.

leaving the scene of the crime) and innocent in states s_1 (in which he was far away walking alone in the woods) and s_3 (in which he was also videotaped, but a passerby saw someone else commit the crime).[17] Full reports are usually possible at trials (since each side is free to submit any and all material evidence they wish), but two-way disprovability may fail, as in this example. In this case, m_1 is simply proof-less testimony from the accused that he was "alone in the woods." Message m_2 is the videotape while m_3 is the videotape together with testimony from the passerby. Suppose the prosecutor speaks first and then the defense. Then there is a BUR rule which has a natural interpretation as "innocent until proven guilty beyond a reasonable doubt." If the prosecutor's case consists only of the message m_1, he has proven nothing. Under this rule, the inference is innocence, which can be interpreted as "innocent until proven guilty." If the prosecutor's case consists of m_2, then the BUR rule requires a verdict of guilt unless the defense presents m_3. Note that m_2 does not *prove* guilt, only guilt "beyond a reasonable doubt" in the sense that if the accused is innocent, he can establish this. The jury relies on conflicting preferences (given our adversarial system of justice) to provide the defense attorney with the proper incentives to bring forward any extenuating evidence needed to refute a false charge.

IV. Necessary and Sufficient Conditions for Robust Inference in Sequential Games

In this section, we generalize the results in the previous section in two ways. First, we provide necessary conditions to clarify the role of the various assumptions made there. Second, we consider sequential games

[17] While symmetric information is not implausible in criminal cases (where concerns about abuse of prosecutorial power are an important part of the U.S. legal tradition), it is even more descriptive of civil cases in which both parties usually know precisely what transpired between them.

in general, rather than just the open forum. Our first result is perhaps unsurprising.

PROPOSITION 4. *In any sequential game* $(K, \mathcal{I})$, *one-way disprovability is necessary for the existence of a robust inference rule for any class of preferences* $\mathcal{P}$ *containing at least one state independent preference profile and, in particular, for* $\mathcal{P}^*$.

Perhaps more surprising is that one-way disprovability is not necessary for separation *per se*. As a trivial example, suppose one sender is indifferent over all possible inferences by the receiver. Clearly, there is an equilibrium in which he tells the truth regardless of whether one-way disprovability holds. However, holding the receiver's inference rule fixed, this sender in general has many best replies. If some lead to incorrect inferences, then the inference rule is not robust. Later, we will also show that if the game itself is a choice variable, then one-way disprovability alone is sufficient for the existence of a game with a robust inference rule for $\mathcal{P}^*$.

Our next result is a characterization of robust inference rules for sequential games. This result is the key to much of our analysis.

DEFINITION. δ is *forceable* in $(K, \mathcal{I})$ if for every sender i, every state s, and every $\sigma^s_{\sim i} \in \Sigma^s_{\sim i}$, there exists $\sigma^s_i \in \Sigma^s_i$ such that $\delta(h_R(\sigma^s_i, \sigma^s_{\sim i})) = s$.

In other words, a forceable inference rule is one which allows any sender to "force" the inference to be correct given any strategies by the other senders.

PROPOSITION 5. δ *is a robust inference rule for* $\mathcal{P}^*_I$ *for a sequential game* $(K, \mathcal{I})$ *only if it is forceable. Furthermore, if a forceable* δ *exists for* $(K, \mathcal{I})$, *then a forceable, degenerate* δ *exists. Any forceable, degenerate* δ *is a robust inference rule for* $\mathcal{P}^*$ *for* $(K, \mathcal{I})$.

The intuition for sufficiency is much the same as in Proposition 1. Consider any forceable, degenerate δ together with conflicting preferences. Fix any equilibrium of the induced game in any state s. If the inference is not s, it must be some other degenerate inference s'. But then conflicting preferences guarantees that some sender i prefers that the receiver infer s. Because δ is forceable, there must be an alternative strategy for this sender which makes the inference s. Hence the original strategies could not have been an equilibrium. Thus a forceable, degenerate rule must be robust for $\mathcal{P}^*$.

To see why forceability is necessary, suppose it does not hold. Intuitively, the other senders may then be able to "gang up" on sender i and prevent him from making s the inference even when s is the true state. Furthermore, there are preference profiles in $\mathscr{P}_I^*$ such that they would do so. Hence an inference rule which is not forceable cannot be robust for $\mathscr{P}_I^*$ or $\mathscr{P}^*$. One particular way in which this could occur would be if some senders can be prevented from speaking altogether. Thus, the only games which permit robust inference are those which are *balanced* in the sense of:

DEFINITION. A sequential game $(K, \mathscr{I})$ is *balanced* if for every s, every $h \in [M(s)]^K$ such that $F(h)$ is not a singleton, and every sender i, there is a subhistory h' of h such that $\mathscr{I}(h') = i$.

That is, a game is balanced if every sender gets at least one chance to speak, regardless of what messages the other senders use.

COROLLARY 1. *If a robust inference rule exists for $\mathscr{P}_I^*$ for $(K, \mathscr{I})$, then $(K, \mathscr{I})$ is balanced.*

Balanced games are a natural generalization of the open forums studied in Section III. There is also, as one might imagine, a close connection between forceable inference rules and BUR rules. In the special case of an open forum, these notions are equivalent.

PROPOSITION 6. *If δ is a forceable inference rule for an open forum, then it is a BUR rule. If δ is a BUR rule, then it is forceable for any balanced sequential game.*

Hence Propositions 5 and 6 provide the following restatement of Proposition 1 along with a partial converse:

COROLLARY 2. *If δ is a degenerate BUR rule, then it is a robust inference rule for $\mathscr{P}^*$ for any open forum. If δ is a robust inference rule for $\mathscr{P}^*$ for an open forum, then it must be a BUR rule.*

Note that Corollary 2 is a partial converse because it only establishes that a robust inference rule for an open forum must be a BUR rule, not that it must be degenerate.

A forceable rule is not necessarily a BUR rule in balanced sequential games other than open forums for two reasons. First, in a balanced game a sender might send several messages without "yielding the floor." Second, a sender may get nonconsecutive chances to speak. For example, after initially claiming a state he may later be asked to "back up" his claim if subsequently challenged by another sender. In such a game, robust inference may well involve disbelieving the initial claim—even if it is never explicitly refuted—if later the sender does not provide this back-up evidence.

Although refutability ensures existence of a degenerate BUR rule, it can be relaxed to a condition which—if less straightforward to check—is both necessary and sufficient for such a rule to exist. To present this condition we use the following construction. For any s and any h with $s \in F(h)$, let $\tau_1(s \mid h) = M(s)$ and then recursively define[18]

$$\tau_k(s \mid h) = \{m \in M(s) \mid \nexists s' \in F(h \cdot m),\, s' \neq s, \text{ with } \tau_{k-1}(s' \mid h \cdot m) \subseteq M(s)\}.$$

Recalling that $F(h_0) = S$, we define $\tau^*(s) = \tau_K(s \mid h_0)$.

DEFINITION. The message technology satisfies *weak refutability* in a game $(K, \mathscr{I})$ if $\tau^*(s) \neq \varnothing$ for all s.

Figure 3 gives an example in which refutability does not hold (since $T(s_1) = \{m_3\} \subset M(s_2)$), but weak refutability does (i.e., $\tau^*(s_1) = \{m_3\}$, $\tau^*(s_2) = \{m_3, m_4\}$ and $\tau^*(s_3) = \{m_1, m_2\}$ for $K \geqslant 3$).

PROPOSITION 7. *A degenerate BUR rule exists iff the message technology satisfies weak refutability.*

Intuitively, $\tau_k(s \mid h)$ describes the minimum amount of provability needed with a degenerate BUR rule to ensure that the kth speaker from the end can make a trustworthy claim of state s after a history h.[19] That is, using a message in $\tau_k(s \mid h) \backslash M(s')$, he can simultaneously refute whatever claim s' is on the table and replace it with s. To see why, note that the last sender to speak cannot "block" later challenges since there are none. Hence, aside from requiring him to refute the claim on the table in making a new claim, there is no need to restrict him to trustworthy claims or to satisfy any other burden of proof. Hence $\tau_1(s \mid h) = M(s)$. At earlier stages, however, we must prevent senders from making claims in ways which block later feasible challenges. That is, $\tau_k(s \mid h)$ includes only those messages m which rule out enough other states so that if any of the remaining states in $F(h \cdot m)$ were true, the next speaker could disprove the claim of s and claim the true state s' using a message in $\tau_{k-1}(s' \mid h \cdot m) \backslash M(s)$. This is precisely what the recursive definition of $\tau_k(s \mid h)$ requires. In short, for a BUR rule to exist, we must be able to constrain senders in this fashion and so need the $\tau^*(s)$ sets to be nonempty for all s. As we show in the Proof of Proposition 7, it then follows that all the τ_k sets are also nonempty.

It is not hard to show that if $K = 2$, then $\tau^*(s) = T(s)$. Hence we have the following corollary:

[18] We thank Debra Holt for suggesting a recursive approach to obtaining a necessary condition.

[19] This intuition suggests that $\tau_k(s \mid h)$ is only relevant for histories with k stages left to go. It is more convenient, however, to define $\tau_k(s \mid h)$ for every s and h such that $s \in F(h)$.

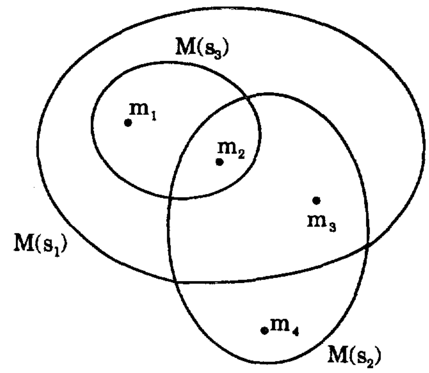

FIG. 3. Weak refutability without refutability.

COROLLARY 3. *If $K = n = 2$ and $T(s) \neq \emptyset$ for all s, then there is a robust inference rule for $\mathscr{P}^*$ for any balanced sequential game.*

We conclude by relating weak refutability to refutability. As the names suggest, the latter is stronger. Each condition enables us to meet a certain burden of proof which is sufficient to generate robust inference. The two burdens differ in two ways. First, the burden of proof that refutability allows—that is, requiring claims to be trustworthy—may require a sender to rule out certain states even though they have already been disproved by the messages of previous senders. The burden of proof associated with weak refutability avoids such redundancies. In this regard, the burden associated with refutability is stricter. Second, though, the burden associated with weak refutability may require a sender to rule out a state in order to "help out" a later sender—that is, to enable a later sender to satisfy his burden of proof.[20] However, when refutability holds, this "help" is never needed. Thus whenever the burden of proof associated with refutability can be satisfied, the burden of proof associated with weak refutability can also be satisfied. That is:

PROPOSITION 8. *Refutability implies weak refutability.*

V. COMMENTS AND EXTENSIONS

1. *Optimal Mechanisms*

Proposition 4 shows that one-way disprovability is necessary for any sequential game to have a robust inference rule for $\mathscr{P}^*$. More surprisingly,

[20] For example, this might involve requiring a sender who claims s to rule out s' even though $M(s') \not\subset M(s)$ because otherwise a later challenge to s'' would be blocked because $M(s') \subset M(s'')$ and $M(s'')\backslash M(s') \subset M(s)$.

it is also true that given any message structure satisfying one-way disprovablity, there is always *some* game $(K, \mathscr{I})$ for which a robust inference rule for $\mathscr{P}^*$ exists. Thus if the game itself is a choice variable (as may be true for congressional hearings, trials, etc.) and if conflicting preferences and one-way disprovability hold, a decision maker can ensure robust inference.

We can also give an upper bound on the number of stages needed for robust inference. Let

$$\mathscr{L} = \max_s \# \{s' \neq s \mid M(s') \subseteq M(s)\}.$$

PROPOSITION 9. *If one-way disprovability holds, there is a sequential game with* $n(\mathscr{L}+1)$ *stages for which a robust inference rule for* $\mathscr{P}^*$ *exists.*

To see the intuition, consider the following game. First, sender 1 is allowed to send any $\mathscr{L}+1$ messages, then sender 2 sends $\mathscr{L}+1$, etc., for a total of $n(\mathscr{L}+1)$ stages. We again interpret the receiver's inference rule as if each sender attaches a cheap-talk claim of a state with his set of messages. We extend the notion of trustworthiness to collections of messages by saying that a claim of s is *weakly trustworthy* if the $\mathscr{L}+1$ messages accompanying it rule out every s' with $M(s') \subset M(s)$. Thus, if a false weakly trustworthy claim is on the table, it can clearly be refuted and the true state claimed in a weakly trustworthy fashion using no more than $\mathscr{L}+1$ messages. Finally, it is easy to show that this inference rule is forceable and degenerate, so that it must be robust for $\mathscr{P}^*$.

In essence, this construction involves satisfying refutability with vectors of messages, rather than single messages. Although these vectors may be long, we are not, however, creating a "back door" form of complete provability. Indeed, such a vector may well prove less than a full report would prove, since it only disproves "subset" states as well as any false claim on the table.

It is not difficult to derive tighter bounds on the number of stages. For example, the last sender does not really need to make his claim in a weakly trustworthy way since there are no later claims that he might block. Hence we can reduce the number of stages to $(n-1)(\mathscr{L}+1)+1$. Also, it may not always require the full $\mathscr{L}+1$ messages to refute an existing claim and make a new one in a weakly trustworthy way.

Nonetheless, the number of stages in this approach can still be quite large. Realistically, one of the most important restrictions on provability is precisely constraints on time—the receiver simply cannot listen to messages for a long period of time. Hence results with a relatively small number of stages seem to us more important. In particular, recall that Corollary 1 states that a robust inference rule exists for $\mathscr{P}^*$ only if the game is balanced. Since balancedness requires that each sender have a turn to speak, the number of rounds of message-sending K must at least equal the number of senders, n. Thus, our results on open forums, where $K = n$, can be seen as sufficient conditions for robust inference with the minimum possible number of stages to the game.

2. *Robust Inference with More Information about Preferences*

To this point, all the receiver knows about the senders' preferences is that they satisfy conflicting preferences. If more is known, weak refutability may no longer be necessary for robust inference in an open forum. For example, suppose that he also knows that the first sender's preferences satisfy an *ordered subset condition*

$$M(s) \subset M(s') \Rightarrow s \succ_{1,s} s'.$$

It is not hard to show that for any message sets, such a preference order exists. With this additional knowledge, one-way disprovability becomes the only necessary and sufficient condition on provability for robust inference. Put differently, if the receiver has this additional information, robust inference is possible with only n stages if and only if one-way disprovability is satisfied.

PROPOSITION 10. *Let $\mathscr{P}^*_{\text{OS}}$ be any set of preference profiles satisfying ordered subsets for sender* 1 *and conflicting preferences. If one-way disprovability holds, then a robust inference rule exists for $\mathscr{P}^*_{\text{OS}}$ for any balanced sequential game in which sender* 1 *speaks first.*

Consider an inference rule δ which lets sender 1 use any message in $M(s)$ to claim state s. Messages in $M(s)$ from subsequent senders are interpreted as confirming the current claim s while messages in $M(s')\backslash M(s)$ replace the current claim s with a new claim s'. Thus the only burden of proof for claims is feasibility. Whatever claim is on the table at the end of play is the receiver's inference. To see that this inference rule is robust for these preferences, fix any true state s and a profile $\succcurlyeq$ from $\mathscr{P}^*_{\text{OS}}$ and suppose there is an equilibrium in the induced game in which $s' \neq s$ is the receiver's final inference. There are only two possibilities. First, we could have $M(s) \subset M(s')$. However, by the ordered subset condition, sender 1 would have been strictly better off claiming the truth which would make s the final inference. Alternatively, we could have $M(s) \not\subset M(s')$. However, then by conflicting preferences, there is some sender i for whom $s \succ_{i,s} s'$. If this is sender 1, then he again should have claimed s instead. If i is not sender 1, then at i's turn to speak, he would want to refute the current claim on the table, say s'', and replace it with s. For this not to have happened, challenges to s'' must be blocked in the sense that $M(s) \subset M(s'')$. In this case, s'' is also the final inference (that is, $s'' = s'$) since no challenges to s'' are possible. However, this contradicts our assumption that $M(s) \not\subset M(s')$. Thus the receiver always infers correctly. More colloquially, with this inference rule the only lies by 1 which the other senders cannot refute are lies sender 1 would not want to tell.

3. *Robust Inference without Conflicting Preferences*

Robust inference with respect to *all* possible preference profiles is unachievable except under very stringent conditions. For example, if $M(s) \subset M(s')$ and all senders have s as their favorite inference in every state, then there is no separating equilibrium. To see this, suppose there were and let h be a history for which $\delta(h) = s$. Since $M(s) \subset M(s')$, the history h must also be feasible in state s'. It is not hard to see then that the only equilibrium outcome in the induced game in state s' has s as the inference.

Conflicting preferences are not, however, necessary for robust inference. If the set of preference profiles $\mathscr{P}$ does not satisfy conflicting preferences but is sufficiently small, then robust inference with respect to $\mathscr{P}$ is still possible. On the other hand, it may require strong conditions. To illustrate this, we consider the polar opposite case to conflicting preferences—the case where all senders have identical preferences. To simplify further, we restrict attention to state independent preferences.[21]

DEFINITION. The message sets satisfy *ordered provability* if we can number the states so that

$$[M(s_l)]^n \setminus \bigcup_{k>l} [M(s_k)]^n \neq \varnothing, \qquad l = 1, 2, ..., L.$$

In other words, message sets satisfy ordered provability if we can number the states so that there is an n vector of messages feasible in state s_1 which proves that s_1 is the true state, an n vector of messages feasible in s_2 which proves that either s_1 or s_2 is the true state, etc. Clearly, this is a very special structure. While it holds if provability is complete, it is not implied by either two-way disprovability or the full reports condition alone. Hence the following proposition indicates that robust inference is unlikely to be possible when all senders have identical preferences.

PROPOSITION 11. *Suppose* $\succcurlyeq \in \mathscr{P}$ *has the property that* $\succcurlyeq_{i,s} = \succcurlyeq_{j,s'}$ *for all i, j, s, and s'. Then a robust inference rule for* $\mathscr{P}$ *for an open forum exists*

[21] When all senders have the same preferences, we are effectively in the one-sender case. The structure of separating equilibria with state dependent preferences in the one-sender case and no provability is already well understood, so our focus is to emphasize what is unusual here.

only if (a) *the message sets satisfy ordered provability and* (b) *given the numbering of the states for which ordered provability holds, the profile* $\succcurlyeq$ *satisfies* $s_l \succ_{i,s} s_k$ *for all* $k > l$. *If this condition holds, if* $s' \nsucc_{i,s} s''$ *for all* i, s', *and* s'' (*with* $s' \neq s''$), *and if* $\mathscr{P} = \{\succcurlyeq\}$, *then a robust inference rule does exist.*

In short, robust inference with identical state independent preferences is possible, but only under strong conditions on the message sets and only for relatively special sets of preferences.

VI. Conclusion

This article has investigated the ability of an uninformed decision maker to elicit private information from self-interested parties. We find that with more than one speaker, conflicting preferences can lead to the revelation of a surprising amount of information—even with only very limited provability and with little information on the part of the decision maker about the speakers' preferences or strategies. Since even minimal amounts of provability may radically affect predicted outcomes, it is important to take account of communication with partial provability in models of economic and other decision making.

In many real-world settings, speakers may be asymmetrically informed. Our modeling choice of symmetric information is based on the view that the informational asymmetries between speakers and decision makers are typically more substantial than those among the speakers themselves (e.g., consider competing car salesmen vs car buyers, litigants in a law suit vs the judge/jury, rival management teams vs small shareholders). While the more general model with asymmetrically informed senders is an interesting topic for future research, the symmetrically informed sender model is likely, in our opinion, to capture much of what is important in many practical applications.

Our analysis does generalize to certain special cases of asymmetric information. In particular, our results have straightforward analogues in the case where (a) the senders' preferences are state independent and (b) the senders' message sets are the same. The latter assumption says that some senders may have more information but do not have more ability to provide proof. The case of asymmetric access to proof, with or without asymmetric information, is another intriguing area for further research.

Appendix

Proof of Proposition 1. Implied by Propositions 5 and 6. ∎

Proof of Proposition 2. First, we show that complete provability holds only if two-way disprovability and the full reports condition hold. Given complete provability, for each s there is a message $m_s \in M(s)$ such that $F(m_s) = \{s\}$. Thus for any s and s', $s \neq s'$, we have $m_s \notin M(s')$, so $M(s) \nsubseteq M(s')$ and hence two-way disprovability holds. Full reports also holds since

$$F(m_s) = \{s\} \subseteq \bigcap_{m \in M(s)} F(m).$$

To show the converse, suppose two-way disprovability and the full reports condition hold. Fix any s and let m_s^* be the full report for state s. Could we then have $s' \in F(m_s^*)$ for any state $s' \neq s$? If there is any $m' \in M(s)$ such that $m' \notin M(s')$, then

$$s' \notin \bigcap_{m \in M(s)} F(m),$$

so $s' \notin F(m_s^*)$. But by two-way disprovability, $M(s) \nsubseteq M(s')$, so such a message m' must exist. Hence $s' \notin F(m_s^*)$, so $F(m_s^*) = \{s\}$. Hence complete provability holds.

Next, we show that two-way disprovability implies refutability. Clearly, two-way disprovability implies that $T(s) = M(s)$ for all s. Hence $T(s) \subseteq M(s')$ iff $M(s) \subseteq M(s')$, precisely what two-way disprovability rules out. Hence two-way disprovability implies refutability.

Last, suppose that one-way disprovability and full reports hold. Fix two states, s and s', such that $M(s) \not\subset M(s')$. Let m_s^* be the full report for state s. Clearly, if $T(s) = M(s)$, then $T(s) \nsubseteq M(s')$. So suppose $T(s) \neq M(s)$—that is, there is some s'' such that $M(s'') \subseteq M(s)$. Since one-way disprovability implies that the inclusion is strict, there is a message $m \in M(s) \backslash M(s'')$. Since the full report must have at least as much information content as m, it must be true that $m_s^* \notin M(s'')$. Since this holds for any s'' with $M(s'') \subset M(s)$, we must have $m_s^* \in T(s)$. Also, since $M(s) \not\subset M(s')$, there must be a message $m' \in M(s) \backslash M(s')$. Again, m_s^* must prove at least as much as m', so that $m_s^* \notin M(s')$. Hence $T(s) \nsubseteq M(s')$. Hence refutability holds. ▮

Proof of Proposition 3. Implied by Propositions 7 and 8. ▮

Proof of Proposition 4. Suppose one-way disprovability fails. Specifically, suppose $s_1 \neq s_2$ but $M(s_1) = M(s_2)$. Fix any state-independent preference profile and any sequential game. Clearly, if there is no separating equilibrium, there is no robust inference rule. So suppose (σ, δ) is a separating equilibrium. Consider the strategies $\hat{\sigma}$, then, where $\hat{\sigma}_i^s = \sigma_i^s$ for every $s \neq s_2$ and $\hat{\sigma}_i^{s_2} = \sigma_i^{s_1}$. Since $M(s_1) = M(s_2)$, these strategies are feasible. Furthermore, by definition, (1) holds for every i for σ which implies that the same is true of $\hat{\sigma}$. Clearly, though, $\delta(h_R(\hat{\sigma}^{s_2})) = \delta(h_R(\sigma^{s_1})) = s_1$. Hence δ is not a robust inference rule for this game and preference profile. ∎

Proof of Proposition 5. First, suppose δ is a robust inference rule for $\mathscr{P}_I^*$ for sequential game $(K, \mathscr{I})$, but is not forceable. Then there is a state s, sender i, and $\hat{\sigma}_{\sim i}^s \in \Sigma_{\sim i}^s$ such that there is no strategy $\sigma_i^s \in \Sigma_i^s$ with $\delta(h_R(\sigma_i^s, \hat{\sigma}_{\sim i}^s)) = s$. Fix a preference profile $\succcurlyeq \in \mathscr{P}_I^*$ such that all senders other than i have identical preferences with s being the best inference for sender i and the worst for the other senders. Clearly, many such profiles exist. Fix an equilibrium in the induced game at state s, say, $(\bar{\sigma}_i^s, \bar{\sigma}_{\sim i}^s)$. Since δ is robust, we must have $\delta(h_R(\bar{\sigma}_i^s, \bar{\sigma}_{\sim i}^s)) = s$. However, the facts that all senders other than i have identical preferences and that the induced game is a finite game of perfect information imply (by backward induction) that any equilibrium of the induced game must also be an equilibrium when all senders other than i are joined together into one player. But these other senders can deviate to $\hat{\sigma}_{\sim i}^s$ and guarantee a better inference for themselves than s as i has no strategy against $\hat{\sigma}_{\sim i}^s$ which leads to an inference of s. This contradicts the assumption that $\bar{\sigma}^s$ is an equilibrium of the induced game at state s. Hence δ must be forceable for it to be robust for $\mathscr{P}_I^*$ and thus for $\mathscr{P}^*$.

Clearly, if a forceable δ exists, then we can construct a forceable, degenerate δ simply by changing δ on any history h such that $\delta(h)$ is nondegenerate to any feasible degenerate inference. This change will not interfere with forceability, since forceability only refers to the parts of δ which specify degenerate inferences. We complete the proof by showing that any forceable, degenerate δ is a robust inference rule for $\mathscr{P}^*$.

Fix any such δ and any preference profile $\succcurlyeq \in \mathscr{P}^*$. Suppose there is an equilibrium in the induced game at state s, say $\bar{\sigma}^s$ with $\delta(h_R(\bar{\sigma}^s)) \neq s$. Since δ is degenerate $\delta(h_R(\bar{\sigma}^s)) = s'$ for some state s'. By conflicting preferences, $s \succ_{i,s} s'$ for some sender i. But since δ is forceable, i has an alternative strategy, say $\hat{\sigma}_i^s \in \Sigma_i^s$ such that $\delta(h_R(\hat{\sigma}_i, \bar{\sigma}_{\sim i}^s)) = s$ which is strictly better for him. This contradicts the assumption that $\bar{\sigma}^s$ was an equilibrium in the induced game at state s. Hence there is no equilibrium in the induced game at s in which the receiver's inference differs from s. Since the induced game is a finite game of perfect information, there must be a pure strategy equilibrium. Since every such equilibrium has the receiver inferring s, we see that δ is robust for $\mathscr{P}^*$. ∎

Proof of Corollary 1. Suppose not. Then there is a sender i and a feasible history $h \in H_R$ such that $F(h)$ is not a singleton and $\mathscr{I}(h') \neq i$ for every subhistory h' of h. Let s and s' be distinct states in $F(h)$. Suppose δ is robust for $\mathscr{P}_I^*$ for $(K, \mathscr{I})$. Then by Proposition 5, δ is forceable. Let $\sigma_{\sim i}^s$

be any vector of strategies for the senders other than i which is feasible in both s and s' and produces the history h. Since both s and s' are in $F(h)$, such a strategy must exist. By forceability, there must be strategies σ_i^s and $\hat{\sigma}_i^s$ such that $\delta(h_R(\sigma_i^s, \sigma_{\sim i}^s)) = s$ and $\delta(h_R(\hat{\sigma}_i^s, \sigma_{\sim i}^s)) = s'$. But since sender i is never able to send a message when the other senders use $\sigma_{\sim i}^s$, we must have $h_R(\sigma_i^s, \sigma_{\sim i}^s) = h_R(\hat{\sigma}_i^s, \sigma_{\sim i}^s)$, a contradiction. ▮

Proof of Proposition 6. The condition that δ be a BUR rule is nothing more than the translation of forceability for an open forum, implying the first statement. As to the second, suppose δ is a BUR rule. In any balanced sequential game, either the sequence of messages proves what the state is or every sender has at least one turn to send a message. The fact that δ is a BUR rule implies that each sender can, at his turn, force the inference to be s if all other senders are restricted to strategies which are feasible in state s. This is precisely what forceability requires. ▮

The proof of Proposition 7 uses the following lemma.

LEMMA 1. *For any h and h' with $F(h) \subseteq F(h')$, for all k, and for all $s \in F(h)$, $\tau_k(s \mid h') \subseteq \tau_k(s \mid h)$. Also, for all k, h, and $s \in F(h)$, $\tau_{k+1}(s \mid h) \subseteq \tau_k(s \mid h)$.*

Proof of Lemma 1. Both statements are proved by induction. To show the first statement, note that $\tau_1(s \mid h) = \tau_1(s \mid h') = M(s)$, so the statement holds trivially for $k = 1$. So suppose we have demonstrated that the statement holds for $k < j$. We now show that it holds for $k = j$. So suppose $m \in \tau_j(s \mid h')$. Then for all $s' \in F(h')$, $s' \neq s$, we have $\tau_{j-1}(s' \mid h') \nsubseteq M(s)$. But $F(h) \subseteq F(h')$ and, by the induction hypothesis, $\tau_{j-1}(s' \mid h') \subseteq \tau_{j-1}(s' \mid h)$. Hence for every $s' \in F(h)$, $s' \neq s$, we have $\tau_{j-1}(s' \mid h) \nsubseteq M(s)$, so $m \in \tau_j(s \mid h)$. Hence $\tau_j(s \mid h') \subseteq \tau_j(s \mid h)$.

To show the second statement, note that $\tau_2(s \mid h) \subseteq M(s) = \tau_1(s \mid h)$, so that the statement holds trivially for $k = 1$. Suppose we have shown that the statement holds for $k < j$. We now show that it holds for $k = j$. Suppose $m \in \tau_{j+1}(s \mid h)$. Then for all $s' \in F(h \cdot m)$, $s' \neq s$, $\tau_j(s' \mid h \cdot m) \nsubseteq M(s)$. By the induction hypothesis, though, $\tau_j(s' \mid h \cdot m) \subseteq \tau_{j-1}(s' \mid h \cdot m)$. Hence for all $s' \in F(h \cdot m)$, $s' \neq s$, $\tau_{j-1}(s' \mid h \cdot m) \nsubseteq M(s)$. Hence $m \in \tau_j(s \mid h)$, so $\tau_{j+1}(s \mid h) \subseteq \tau_j(s \mid h)$. ▮

Proof of Proposition 7. First, suppose a BUR rule exists. Fix any feasible history h_1 up to the last message to be sent—that is, with $K-1$ messages. Suppose $F(h_1)$ is not a singleton. Then by the definition of BUR,

$$\forall s \in F(h_1), \quad \exists m_{s, h_1} \in M(s) \quad \text{such that } \delta(h_1 \cdot m_{s, h_1}) = s. \tag{A1}$$

Note also that this property holds trivially if $F(h_1)$ is a singleton.

Now consider any feasible history h_2 up to the next to last round—that is, a history with $K-2$ messages. Suppose $F(h_2)$ is not a singleton. By definition of BUR,

$$\forall s \in F(h_2), \quad \exists m_{s,h_2} \in M(s) \quad \text{such that } \delta(h_2 \cdot m_{s,h_2} \cdot m) = s, \quad \forall m \in M(s). \tag{A2}$$

As before, if $F(h_2)$ is a singleton, this property holds trivially.

We claim that for every h_2 with $K-2$ messages and every $s \in F(h_2)$, $m_{s,h_2} \in \tau_2(s \mid h_2)$. To see this, first, suppose $F(h_2)$ is a singleton. Then there are no $s' \neq s$ with $s' \in F(h_2 \cdot m)$ for any m. Hence $\tau_2(s \mid h_2) = M(s)$. Hence the requirement follows trivially from (A2). So suppose $F(h_2)$ is not a singleton. Suppose $m_{s,h_2} \notin \tau_2(s \mid h_2)$. By definition, then, there is a state $s' \in F(h_2)$ such that $m_{s,h_2} \in M(s') \subset M(s)$. Let $h'_1 = h_2 \cdot m_{s,h_2}$. Since $M(s') \subset M(s)$, we must have $m_{s',h'_1} \in M(s)$. But by (A1), $\delta(h_2 \cdot m_{s,h_2} \cdot m_{s',h'_1}) = s'$, while (A2) implies that $\delta(h_2 \cdot m_{s,h_2} \cdot m_{s',h'_1}) = s$, a contradiction.

Given this, BUR requires that for every h_2 with $K-2$ messages and every $s \in F(h_2)$, $\tau_2(s \mid h_2)$ must be nonempty. Otherwise, it is impossible to find m_{s,h_2} satisfying (A2) and allowing (A1) to be satisfied.

From here, the rest of the necessity proof is by induction. Fix any feasible history h_j with exactly $K-j$ messages, so there are j rounds left counting this turn. By BUR,

$$\forall s \in F(h_j), \quad \exists m_{s,h_j} \in M(s_j) \quad \text{such that } \delta(h_j \cdot m_{s,h_j} \cdot h') = s, \quad \forall h' \text{ with } s \in F(h'). \tag{Aj}$$

Suppose that we have shown for $j \leqslant k-1$ that for every feasible h_j with $K-j$ messages, each m_{s,h_j} satisfying (Aj) must be an element of $\tau_j(s \mid h_j)$. (This is what we have shown for $k=2$.) We now show that the same must hold for $j=k$.

Suppose that for some history h_k with $K-k$ messages, there is a state $s \in F(h_k)$ such that the m_{s,h_k} satisfying Eq. (Ak) is not an element of $\tau_k(s \mid h_k)$. Let $h'_{k-1} = h_k \cdot m_{s,h_k}$. By definition of τ_k, then, there must be some $s' \neq s$ such that $s' \in F(h'_{k-1})$ and $\tau_{k-1}(s' \mid h'_{k-1}) \subseteq M(s)$. Hence $m_{s',h'_{k-1}} \in M(s)$ (using the induction hypothesis). But then fix any h'' feasible in both $M(s)$ and $M(s')$. (Since $m_{s',h'_{k-1}}$ is feasible in both states, such a history must exist.) By Eq. (Ak), $\delta(h_k \cdot m_{s,h_k} \cdot h'') = s$, but Eq. (A$k-1$) implies that $\delta(h'_{k-1} \cdot h'') = s'$, a contradiction.

Hence for every k, every h_k with $K-k$ messages, and every $s \in F(h_k)$, the message m_{s,h_k} of Eq. (Ak) must be an element of $\tau_k(s \mid h_k)$. Hence for every k and every such h_k and s, we must have $\tau_k(s \mid h_k) \neq \varnothing$. Since this also applies for $k=K$, we must have $\tau_K(s \mid h_0)$—that is, $\tau^*(s)$—nonempty for every s. Hence weak refutability must hold.

To complete the proof, suppose $\tau^*(s) \neq \varnothing$ for all s. First, note that the second statement in Lemma 1 implies that $\tau_K(s \mid h_0) \subseteq \tau_k(s \mid h_0)$ for all k and all s. Hence, using the first statement of Lemma 1, $\tau_K(s \mid h_0) \subseteq \tau_k(s \mid h_0) \subseteq \tau_k(s \mid h)$, for all k, all h, and all $s \in F(h)$. Hence $\tau^*(s) \neq \varnothing$ for all s implies that $\tau_k(s \mid h) \neq \varnothing$ for all k, all h, and all $s \in F(h)$.

Next, suppose that $m \in \tau_k(s \mid h)$ and m' has the same pure information content as m—i.e., $F(m) = F(m')$. By definition, for all $s' \in F(h \cdot m)$, we have $\tau_{k-1}(s' \mid h \cdot m) \not\subseteq M(s)$. But since $F(m) = F(m')$, we have $F(h \cdot m) = F(h \cdot m')$. By the first statement of Lemma 1, we also have $\tau_{k-1}(s' \mid h \cdot m) = \tau_{k-1}(s' \mid h \cdot m')$. Hence $m' \in \tau_k(s \mid h)$.

We next claim that our rich language condition, $\#\{m' \mid F(m') = F(m)\} \geqslant \# F(m)$, implies that:

LEMMA 2. *For any feasible h with $K-k$ messages, there exists a set of messages $\{m_{s,h} \mid s \in F(h)\}$ such that*

$$m_{s,h} \in \tau_k(s \mid h), \quad \forall s \in F(h) \text{ and } m_{s,h} \neq m_{s',h} \text{ whenever } s \neq s'. \tag{3}$$

Furthermore, for any feasible h with $K-k$ messages and any s' such that $\tau_k(s \mid h) \not\subseteq M(s')$ for all $s \in F(h) \backslash \{s'\}$, there is a set of messages $\{m_{s,h} \mid s \in F(h) \backslash \{s'\}\}$ such that

$$m_{s,h} \in \tau_k(s \mid h) \backslash M(s'), \quad \forall s \in F(h) \backslash \{s'\} \quad \text{and}$$
$$m_{s,h} \neq m_{s',h} \text{ whenever } s \neq s'. \tag{4}$$

Proof of Lemma 2. We demonstrate both statements by the following algorithm.

Step 1. Fix any $s_1 \in F(h)$ (or $F(h) \backslash \{s'\}$ for the second statement). Since $\tau_k(s_1 \mid h)$ is nonempty (and, for the second statement $\tau_k(s_1 \mid h) \backslash M(s') \neq \varnothing$ by assumption), we can choose any message from $\tau_k(s_1 \mid h)$ (or $\tau_k(s_1 \mid h) \backslash M(s')$) and call it $m_{s_1,h}$.

Step 2. Let $\hat{S}_1$ denote the set of $s \in F(h)$, $s \neq s_1$ (and $s \neq s'$ for the second statement), such that $m_{s_1,h} \in \tau_k(s \mid h)$. Clearly, if $m_{s_1,h} \in \tau_k(s \mid h)$, we must have $m_{s_1,h} \in M(s)$. Hence by our rich language condition, there must be at least $\# \hat{S}_1$ different messages with the same pure information content as $m_{s_1,h}$. By the above argument, these messages must all be elements of $\tau_k(s \mid h)$ (and, for the second statement, not in $M(s')$) for each $s \in \hat{S}_1$. Furthermore, none of them can be elements of $\tau_k(s \mid h)$ for $s \notin \hat{S}_1$ since $m_{s_1,h}$ is not an element of any of these sets. Hence we can choose distinct messages from each of the $\tau_k(s \mid h)$ (or $\tau_k(s \mid h) \backslash M(s')$) sets for $s \in \hat{S}_1$.

Step 3. Let s_2 denote any $s \in F(h) \backslash \hat{S}_1$ which is different from s_1 (and, for the second part, different from s'). We can proceed exactly as in Step 1 to find a message $m_{s_2, h}$ and as in Step 2 to find messages $m_{s, h}$ for each state s with $m_{s_2, h} \in \tau_k(s \mid h)$ (or $\tau_k(s \mid h) \backslash M(s')$). Since none of the messages used in Steps 1 and 2 are in $\tau_k(s \mid h)$ for these states, we cannot be choosing any message for two different states.

Clearly, we can continue this procedure until we have found messages $m_{s, h}$ satisfying our requirements for any given h (and, if appropriate, s'). ∎

We use Lemma 2 to recursively construct a BUR rule. First, fix messages $\{m_{s, h_0} \mid s \in S\}$ satisfying (3). Let $H_0^* = \varnothing$ and let H_j denote the set of feasible histories with exactly j messages.

To carry the recursion forward, suppose we have defined sets $H_0^*, ..., H_{k-1}^*$ with $H_j^* \subseteq H_j$ for each j and functions $\xi_1, ..., \xi_{k-1}$ where ξ_j: $H_j^* \to S$. (Since $H_0^* = \varnothing$, there is no need to define a function ξ_0 since its domain would be empty.) Suppose also that for each history h with $k-1$ or fewer messages, we have also defined a set of messages with the following properties. First, if $h \notin H_j^*$ for any j, then we have a set of messages $\{m_{s, h} \mid s \in F(h)\}$ satisfying (3). If, instead, $h \in H_j^*$, then we have a set of messages $\{m_{s, h} \mid s \in F(h) \backslash \{\xi_j(h)\}\}$ satisfying (4). Given these objects, we define an appropriate H_k^* and ξ_k and then sets of messages for each feasible history with k messages if $k \leqslant K-1$.

For the first part, let

$$H_k^* = \{h \in H_k \mid h = h' \cdot m \text{ with } m = m_{s, h'}$$
$$\text{for some } s \text{ or } h' \in H_{k-1}^* \text{ and } m \in M(\xi_{k-1}(h'))\}.$$

Also, define ξ_k: $H_k^* \to S$ by

$$\xi_k(h' \cdot m) = \begin{cases} s, & \text{if } \quad m = m_{s, h'}; \\ \xi_{k-1}(h'), & \text{otherwise.} \end{cases}$$

Next, fix any $h \in H_k$. If $k = K$, we are done. If $k \leqslant K-1$ and $h \notin H_k^*$, choose a set of messages $\{m_{s, h} \mid s \in F(h)\}$ satisfying (3). By Lemma 2, this is possible.

To complete the argument, consider $h \in H_k^*$. We claim that there is no $s \in F(h) \backslash \{\xi_k(h)\}$ such that $\tau_{K-k}(s \mid h) \subseteq M(\xi_k(h))$. To see this, suppose the contrary, so $s \in F(h) \backslash \{\xi_k(h)\}$ but

$$\tau_{K-k}(s \mid h) \subseteq M(s')$$

where $s' = \xi_k(h)$. By construction, we can write $h = h_1 \cdot m_{s', h_1} \cdot h_2$ where $s' \in F(h)$ and $m_{s', h_1} \in \tau_{K-j-1}(s \mid h_1)$ where $j \leqslant k$ is the number of messages in h_1. Because $K-j-1 > K-k$, Lemma 1 implies that

$$\tau_{K-j-1}(s \mid h) \subseteq \tau_{K-k}(s \mid h).$$

Clearly, $F(h) \subseteq F(h_1 \cdot m_{s', h_1})$. Hence Lemma 1 also implies that

$$\tau_{K-j-1}(s \mid h_1 \cdot m_{s', h_1}) \subseteq \tau_{K-j-1}(s \mid h).$$

Hence

$$\tau_{K-j-1}(s \mid h_1 \cdot m_{s', h_1}) \subseteq M(s').$$

But this contradicts the definition of τ_{K-j-1}. Hence, as asserted, there is no $s \in F(h) \backslash \{s'\}$ with $\tau_{K-k}(s \mid h) \subseteq M(s')$. Hence by Lemma 2, we can find a set of messages $\{m_{s,h} \mid s \in F(h) \backslash \{s'\}\}$ satisfying (4).

This procedure gives us a set $H_K^* \subseteq H_K$ and a function $\xi_K\colon H_K^* \to S$. Define δ so that $\delta(h) = \xi_K(h)$ if $h \in H_K^*$. The choice of δ for $h \notin H_K^*$ is arbitrary. It is easy to see that δ is well defined and a BUR rule. ∎

Proof of Proposition 8. Suppose refutability holds. We show by induction that this implies that $T(s) \subseteq \tau_k(s \mid h)$ for all k, all s, and all h with $s \in F(h)$. First, note that this holds trivially for $k = 1$ since $\tau_1(s \mid h) = M(s)$. So suppose we have shown that this holds for all $k < j$. We now show that it holds for $k = j$. So fix any h and any $s \in F(h)$. Suppose $m \in T(s)$. For any $s' \neq s$ with $M(s') \subseteq M(s)$, we have $s' \notin F(m)$ and hence $s' \notin F(h \cdot m)$. For any s' with $M(s') \nsubseteq M(s)$, refutability implies that $T(s') \nsubseteq M(s)$. By the induction hypothesis, for every $s' \in F(h \cdot m)$, $T(s') \subseteq \tau_{j-1}(s' \mid h \cdot m)$. Hence $\tau_{j-1}(s' \mid h \cdot m) \nsubseteq M(s)$, as otherwise we would contradict refutability. Hence $m \in \tau_j(s \mid h)$, so $T(s) \subseteq \tau_j(s \mid h)$.

We complete the proof by showing that refutability implies that $T(s) \neq \varnothing$ for all s. Since $T(s) \subseteq \tau^*(s)$, this clearly implies weak refutability. So suppose refutability holds but that for some s, $T(s) = \varnothing$. Clearly, this implies that there is at least one $s' \neq s$ with $M(s') \subset M(s)$. Hence we cannot have $M(s) \subset M(s')$. Therefore, by refutability, it must be true that $T(s) \nsubseteq M(s')$. But if $T(s) = \varnothing$, this cannot hold, a contradiction. Hence $T(s) \neq \varnothing$ for all s. ∎

Proof of Proposition 9. Consider the game where $\mathscr{I}(h) = i$ for every history of length k with $(i-1)(\mathscr{L}+1) \leqslant k < i(\mathscr{L}+1)$. That is, sender 1 speaks on the first $\mathscr{L}+1$ turns, sender 2 on the next $\mathscr{L}+1$, etc. Let $M^*(s) = [M(s)]^{\mathscr{L}+1}$, where m^* will denote a typical element of $M^*(s)$ for some s. Clearly, this game is equivalent to an open forum where we replace the original message sets with the $M^*(s)$ sets. It is also easy to see that the fact that the $M(s)$ sets satisfy one-way disprovability implies that the $M^*(s)$ sets also satisfy this condition. That is, if $s \neq s'$, then $M^*(s) \neq M^*(s')$. We now show that the $M^*(s)$ sets satisfy refutability. By Propositions 1 and 3, this will imply the existence of a robust inference rule.

To see this, first note that $M^*(s) \subseteq M^*(s')$ if and only if $M(s) \subseteq M(s')$. Let

$$\tau^*(s) = M^*(s) \Big\backslash \left[\bigcup_{s' \in S^*(s)} M^*(s') \right]$$

where, as before, $S^*(s) = \{s' \neq s \mid M(s') \subset M(s)\}$. Fix any s'. We claim that for every $s \notin S^*(s')$, we have $\tau^*(s) \not\subseteq M^*(s')$. Obviously this is true if every s is in $S^*(s')$. So suppose this is not the case. Let s be any state satisfying $s \notin S^*(s')$ (that is, $M(s) \not\subset M(s')$). Let m_1 denote any $m \in M(s) \backslash M(s')$. By definition, we must have $\# S^*(s) \leqslant \mathscr{L}$. For each $s'' \in S^*(s)$, $s'' \neq s'$, there is a message $m_{s''} \in M(s) \backslash M(s'')$. Let m^* be any vector of $\mathscr{L} + 1$ messages containing $m_{s''}$ for each $s'' \in S^*(s)$, $s'' \neq s'$, and also containing m_1. Clearly, regardless of whether $s' \in S^*(s)$, we must have $m^* \in \tau^*(s) \backslash M^*(s')$. Hence $\tau^*(s) \not\subseteq M^*(s')$. ∎

Proof of Proposition 10. Fix any balanced sequential game, $(K, \mathscr{I})$. To aid with intuition, we will describe the inference rule as though each message is interpreted as a claim of a particular state, where the claim may depend on the history to that point. Formally, this is just a step in the construction of the rule. We take this specification to have the following properties. Suppose on history h, sender $\mathscr{I}(h)$ sends message m. The claim $s(h, m)$ associated with message m on this history must be feasible—that is, $s(h, m) \in F(h \cdot m)$. Second, every claim is possible. More precisely, if $s \in F(h \cdot m)$, then there is a message m' with the same pure information content as m (i.e., $F(m) = F(m')$) such that $s = s(h, m')$. The rich language condition implies that there are enough messages with the same pure information content for this to be possible.

The receiver's inference rule is as follows. Fix any $s(\cdot)$ function satisfying the above and let the vector of messages be $\underline{m} = (m_1, ..., m_K)$. We can associate with this sequence of messages a sequence of claims $(s^1, ..., s^K)$ defined by $s^1 = s(h_0, m_1)$, $s^2 = s(m_1, m_2)$, etc. If $s^1 \in F(\underline{m})$, then the inference is s^1. Otherwise, let k_1 be the smallest k such that $m_k \notin M(s^1)$. If $s^{k_1} \in F(\underline{m})$, then the inference is s^{k_1}. Otherwise, let k_2 be the smallest $k > k_1$ such that $m_k \notin M(s^{k_1})$, etc.

We claim that for any s, there is no equilibrium in the induced game given this inference rule and any preference profile in $\mathscr{P}^*_{\text{OS}}$ such that the receiver's inference differs from s. Suppose, to the contrary, that such an equilibrium exists. Let the receiver's inference in this equilibrium be $s' \neq s$.

Suppose $M(s) \subset M(s')$. By the ordered subset condition on sender 1's preferences, $s \succ_{1,s} s'$. Clearly, if sender 1 had claimed s, this would be the inference since it could not be refuted. Hence he would have been better off claiming s, contradicting the assumption that we have an equilibrium.

Hence, using one-way disprovability, it must be true that $M(s) \not\subset M(s')$. Furthermore, precisely the argument above shows that $s' \succcurlyeq_{1,s} s$. By conflicting preferences, there must be some sender i such that $s \succ_{i,s} s'$. Clearly, we cannot have $i = 1$. Since the game is balanced, sender i must have a turn to speak. Let s'' be the state on the table at sender i's turn. (That is, if h is the history up to sender i's turn and $s^1 \in F(h)$, then $s'' = s^1$. If not, define k_1 as above. If $s^{k_1} \in F(h)$, then $s'' = s^{k_1}$, etc.) If $M(s) \not\subset M(s'')$, then there must be some message $m \in M(s) \backslash M(s'')$ which sender i could send which would guarantee that the final inference is s. Since $s \succ_{i,s} s'$, this deviation would be strictly better for i, again contradicting the hypothesis that we have an equilibrium. Hence it must be true that $M(s) \subset M(s'')$. But then no sender following i can refute s'' either, so s'' must be the final inference. That is, we must have $s'' = s'$. But $M(s) \not\subset M(s')$ and $M(s) \subset M(s'')$, a contradiction.

Therefore, there is no equilibrium in the induced game for any state s in which the receiver's inference differs from s. As before, the induced game must have a pure strategy equilibrium, so the inference rule is robust. ∎

Proof of Proposition 11. Fix any preference profile $\succcurlyeq$ such that $\succcurlyeq_{i,s} = \succcurlyeq_{j,s'}$ for all i, j, s, and s'. Let $\succcurlyeq^* = \succcurlyeq_{i,s}$. A key fact about equilibria of the induced game with this preference profile and any inference rule is that equilibria of this game are the same as the equilibria where there is only one sender and he has preferences $\succcurlyeq^*$. We will use this fact repeatedly.

Suppose $\succcurlyeq \in \mathscr{P}$ and suppose δ is robust for $\mathscr{P}$ for the open forum. For each s, let $h(s)$ denote any sequence of messages feasible in s such that $\delta(h(s)) = s$. Since δ is robust, such sequences must exist for every s. Suppose there are states s and s', $s \neq s'$, such that $h(s) \in [M(s')]^K$ and $s \succcurlyeq^* s'$. Clearly, there is an equilibrium of the induced game at state s' in which the history $h(s)$ is generated. But this contradicts the robustness of δ.

Without loss of generality, number the states so that $s_1 \succcurlyeq^* s_2 \succcurlyeq^* \cdots \succcurlyeq^* s_L$. By the above reasoning, $h(s_1)$ must not be feasible in states $s_2, ..., s_L$. But this says exactly that there is an h in

$$M(s_1) \backslash \bigcup_{l=2}^{L} M(s_l).$$

Similarly, $h(s_2)$ must not be feasible in states $s_3, ..., s_L$, etc. The implied condition is precisely ordered provability with the states numbered as stated in the proposition.

If ordered provability holds and $s_1 \succ^* s_2 \succ^* \cdots \succ^* s_L$, then it is easy to show that a robust inference rule for $\{\succcurlyeq\}$ for the open forum exists. For any history h, let $\delta(h) = s_k$ for the largest k such that

$$h \in M(s_k) \setminus \bigcup_{l > k} M(s_l).$$

If no such k exists, let $\delta(h)$ be the worst degenerate inference according to $\succcurlyeq^*$. It is easy to see that in any state s, the best inference (according to $\succcurlyeq^*$) which can be generated by a feasible h is s. Hence there is an equilibrium in the induced game in s in which all senders choose the messages in h. If $s = s_L$, there is no other inference which can be generated, so there is no other equilibrium outcome. If $s \neq s_L$, then every inference different from s which can be generated is strictly worse than s. Hence there is no other equilibrium. ▌

REFERENCES

1. D. AUSTEN-SMITH AND J. WRIGHT, Competitive lobbying for a legislator's vote, *Social Choice and Welfare* **9** (1992), 229–257.
2. J. BANKS, A model of electoral competition with incomplete information, *J. Econ. Theory* **50** (1990), 309–325.
3. V. CRAWFORD AND J. SOBEL, Strategic information transmission, *Econometrica* **50** (1982), 1431–1451.
4. S. MATTHEWS AND D. FERTIG, Advertising signals of product quality, Northwestern University working paper, 1990.
5. J. FARRELL, Meaning and credibility in cheap-talk games, *Games Econ. Behav.* **5** (1993), 514–531.
6. M. FISHMAN AND K. HAGERTY, The optimal amount of discretion to allow in disclosure, *Quart. J. Econ.* (1990), 427–444.
7. D. FUDENBERG AND J. TIROLE, Perfect Bayesian equilibrium and sequential equilibrium, *J. Econ. Theory* **53** (1991), 236–260.
8. J. GREEN AND J.-J. LAFFONT, Partially verifiable information and mechanism design, *Rev. Econ. Stud.* **53** (1986), 447–456.
9. S. GROSSMAN, The informational role of warranties and private disclosure about product quality, *J. Law Econ.* (1981), 461–483.
10. J. HARRINGTON, The revelation of information through the electoral process, Johns Hopkins University working paper, 1989.
11. M. HARRIS AND R. TOWNSEND, Resource allocation under asymmetric information, *Econometrica* (1981), 33–64.
12. E. MASKIN, Nash equilibrium and welfare optimality, working paper, 1977.
13. P. MILGROM, Good news and bad news: Representation theorems and applications, *Bell J. Econ.* **12** (1981), 380–391.
14. P. MILGROM AND J. ROBERTS, Relying on the information of interested parties, *Rand J. Econ.* **17** (1986), 18–32.
15. J. MOORE AND R. REPULLO, Subgame perfect implementation, *Econometrica* **56** (1988), 1191–1220.

16. M. OKUNO-FUJIWARA, A. POSTLEWAITE, AND K. SUZUMURA, Strategic information revelation, *Rev. Econ. Stud.* **57** (1990), 25–47.
17. T. PALFREY AND S. SRIVASTAVA, Nash implementation using undominated strategies, *Econometrica* **59** (1991), 479–501.
18. H. S. SHIN, The burden of proof in a game of persuasion, *J. Econ. Theory* **64** (1994), 253–264.
19. A. M. SPENCE, "Market Signaling," Harvard Univ. Press, Cambridge, 1974.