Available online at www.sciencedirect.com

ScienceDirect

Journal of Economic Theory 151 (2014) 304–341

JOURNAL OF Economic Theory

www.elsevier.com/locate/jet

Dynamic strategic information transmission ☆

Mikhail Golosov [a,b], Vasiliki Skreta [c], Aleh Tsyvinski [d,b], Andrea Wilson [e]

[a] *Princeton, United States*
[b] *NES, United States*
[c] *UCL, United Kingdom*
[d] *Yale, United States*
[e] *NYU, United States*

Received 6 February 2013; final version received 21 December 2013; accepted 21 December 2013

Available online 8 January 2014

Abstract

This paper studies strategic information transmission in a finite horizon environment where, each period, a privately informed expert sends a message and a decision-maker takes an action. We show that communication in this dynamic environment drastically differs from a one-shot game. Our main result is that, under certain conditions, full information revelation is possible. We provide a constructive method to build such fully revealing equilibria; our result obtains with rich communication, in which non-contiguous types pool together, thereby allowing dynamic manipulation of beliefs. Essentially, conditioning future information release on past actions improves incentives for information revelation.

JEL classification: D82; D83

Keywords: Asymmetric information; Cheap talk; Dynamic strategic communication; Full information revelation

☆ Golosov and Tsyvinski thank National Science Foundation and EIEF for hospitality. We thank Manuel Amador, Luca Anderlini, Sandeep Baliga, Abhijit Banerjee, Marco Battaglini, Andreas Blume, Adam Brandenburger, Amanda Friedenberg, Johannes Hőrner, Jerome Mathis, Ronny Razin, Larry Samuelson, and Andy Skrzypacz for inspiring discussions and comments. Skreta gratefully acknowledges financial support from French National Research Agency (A.N.R., project "Regulate to Promote Information Sharing").

0022-0531/$ – see front matter
http://dx.doi.org/10.1016/j.jet.2013.12.012

1. Introduction

That biased experts impede information transmission often has serious consequences: Worse projects are financed, beneficial reforms are blocked, and firms may fail to reward the most productive employees. The seminal analysis of strategic information transmission by Crawford and Sobel [10] has seen applications ranging from economics and political science to philosophy and biology.[1] They assume that a biased and privately informed expert and a decision-maker interact only once. The conflict of interest results in coarse information revelation, and in some cases, no information revelation at all. There are, however, many environments in which the expert and receiver interact repeatedly and information transmission is dynamic. This paper explores sequential choice contexts in which the decision-maker seeks the expert's advice prior to each decision.

We study a dynamic, finite-horizon extension of the strategic information transmission of Crawford and Sobel [10]. In each period, an expert sends a message and a decision-maker takes an action. Only the expert knows the state of the world, which remains constant throughout the game. We maintain all other features of the Crawford and Sobel environment, and in particular, the conflict of interest between the expert and decision-maker. Our goal is to investigate the extent to which conflicts of interest limit information transmission in multi-period interactions, and whether or not it is possible for the decision maker to learn the truth.

We find that when the decision-maker and the expert are equally patient, it is often *possible* to elicit the precise truth from the expert, but this may be very costly in the early part of the game. Our analysis suggests that fully revealing equilibria become easier to construct, with better welfare properties, if the decision-maker places a greater weight on the future than does the expert. Many situations have this feature. For example, consider an individual consulting a doctor or a lawyer for the best course of action. Final decisions – e.g. whether or not to pursue a costly course of treatment, whether or not to file a lawsuit – are clearly the most important.

Our most surprising and difficult-to-prove result, Theorem 1, establishes that full information revelation is possible. This result obtains in a finite horizon environment where the two players are equally patient. The construction of the fully revealing equilibrium relies on two key features. The first is the use of what we call "separable groups": the expert employs a signaling rule in which far-apart types pool together initially, and eventually find it optimal to separate and reveal the truth. The second feature is to make advice contingent on actions: the expert promises to reveal the truth later, but only if the decision-maker follows his advice now; this initial advice, in turn, is designed to reward the expert for revealing information. In a nutshell, communication in a multi-period interaction is facilitated via an initial signaling rule that manipulates posteriors (in a way that enables precise information release in the future), initial actions which reward the expert for employing this signaling rule, and trigger strategies which reward the decision-maker for choosing these initial actions.

More broadly, our equilibrium construction combines mechanism design techniques with insights and ideas from the repeated games literature. On the expert's side, we characterize the shape of the "reward functions" (measured in payoffs) that would incentivize information revelation in a direct mechanism, and then determine the action sequences (ending with the action that a fully informed decision-maker will choose) that give rise to these payoffs. On the decision-

[1] For a survey with applications across disciplines, see Sobel [28].

maker's side, we use folk-theorem-type arguments to ensure that he will comply with the desired equilibrium action sequences.

We now explain in more detail our construction of a fully revealing equilibrium. We first show that it is possible to divide all states into separable groups. A separable group is a finite set of states (types), which are sufficiently far apart that each type would rather reveal the truth than mimic any other type in his group. The expert's initial signaling rule reveals the separable group containing the truth; therefore, this creates histories after which it is common knowledge that the decision-maker puts probability one on a particular separable group, at which point the types in this group will find it optimal to separate. The idea of initially pooling together far-away intervals of types, who will then later have an incentive to separate, was first proposed in Krishna and Morgan [19]. They demonstrated how this could increase information revelation in a variant of Crawford and Sobel's [10] game where the expert and the decision maker engage in a "conversation" (exchange of messages through a jointly-controlled lottery) before the decision-maker chooses his action.

In this paper, we study a dynamic extension with many communication-decision rounds, and demonstrate that if the initial groups of pooled types are finite and chosen in the right way, it is possible for the decision-maker to extract *all* information from a biased expert. This is technically challenging for two somewhat interrelated reasons: There is a continuum of states and finitely many periods over which learning can take place. We then have to divide the state in a continuum of separable groups, so that eventually the expert will be willing to tell the truth.

The division of all types into separable groups is quite delicate: The expert anticipates that once he joins a separable group, he will forgo his informational advantage. Thus, for the expert to join the separable group containing his true type, we have to make sure that he does not want to mimic a nearby type by joining some other separable group. This is difficult since in our model there are no transfers. We succeed in designing initial actions, which ensure that any future gain to the expert from mimicking some other type is offset by the initial cost. These expert-incentivizing actions are not myopically optimal for the decision-maker, so we employ trigger strategies: the expert (credibly) threatens to babble in the future if the decision-maker fails to choose the actions that he recommends at the beginning. The final part of the proof, then, shows that we can design the separable groups and initial actions such that the decision-maker would rather follow the expert's initial advice, knowing that he will then eventually learn the exact truth, than choose the myopically optimal action in the initial periods, knowing that he will then never learn more than the separable group containing the truth. This latter part – finding expert-incentivizing actions which the DM is actually willing to choose – is perhaps the most difficult aspect of our construction.

We emphasize several additional differences between dynamic and static communication games. First, in contrast to Crawford and Sobel [10], not all equilibria are equivalent to ones with partitional structure (equilibria where a message is send by a connected set of types). If attention is restricted to partition equilibria, learning quickly stops, whereas, as we discussed, fully revealing equilibria exist. Second, welfare properties of equilibria also differ in a dynamic setup. Crawford and Sobel [10] show that, ex-ante, both the expert and the decision-maker will (under typical assumptions) prefer equilibria with finer partitions. We provide an example that shows that it is not necessarily the case for dynamic equilibria.[2] We also present an example in

[2] A similar phenomenon occurs when communication is noisy, as shown in an example of the working paper version of Chen, Kartik, and Sobel [9]. In their example, a two-step partition welfare dominates a three-step partition.

which dynamic partition equilibria can strictly welfare-dominate the best static equilibrium, and an example showing that non-partition equilibria can strictly welfare dominate the best partition equilibrium.

Our work shows that the nature of dynamic strategic communication is quite distinct from its static counterpart. In the static case, because of the conflict of interest between the decision-maker and the expert, nearby expert types have an incentive to pool together, precluding full information revelation. The single-crossing property also implies that at equilibrium, the action is a monotonic step function of the state. These two forces make complex signaling (even though possible) irrelevant. In the dynamic setup, the key difference is that today's communication sets the stage for tomorrow's communication. Complex signaling helps in the dynamic setup, because it can generate posteriors that put positive probability only on expert types who are so far apart, they have no incentive to mimic each other. This is what enables fully revealing equilibria.

Related literature. Crawford and Sobel [10] is the seminal contribution on strategic information transmission. That paper has inspired an enormous amount of theoretical work and myriads of applications. Here we study a dynamic extension. Much of the previous work on dynamic communication has focused on the role of reputation; see, for example, Sobel [27], Morris [21], and Ottaviani and Sorensen [22,23]. Some other dynamic studies allow for multi-round communication protocols, but with a single round of action(s). Aumann and Hart [4] characterize geometrically the set of equilibrium payoffs when a long conversation is possible. In that paper, two players – one informed and one uninformed – play a finite simultaneous-move game. The state of the world is finite, and players engage in direct (no mediator) communications, with a potentially infinitely long exchange of messages, before simultaneously choosing costly actions. In contrast, in our model, only the informed party sends messages, the uninformed party chooses actions, and the state space is infinite.

Krishna and Morgan [19] add a long communication protocol to Crawford and Sobel's [10] game, and Goltsman, Hőrner, Pavlov and Squintani [14] characterize such optimal protocols.[3] Forges and Koessler [12,13] allow for a long protocol in a setup where messages can be certifiable. In all those papers, once the communication phase is over, the decision-maker chooses one action. In our paper, there are multiple rounds of communication and actions (each expert's message is followed by an action of the decision-maker). The multiple actions align incentives in a way that was not possible in these earlier works: the expert is able to condition his advice on the decision-maker's past behavior, and additionally, the decision-maker is able to choose actions which reward the expert appropriately for following a path of advice that ultimately leads to revelation of the true state.

In our setup, the dynamic nature of communication enables full information revelation. In contrast, full information revelation is not possible in the dynamic setup of Anderlini, Gerardi, and Lagunoff [2], who consider dynamic strategic communication in a dynastic game, and show that if preferences are not fully aligned, "full learning" equilibria do not exist.[4] Renault, Solan,

[3] They examine the optimal use of a third party, such as a mediator or negotiator, to relay messages. For the *expert*, this model is strategically equivalent to ours: his expected payoff is the same whether he induces a sequence of actions; or a probability distribution over these actions. For the *decision-maker*, however it is not: In our dynamic setup, the DM's past actions can affect future communication, and so it is possible to induce him to choose actions which are not myopically optimal.

[4] In their model, the state space is finite (0 or 1), and there is no perfectly informed player: each receiver gets a signal about the state and a message from his predecessor, and then becomes the imperfectly informed advisor to the next player.

and Vielle [26] examine dynamic sender-receiver games, and characterize equilibrium payoffs (for quite general preferences) in an infinite-horizon model with a finite state space, and a state that evolves according to a stationary Markov process. In contrast, we assume a continuous state space with persistent information, and our focus is on the possibility of full information revelation in finite time.[5]

Our model bears some similarities to models of static strategic communication with multiple receivers. In those models, see, for instance, Goltsman and Pavlov [15], the expert cares also about a sequence of actions, but in contrast to our model, those actions are chosen by different individuals. An important difference is that in our model, the receiver cares about the entire vector of actions chosen; in those models, each receiver cares only about his own action. This enables our use of trigger strategies, which we find is a necessary feature of equilibria with eventual full information revelation. Still, some of the properties of the equilibria that we obtain also appear in the models with multiple receivers. In particular, our non-monotonic example (Example 3) in Section 4 resembles Example 2 of Goltsman and Pavlov [15]. It is also similar to Example 2 in Krishna and Morgan [19], whereas our Example 2 is similar to Example 1 in Krishna and Morgan [19].[6]

Conceptually, our result relates to a large literature on repeated games with incomplete information. The expert knows that if he reveals the state, then he loses any ability to influence the decision-maker's subsequent action choices; due to a conflict of interest between players, this creates an incentive to conceal some information. At the same time, a completely uninformed decision-maker may choose actions which are bad for both players. The question is then, how much information should the expert reveal.

In an early related result, Aumann and Maschler [5] studied optimal revelation policies in infinitely repeated zero-sum games, with only one player informed about the (binary) state. They found that the amount of information revealed at equilibrium depends on the degree of conflict of interest,[7] just as we find that fully revealing equilibria are possible only when the conflict of interest is sufficiently small. Information revelation was further facilitated in their model by the fact that both players choose actions (which allows the expert to restrict the outcome choices available to uninformed player), and the fact that their state space was finite. More recently, there has been work on constructing fully revealing equilibria in the *known-own-payoffs* case, where each player's payoff depends only on his own private type (see, for example, Athey and Bagwell [3] and Peski [25]). In contrast, we assume that *both* players' payoffs depend on a

[5] Ivanov [16] allows for a dynamic communication protocol in a setup where the expert is also initially uninformed, and the decision-maker controls the quality of information available to the expert. He employs separable groups, but in a much different informational setting: His decision-maker has a device that initially reveals (to the expert only) the separable group containing the truth, and contains a built-in threat to only reveal the exact state if the expert reports this information truthfully. Compared to our model, this eliminates all incentive requirements for the decision-maker, and imposes an additional cost on the expert (namely, he will fail to learn the truth himself) if he fails to follow the prescribed strategy, thus weakening the required incentive constraints.

[6] Equilibria can be non-partitional also in environments where the decision-maker consults two experts as in Krishna and Morgan [18].

[7] For example, they trivially obtain full information revelation in a game where the informed player chooses either T or B, the uninformed player chooses L or R, and the informed player (P1) earns payoff 0 from outcomes (B, R) in state 1 and (T, L) in state 2, and payoff 4 in all other cases. Clearly, it is optimal for player 1 to play T in state 1 and B in state 2, guaranteeing his highest possible payoff (4) regardless of P2's action choice. Of course, such a fully revealing strategy would not be optimal if it led P2 to choose an action which P1 strongly disliked. P1 also would not reveal the truth if he could send only a payoff-irrelevant message, then letting P2 choose the outcome (as is the case in our model).

common state variable. Finally, there is a large literature on reputation, in which one player with private information wishes to convince his opponent that he is a particular "type", in the hope of inducing desirable action choices. This idea is present here, too: indeed, much of the difficulty in constructing fully revealing equilibria arises from the fact that the expert is "biased", and in particular, would like to convince the DM that the state is something other than it is. With a myopic uninformed player, it is generally the case that the informed player successfully builds a reputation as his most-preferred type; in other words, fully revealing equilibria do not exist.

However, this existing literature deals almost exclusively with a finite (usually binary) state space, whereas our state space is a continuum. This makes it much more difficult to elicit the truth from a biased expert: As described in the equilibrium outline above, initial actions must be chosen carefully to incentivize the expert to tell the truth, rather than mimicking nearby types. This constraint would of course not be present in a finite model, with types spaced sufficiently far apart. Additionally, in most existing models, both players choose actions, and the horizon is infinite. In our model, the horizon is finite, and only the informed player chooses actions (which then determine both players' payoffs). Therefore, existing results say little about the extent to which information revelation is possible in our game.

Summarizing, we study a finite-horizon setting in which both players' payoffs depend on a common state variable, about which only one player is informed. Preferences are partially aligned, but with a conflict of interest, as in Crawford and Sobel [10]. For this setting, fully revealing equilibria have proved difficult, and have previously been found only for the following modifications of the model: if the expert consults two experts as in Battaglini [6], Eso and Fong [11], and Ambrus and Lu [1]; when information is completely or partially certifiable, as in Mathis [20]; and when there are lying costs and the state is unbounded as in Kartik, Ottaviani, and Squintani [17]. In the case of multiple experts, playing one against the other is the main force that supports truthful revelation. In the case of an unbounded state, lying costs become large and support the truth. In the case of certifiable information, one can exploit the fact that messages are state-contingent to induce truth-telling. All these forces are very different from the forces behind our fully revealing construction, which exploits the alignment of incentives that arise in a dynamic setting when both players' payoffs depending on a common state variable.

2. Motivating example: an impatient financial advisor

One of the striking results of the static strategic communication game is that there exist no equilibria with full information revelation. Although the state can take a continuum of values, all equilibria are equivalent to one in which the expert sends at most finitely distinct many signals to the decision-maker. That is, a substantial amount of information is not transmitted. This example motivates the general construction used to establish Theorem 1.

We show how to construct a fully revealing perfect Bayesian equilibrium when the expert is myopic, using just two periods. There are two essential ingredients of this example. First, the set of types that pool together in the first period are far enough apart that they can be separated in the second period: that is, each possible first-period message is sent by a separable group of types. Second, each separable group induces the same optimal (for the decision-maker) first-period action. This implies that the expert does not care which group he joins (since a myopic expert cares only about the 1st-period action, which is constant across groups).

Example 1 *(Fully revealing equilibrium with impatient experts).* Suppose there is an expert E (financial advisor) and a decision-maker DM (an employee). The expert knows the true state of the world θ, which is drawn from a uniform distribution on [0, 1] and remains constant over time. The players' payoffs in period $t \in \{1, 2\}$ depend on both the state, θ, and on the action chosen by the decision-maker, $y_t \in \mathbb{R}$. More precisely, payoffs in period t are given by

$$u_t^E(y_t, \theta, b) = -(y_t - \theta - b)^2 \quad \text{and} \quad u_t^{DM}(y, \theta) = -(y_t - \theta)^2, \tag{1}$$

where $b > 0$ is the expert's "bias". The expert is myopic, with discount factor $\delta^E = 0$; the construction works for any discount factor for the decision-maker.

The expert employs the following signaling rule. In period 1, expert types $\{\frac{1}{8} - \varepsilon, \frac{3}{8} + \varepsilon, \frac{4}{8} + \varepsilon, 1 - \varepsilon\}$ pool together and send message m_ε, for all $\varepsilon \in (0, \frac{1}{8})$. For all type pairs $\{\frac{1}{8} + \tilde{\varepsilon}, \frac{7}{8} - \tilde{\varepsilon}\}$ with $\tilde{\varepsilon} \in [0, \frac{1}{4}]$, the expert sends a message $m_{\tilde{\varepsilon}}$. Expert types $\{0, \frac{4}{8}, 1\}$ send message m_b. That is, we have two families of separable groups indexed by ε and $\tilde{\varepsilon}$ that cover the entire interval except the types $\{0, \frac{4}{8}, 1\}$, and 1 additional separable group consisting of these remaining states. Noting that the *expected* type in each of these information set is $\frac{1}{2}$, it follows that the DM's best response in period 1 is to choose the action $y_1(m_\varepsilon) = y_1(m_{\tilde{\varepsilon}}) = y_1(m_b) = \frac{1}{2}$, for all equilibrium messages m_ε, $m_{\tilde{\varepsilon}}$, and m_b. In period 2, the expert reveals the truth, and so the DM chooses an action equal to the true state. After any out-of-equilibrium initial message, the DM assigns equal probability to all states, leading to action $y_1^{out} = 0.5$. After any out-of-equilibrium second-period message, the DM assigns probability 1 to the lowest type in his information set (prior to the off-path message), and accordingly chooses an action equal to this type.

We now argue that this is an equilibrium for any $b \leqslant \frac{1}{16}$: First, notice that all messages (even out-of-equilibrium ones) induce the same action in period 1. Hence, the expert is indifferent between all possible first-period messages if he puts zero weight on the future. So, in particular, a myopic expert will find it optimal to send the "right" message, following the strategy outlined above. Now consider, for example, the history following an initial message m_ε. The DM's posterior beliefs assign probability $\frac{1}{4}$ to each of the types in $\{\frac{1}{8} - \varepsilon, \frac{3}{8} + \varepsilon, \frac{4}{8} + \varepsilon, 1 - \varepsilon\}$. The expert's strategy at this stage is to tell the truth: so, if he sends a message that he is type $k \in \{\frac{1}{8} - \varepsilon, \frac{3}{8} + \varepsilon, \frac{4}{8} + \varepsilon, 1 - \varepsilon\}$, then the DM will believe that k is the true state, and accordingly will choose action k; if the expert deviates to some off-path message, then the DM will assign probability 1 to the lowest type in his information set, $\frac{1}{8} - \varepsilon$, and accordingly choose action $\frac{1}{8} - \varepsilon$. Therefore, to prove that the expert has no incentive to deviate, we need only show that each expert type $k \in \{\frac{1}{8} - \varepsilon, \frac{3}{8} + \varepsilon, \frac{4}{8} + \varepsilon, 1 - \varepsilon\}$ would rather tell the truth, than mimic any of the other types in his group. Type k prefers action k to k' whenever

$$-(k - k - b)^2 \geqslant -\bigl(k' - k - b\bigr)^2 \quad \Leftrightarrow \quad \bigl(k' - k\bigr)\bigl(k' - k - 2b\bigr) \geqslant 0$$

i.e., whenever $k' < k$, or whenever $k' > k + 2b$. So in particular, to make sure that no type in $\{\frac{1}{8} - \varepsilon, \frac{3}{8} + \varepsilon, \frac{4}{8} + \varepsilon, 1 - \varepsilon\}$ wishes to mimic any other type in this group, it is sufficient to make sure that every pair of types are at least $2b$ apart. Since the closest-together types in the group, $\frac{3}{8} + \varepsilon$ and $\frac{4}{8} + \varepsilon$, are separated by $\frac{1}{8}$, we conclude that the group is separable whenever $\frac{1}{8} > 2b \Leftrightarrow b < \frac{1}{16}$. And similarly after messages $m_{\tilde{\varepsilon}}$ and m_b.

This construction does not apply with a more patient expert ($\delta^E > 0$), because it does not provide a forward-looking expert with incentives to join the "right" separable group.[8] For example,

[8] Another possible critique of the construction, is that it is fragile in the sense that the expert is indifferent between any of the messages used in equilibrium. However, this kind of "fragility" is common in game theory, and indeed present in every mixed-strategy equilibrium.

consider type $\frac{3}{8}$, and suppose that $b = \frac{1}{16}$. The truthful strategy is to reveal group $\{\frac{1}{8}, \frac{3}{8}, \frac{4}{8}, 1\}$ in period 1, and the tell the truth in period 2, inducing actions $(y_1, y_2) = (\frac{1}{2}, \frac{3}{8})$. However such strategy cannot be part of an equilibrium if $\delta^E > 0$. The best deviation for $\theta = \frac{3}{8}$ is to mimic type $\frac{3}{8} + \frac{1}{16}$ – initially claiming to be part of the group $\{\frac{1}{8} - \frac{1}{16}, \frac{3}{8} + \frac{1}{16}, \frac{4}{8} + \frac{1}{16}, \frac{7}{8} - \frac{1}{16}\}$, and then subsequently claiming that the true state is $\frac{3}{8} + \frac{1}{16}$ – thereby inducing actions $(y_1, y_2) = (\frac{1}{2}, \frac{3}{8} + \frac{1}{16})$. This deviation then leads to no change in the first-period action, but the 2nd-period action is now equal to type $\frac{3}{8}$'s bliss point, $\frac{3}{8} + \frac{1}{16}$. When $\delta^E > 0$ we need to provide the expert with better incentives to join the "right" separable group: since θ prefers $\theta + b$'s action in the future, he must prefer his own action now. This is much more complex, but in Section 5, we show how to construct such separation-inducing actions.

3. The model

There are two players, an expert (E) and a decision-maker (DM), who interact for finitely many periods. The expert knows the true state of the world $\theta \in [0, 1]$, which is constant over time and is distributed according to the c.d.f. F, with associated density f. Both players care about their discounted payoff sum: when the state is θ and the DM chooses actions $y^T = (y_1, \dots, y_T)$ in periods $1, 2, \dots, T$, payoffs are given by

$$\text{Expert:} \quad U^E(y^T, \theta, b) = \sum_{t=1}^{T} (\delta^E)^{t-1} u^E(y_t, \theta, b),$$

$$\text{DM:} \quad U^{DM}(y^T, \theta) = \sum_{t=1}^{T} (\delta_{DM})^{t-1} u^{DM}(y_t, \theta)$$

where $b > 0$ is the expert's "bias" and reflects a conflict of interest between the players, and δ^E, δ^{DM} are the players' discount factors. We assume that $u^E(y_t, \theta, b)$ and $u^{DM}(y_t, \theta)$ satisfy the conditions imposed by Crawford and Sobel [10]: for $i = DM, E$, $u^i(\cdot)$ is twice continuously differentiable, $u^i_1(y, \theta) = 0$ for some y and $u^i_{11}(\cdot) < 0$ (so that u^i has a unique maximizer y for each pair (θ, b)), and that $u^i_{12}(\cdot) > 0$ (so that the best action from an informed player's perspective is strictly increasing in θ). Most of our main results will make the more specific assumption that preferences are quadratic, as given by (1).

At the beginning of each period t, the expert sends a (possibly random) message m_t to the DM. The DM then updates his beliefs about the state, and chooses an action $y_t \in \mathbb{R}$ that affects both players' payoffs. Let $y^{DM}(\theta)$ and $y^E(\theta)$ denote, respectively, the DM's and the expert's most preferred actions in state θ; we assume that for all θ, $y^{DM}(\theta) \neq y^E(\theta)$, so that there is a conflict of interest between the players regardless of the state.

We assume that the DM observes his payoffs only at the end of the game. This is to rule out cases in which the DM can make inferences about the state from observing his payoff, as we wish to focus solely on learning through communication.

A *strategy profile* $\sigma = (\sigma^i)_{i=E,DM}$, specifies a strategy for each player. Let h_t denote a history that contains all the reports submitted by the expert, $m^{t-1} = (m_1, \dots, m_{t-1})$, and all actions chosen by the DM, $y^{t-1} = (y_1, \dots, y_{t-1})$, up to stage t. The set of all feasible histories at t is denoted by H_t. A behavioral strategy for the expert, σ_E, consists of a sequence of signaling rules that map $[0, 1] \times H_t$ to a probability distribution over reports in the Borel set $\mathcal{M}$. Let $q(m|\theta, h_t)$ denote the probability that the expert reports message m at history h_t when his type is θ. A strategy for the DM, σ_{DM}, is a sequence of maps from H_t to actions. We use $y_t(m|h_t) \in \mathbb{R}$ to denote the action that the DM chooses at h_t given a report m. A *belief system*, μ, maps H_t to the set of probability distributions over [0, 1]. Let $\mu(\theta|h_t)$ denote the DM's beliefs about the experts's type after a history h_t.[9] A strategy profile σ and a belief system μ is an assessment. We seek strategy profiles and belief systems that form *Perfect Bayesian Equilibria*, (*PBE*).[10]

We use the terminology as follows: a *babbling equilibrium* is one in which all expert types $\theta \in [0, 1]$ follow the same strategy, and thus the DM chooses some constant action $\hat{y}$ after all histories. A *partition equilibrium* is one in which each message along the equilibrium path is sent by a connected interval of types, and each type in the interval sends the message with probability one. Following Crawford and Sobel [10], we refer to the expert's strategy in such equilibria as *uniform signaling*. Finally, we say that an equilibrium is *fully revealing* if there exists a time $\widehat{T} \leqslant T$ such that for all $\theta \in [0, 1]$, expert type θ (at equilibrium) sends a message sequence that reveals his true type with probability 1 by time $\widehat{T}$, and accordingly, $y_t(\theta) = y^{DM}(\theta)$, $\forall t \geqslant \widehat{T}$.

4. Dynamic partition equilibria

Recall from Crawford and Sobel [10] that in the one-shot strategic communication game, all equilibria have a partitional structure: Intervals of expert types pool together to send the same message, inducing actions which are increasing step functions of the state. Communication is then coarse; even though the state θ takes a continuum of values, only finitely many different actions are induced.

Equilibria with this monotonic partitional structure preclude full information revelation, even in a dynamic setting:

Proposition 1. *For all horizons T, there exist no fully revealing monotonic partition equilibria.*

This follows almost immediately from Crawford and Sobel [10], whose results can be invoked due to the fact that monotonic partition equilibria imply posterior distributions that are continuous over some interval. Suppose, by contradiction, that there exists a fully revealing monotonic partition equilibrium. Then, there is a period $\widehat{T} \leqslant T$ in which the last subdivision occurs, with $y_t(\theta) = y^{DM}(\theta)$ for all $t \geqslant \widehat{T}$. Then, the incentive constraint at time $\widehat{T}$ for type θ to not mimic $\theta + \varepsilon$ is

$$\begin{aligned}&\big(1+\delta+\delta^2+\cdots+\delta^{T-\widehat{T}-1}\big)u^E\big(y^{DM}(\theta),\theta,b\big)\\&\quad\geqslant\big(1+\delta+\delta^2+\cdots+\delta^{T-\widehat{T}-1}\big)u^E\big(y^{DM}(\theta+\varepsilon),\theta,b\big)\end{aligned}$$

and similarly for type $\theta + \varepsilon$. These conditions are equivalent to the static equilibrium conditions in Crawford and Sobel [10], who proved that they imply that at most finitely many actions can be induced at an equilibrium of a static game, and therefore full information revelation is impossible.

[9] We follow the distributional approach of Milgrom and Weber. For a full discussion of why the formulations leads to regular conditional distributions as posterior beliefs see footnote 2 in Crawford and Sobel [10].

[10] We use the typical extension of the PBE concept for infinite state spaces: both players' strategies must maximize their expected payoffs after all histories, and beliefs must be Bayesian (see Eq. (7)) after all equilibrium message sequences. Our proof of Theorem 1 is by construction, and will ensure that all payoff expressions are well-defined.

By a similar argument (details in Appendix A), we obtain the following result:

Proposition 2. *If the only equilibrium in the static game is babbling, then all monotonic partition equilibria in the dynamic game are babbling.*

Observe that any equilibrium of the one-shot game can be replicated our dynamic game, simply by playing the static equilibrium in the first period, and letting the expert babble thereafter. The DM will then repeat his first-period action choice in all periods, and so both players' average-per-period payoffs will equal their payoffs in the corresponding equilibrium of the one-shot game. We call such equilibria *static partition equilibria*. In the dynamic game, there may exist additional partition equilibria, in which the state space is ultimately partitioned into more intervals. Our next example shows that this can be welfare-improving:

Example 2 *(More intervals can be welfare-improving).* Suppose that $\delta^E = \delta^{DM} = 1$, types are uniformly distributed on $[0, 1]$ and preferences satisfy (1), with bias $b = \frac{1}{12}$. From the analysis of Section 4 in Crawford and Sobel [10], it follows that the static game has only two equilibria: a babbling equilibrium, and an equilibrium with a 2-interval partition, $[0, \frac{1}{3}] \cup [\frac{1}{3}, 1]$, inducing actions $\frac{1}{6}$ and $\frac{4}{6}$.

We look for an equilibrium with the following signaling rule:

types in $[0, \theta_1)$ send message sequence $A = (m_{1(1)}, m_{2(1)})$,

types in $[\theta_1, \theta_2]$ send message sequence $B = (m_{1(2)}, m_{2(2)})$,

types in $(\theta_2, 1]$ send message sequence $C = (m_{1(2)}, m_{2(3)})$.

Thus, the interval $[0, 1]$ is partitioned into $[0, \theta_1] \cup [\theta_1, 1]$ in the first period, and then types in $[\theta_1, 1]$ subdivide further into $[\theta_1, \theta_2] \cup [\theta_2, 1]$ in the second period. The second-period actions induced are $y_{2(1)} = \frac{\theta_1}{2}$, $y_{2(2)} = \frac{\theta_1+\theta_2}{2}$, and $y_{2(3)} = \frac{1+\theta_2}{2}$, and the first-period actions are $y_{1(1)} = \frac{\theta_1}{2}$ and $y_{1(2)} = \frac{1+\theta_1}{2}$. Off-path: the DM assigns probability 1 to type $\theta_1/2$ (and so chooses action $y_{1(1)} = y_{2(1)}$) if he gets any out-of-equilibrium message in the first period, or $m_{1(1)}$ followed by an out-of-equilibrium second-period message. If he gets message $m_{1(2)}$ followed by an off-path second message, he assigns probability 1 to the interval $[\theta_1, \theta_2]$, and so he chooses action $y_{2(2)}$. It is then immediate that the expert cannot gain with an off-path deviation.

In period 2, type θ_2 must be indifferent between the actions $y_{2(2)}$ and $y_{2(3)}$, yielding the following indifference condition:

$$\left(\frac{\theta_1+\theta_2}{2} - \theta_2 - b\right)^2 = \left(\frac{1+\theta_2}{2} - \theta_2 - b\right)^2 \quad \Rightarrow \quad \theta_2 = \frac{1}{3} + \frac{1}{2}\theta_1. \tag{2}$$

And in period 1, type θ_1 must be indifferent between message sequences A and B:

$$\left(\frac{1+\theta_1}{2} - \theta_1 - \frac{1}{12}\right)^2 + \left(\frac{3}{4}\theta_1 + \frac{1}{6} - \theta_1 - \frac{1}{12}\right)^2 = 2\left(\frac{\theta_1}{2} - \theta_1 - \frac{1}{12}\right)^2.$$

Together with (2), this implies cutoffs $\theta_1 = 0.2482$, $\theta_2 = 0.457\,43$; with this, the actions become $y_{1(1)} = y_{2(1)} = 0.124\,1$, $y_{1(2)} = 0.624\,1$, $y_{2(2)} = 0.352\,8$, and $y_{2(3)} = 0.728\,7$.

In our dynamic equilibrium, the expert's (ex ante) payoff is -0.0659 and the DM's (ex ante) payoff is -0.052. If the most informative static equilibrium is played in both periods, payoffs are −0.069 to the expert, −0.055 to the DM, both strictly worse than in our dynamic monotonic partition equilibrium.[11,12]

Next we present an example with a non-partitional equilibrium, in which higher expert types do not always induce (weakly) higher first-period actions. In this example, the bias is so severe that in a static setting, all equilibria would be babbling. We show that even in these extreme bias situations, some information can be revealed with just two rounds. This equilibrium has the feature that the DM learns the state quite precisely when the news is either horrific or terrific, but remains agnostic for intermediate levels. Finally we show that for a range of biases, this equilibrium provides higher expected payoff to both players compared to all partitional equilibria.

Example 3 *(A non-partition equilibrium).* Consider a two-period game where $\delta^E = \delta^{DM} = 1$, types are uniformly distributed on [0, 1] and preferences are given by (1). We will construct an equilibrium with the following "piano teacher" interpretation: a child's parent (the DM) wants the amount of money he spends on lessons to correspond to the child's true talent θ, whereas the piano teacher (expert) wants to inflate this number. In our equilibrium, parents of children who are at either the bottom or top extreme of the talent scale get the same initial message, "you have an interesting child" ($m_{1(1)}$ below), and then find out in the second period whether "interesting" means great ($m_{2(3)}$) or awful ($m_{2(1)}$); parents of average children are told just that in both periods. More precisely, let the expert use the following signaling rule:

In period 1, expert types in $[0, \underline{\theta}) \cup (\bar{\theta}, 1]$ send message $m_{1(1)}$, and types in $[\underline{\theta}, \bar{\theta}]$ send message $m_{1(2)}$. In period 2, types in $[0, \underline{\theta})$ send message $m_{2(1)}$, types in $[\underline{\theta}, \bar{\theta}]$ send a message $m_{2(2)}$, and types in $(\bar{\theta}, 1]$ send $m_{2(3)}$ (all with probability one). With this signaling rule, the optimal actions for the DM in period 1 are $y_{1(1)} = \frac{\underline{\theta}^2 - \bar{\theta}^2 + 1}{2(\underline{\theta} - \bar{\theta} + 1)}$, $y_{1(2)} = \frac{\underline{\theta} + \bar{\theta}}{2}$; in period 2, they are $y_{2(1)} = \frac{\underline{\theta}}{2}$, $y_{2(2)} = \frac{\underline{\theta} + \bar{\theta}}{2}$, $y_{2(3)} = \frac{1 + \bar{\theta}}{2}$. After any out-of-equilibrium first-period message, the DM assigns equal probability to all states in $[\underline{\theta}, \bar{\theta}]$, and chooses action $y_{1(2)}$; after any out-of-equilibrium second-period message following $m_{1(1)}$, the DM assigns equal probability to all types in $[0, \underline{\theta})$, and chooses action $y_{2(1)}$; and after any other off-path message-sequence, the DM assigns equal probability to types in $[\underline{\theta}, \bar{\theta}]$, and so will choose action $y_{2(2)}$. It immediately follows that no expert type can gain by sending an out-of-equilibrium message.

In order for this to be an equilibrium, type $\underline{\theta}$ must be indifferent between message sequences $A \equiv (m_{1(1)}, m_{2(1)})$ and $B \equiv (m_{1(2)}, m_{2(2)})$:

$$-\left(\frac{\underline{\theta}^2 - \bar{\theta}^2 + 1}{2(\underline{\theta} - \bar{\theta} + 1)} - \underline{\theta} - b\right)^2 - \left(\frac{\underline{\theta}}{2} - \underline{\theta} - b\right)^2 = -2\left(\frac{\underline{\theta} + \bar{\theta}}{2} - \underline{\theta} - b\right)^2 \tag{3}$$

and type $\bar{\theta}$ must be indifferent between message sequences B and $C \equiv (m_{1(1)}, m_{2(3)})$:

$$-\left(\frac{\underline{\theta}^2 - \bar{\theta}^2 + 1}{2(\underline{\theta} - \bar{\theta} + 1)} - \bar{\theta} - b\right)^2 - \left(\frac{1 + \bar{\theta}}{2} - \bar{\theta} - b\right)^2 = -2\left(\frac{\underline{\theta} + \bar{\theta}}{2} - \bar{\theta} - b\right)^2. \tag{4}$$

[11] In constructing this strategy profile, we imposed only local incentive compatibility constraints, requiring that type θ_1 is indifferent in period 1 between inducing action sequence $(y_{1(1)}, y_{2(1)})$ and $(y_{1(2)}, y_{2(2)})$, and that type θ_2 is indifferent in period 2 between inducing actions $y_{2(2)}$ and $y_{2(3)}$. It is routine to verify that these conditions are sufficient for global incentive compatibility. Details are available from the authors upon request.

[12] In Appendix B we present an example that demonstrate that equilibria with more partitions can be Pareto inferior to the equilibria with fewer partitions.

At $t = 2$ it must also be the case that type $\underline{\theta}$ prefers $m_{2(1)}$ to $m_{2(3)}$, and the reverse for type $\bar{\theta}$: that is $-(\frac{\underline{\theta}}{2} - \underline{\theta} - b)^2 \geqslant -(\frac{1+\bar{\theta}}{2} - \underline{\theta} - b)^2$ and $-(\frac{1+\bar{\theta}}{2} - \bar{\theta} - b)^2 \geqslant -(\frac{\underline{\theta}}{2} - \bar{\theta} - b)^2$. The global incentive compatibility constraints, requiring that all types $\theta < \underline{\theta}$ prefer sequence A to B and that all types $\theta > \bar{\theta}$ prefer C to B, reduce to a requirement that the average induced action be monotonic, which is implied by indifference constraints (3), (4).

A solution of the system of Eqs. (3) and (4) gives an equilibrium if $0 \leqslant \underline{\theta} < \bar{\theta} \leqslant 1$. We solved this system numerically, and found that the highest bias for which it works is $b = 0.256$. Here, the partition cutoffs in our equilibrium are given by $\underline{\theta} = 0.0581$, $\bar{\theta} = 0.9823$. The corresponding optimal actions for period 1 are $y_{1(1)} = 0.253$, $y_{1(2)} = 0.52$, and for period 2 they are $y_{2(1)} = 0.029$, $y_{2(2)} = 0.52$, $y_{2(3)} = 0.991$. Note that while the first period action is non-monotonic, the average action $\bar{y} = \frac{y_1+y_2}{2}$ is still weakly increasing in the state. Ex ante payoffs are -0.275 for the expert, and -0.144 for the DM.

Recall that in a one-shot game with quadratic preferences, all equilibria are babbling when $b > \frac{1}{4}$. Proposition 2 implies that at $b = 0.256$, if we restricted attention to partition equilibria, we would again find only a babbling equilibrium, in which the DM chooses action $y^B = 0.5$ in both periods: this yields ex-ante payoffs of -0.298 to the expert, -0.167 to the DM, strictly worse than in our above construction.[13]

Our example therefore illustrates how allowing for non-partition equilibria can both increase the amount of information revelation, and can also strictly welfare-dominate the best static equilibrium. By pooling together the best and the worst states in period 1, the expert is willing to reveal in period 2 whether the state is very good or very bad. It also has the following immediate implication:

Proposition 3. *When the expert's bias b is sufficiently large, there exist non-partition equilibria that are welfare superior to all partition equilibria.*

We now move on to our first main result, showing that our dynamic setup aligns the incentives of the expert and DM in such a way that *full information revelation* is possible.

5. Learning the truth when the expert is patient

When the expert is forward-looking, getting him to reveal the truth is much more complicated, as we previewed in Section 2. In this section, we construct a fully revealing equilibrium for the quadratic preferences specified in (1). The equilibrium works as follows: In each period, the expert recommends an action to the DM. Initially, each action is recommended by finitely many (at most four) expert types, who then subdivide themselves further into separable groups of two with an interim recommendation. If the DM chooses all initial actions recommended by the expert, then the expert rewards him by revealing the truth in the final stage of the game, recommending an action $y(\theta) = \theta$. If the DM rejects the expert's early advice, then the expert babbles for the rest of the game, and so the DM never learns more than the separable group containing the truth.

[13] This construction yields strictly higher payoffs compared to best *monotonic* partition equilibrium for all $b \in (0.25, 0.256]$.

↓	↓	↓	↓	↓
Phase 1	**Phase 2**	**Phase 3**	**Phase 4**	**Phase 5**
$(t = 0)$	$(t = 2\alpha_0)$	$(t = 2\alpha_a)$	$(t = 2)$	$(t = T - \tau)$
Group I: v_1	Group I:	Groups II, III:	Groups II, III:	Group I:
Groups II, III: u_1	switch to v_2	switch to u_2	reveal truth	reveal truth

Fig. 1. Timeline.

We provide here a sketch our construction; this is followed by the statement and discussion of our main result, with full proof details in Appendix C.

5.1. Outline

To simplify notation, we rescale the state space by dividing all actions and types by the bias b. We also use the term *disutility* to mean the negative of the utility function. So, when we say that type $\theta \in [0, \frac{1}{b}]$ recommends an action a and earns disutility $(a - \theta - 1)^2$, we mean that in the original state space, type $\theta b \in [0, 1]$ recommends action ab and earns disutility $(ab - \theta b - b)^2 = b^2(a - \theta - 1)^2$.

We first partition the scaled type space $[0, \frac{1}{b}]$ into four intervals, with endpoints $0, \theta_1, \theta_2, \theta_3, \frac{1}{b}$. The separable groups are as follows: at time $t = 0$, each type $\theta \in [0, \theta_1]$ pools with a partner $g(\theta) \in [\theta_2, \theta_3]$ to send a sequence of recommendations $(u_1(\theta), u_2(\theta))$, and then reveal the truth at time $t = 2$ iff the DM followed both initial recommendations. Each type $\theta \in [\theta_1, \theta_2]$ initially pools with a partner $h(\theta) \in [\theta_3, \frac{1}{b}]$ to recommend a sequence $(v_1(\theta), v_2(\theta))$, then revealing the truth at time $T - \tau$ ($\tau < T - 2$ a time parameter to be determined) iff the expert followed their advice.[14] For the purpose of this outline, take the endpoints $\theta_1, \theta_2, \theta_3$ as given, along with the partner functions $g : [0, \theta_1] \to [\theta_2, \theta_3]$; $h : [\theta_1, \theta_2] \to [\theta_3, \frac{1}{b}]$, and recommendation functions u_1, u_2, v_1, v_2. In Appendices A–C, we construct the equilibrium parameters and functions.

For notational purposes it is useful to further subdivide the expert types into three groups: I, *II*, and *III*. Group I consists of types $\theta^I \in [\theta_1, \theta_2]$ with their partners $h(\theta^I) \in [\theta_3, \frac{1}{b}]$. Group *II* consists of all types $\theta^{II} \in [0, \theta_1]$ whose initial recommendation coincides with that of some Group I pair, together with their partners $g(\theta^{II}) \in [\theta_2, \theta_3]$. Group *III* consists of all remaining types $\theta^{III} \in [0, \theta_1]$ and their partners $g(\theta^{III}) \in [\theta_2, \theta_3]$. In other words, we divide the types in intervals $[0, \theta_1] \cup [\theta_2, \theta_3]$ into two groups, *II* and *III*, according to whether or not their initial messages coincide with those of some group I pair.

The timeline of the expert's advice is as shown in Fig. 1 where $0 < \alpha_0 \leqslant \alpha_a < 1$ are specified in Appendices A–C (see Eq. (20)). It should be noted that the times at which the DM is instructed to change his action $(2\alpha_0, 2\alpha_a, T - \tau)$ are not necessarily integers in our construction. In a continuous-time setting, this clearly poses no problem; in discrete time, we can deal with integer constraints via public randomization and/or scaling up the horizon, as explained in Appendix C.4.

In words: in Phase 1, separable groups are formed. Each expert pair $(\theta^I, h(\theta^I))$ recommends $v_1(\theta^I)$, and each pair $(\theta^i, g(\theta^i))$ (with $i = 2, 3$) recommends an action $u_1(\theta^i)$. These initial recommendations overlap: for all $\theta^I \in [\theta_1, \theta_2]$, there exists $\theta^{II} \in [0, \theta_1]$ with $v_1(\theta^I) = u_1(\theta^{II})$.

[14] Note that u_1, u_2, v_1, v_2 are functions of θ, and that in our construction, the expert's messages ("recommendations") are equal to the actions that he wants the decision-maker to take. The DM can then infer the expert's separable group from his recommendation.

Therefore, the DM's information set contains four types $\{\theta^I, h(\theta^I), \theta^{II}, g(\theta^{II})\}$ after any equilibrium message sent by a group *I* or *II* pair, and two types, $\{\theta^{III}, g(\theta^{III})\}$, following all initial recommendations sent by Group *III* pairs. In Phase 2 of the timeline, beginning at time $t = 2\alpha_0$, all pairs $(\theta^I, h(\theta^I))$ switch to the recommendation function $v_2(\cdot)$, thus separating out from any Group *II* pairs $(\theta^{II}, g(\theta^{II}))$ who sent the same initial message. At this point, the DM's information set contains at most two types.[15] In Phase 3, beginning at time $2\alpha_a \geqslant 2\alpha_0$, Group *II* and *III* pairs switch to the recommendation function $u_2(\cdot)$; this conveys no new information to the DM, but we need at least two distinct pre-separation actions in order to provide the expert with appropriate incentives for eventually revealing the truth. During these phases, the DM is able to infer the separable group containing the expert's true type, but, rather than choosing the corresponding myopically optimal action, he chooses the actions recommended by the expert. These expert recommendations, in turn, were chosen to provide the expert with incentives to join the *right* separable group at time 0. Finally, Phases 4 and 5 are the revelation phases: the separable groups themselves separate, revealing the exact truth to the DM, provided that he has followed all of the expert's previous advice. If the DM ever fails to choose a recommended action, then the expert babbles during the revelation phase.

Incentivizing the expert. We now briefly explain the construction of the functions (u_1, u_2) and (v_1, v_2), and the corresponding partner functions g, h. For the expert, three sets of constraints must be satisfied.

The first set of constraints can be thought of as local incentive compatibility constraints – that is, those applying within each type θ's interval $[\theta_i, \theta_{i+1}]$. These (dynamic) constraints ensure that, say, the agent $\theta \in [0, \theta_1]$ prefers to induce actions $u_1(\theta)$ (for $2\alpha_a$ periods), $u_2(\theta)$ (for $2(1-\alpha_a)$ periods), and then reveal his type θ for the final $T-2$ periods, than e.g. to follow the sequence $(u_1(\theta'), u_2(\theta'), \theta')$ prescribed for some other type θ' in the same interval $[0, \theta_1]$ (and analogously within each of the other three intervals). For types $\theta \in [0, \theta_1]$, this boils down to a requirement that u_1, u_2 satisfy the following differential equation,

$$2\alpha_a u_1'(\theta)\big(u_1(\theta) - \theta - 1\big) + 2(1-\alpha_a)u_2'(\theta)\big(u_2(\theta) - \theta - 1\big) = T - 2 \tag{5}$$

and that the "average" action, $2\alpha_a u_1(\theta) + 2(1-\alpha_a)u_2(\theta) + (T-2)\theta$, be weakly increasing in θ.

Note that a longer revelation phase (that is, an increase in the RHS term $(T-2)$ in (5)) requires a correspondingly larger distortion in the action functions u_1, u_2 (they become larger and/or steeper): if the expert anticipates a lengthy phase in which the DM's action will match the true state (whereas the expert's bliss point is to the right of the truth), then it becomes more difficult in the initial phase to provide him with incentives not to mimic the advice of types to his right. This is why a longer horizon does not trivially imply better welfare properties.

The next set of constraints for the expert can be thought of as "global" incentive compatibility constraints, ensuring that no expert type wishes to mimic any type in any other interval. This turns out to impose two additional constraints: each endpoint type $\theta_1, \theta_2, \theta_3$ must be indifferent between the two equilibrium sequences prescribed for his type, and the time-averaged action must weakly increase at each endpoint.

The final constraint requires that each pair of types indeed be "separable": for any pair of types $\theta < \theta'$ who pool together during the first three phases, it must be that type θ would rather

[15] The purpose of ensuring that action functions u_1, v_1 overlap, so that all initial Group *I* messages coincide with the recommendation sent by a Group *II* pair, is that it is otherwise impossible to design strategies which ultimately reveal the true state, and which satisfy both players' incentive constraints.

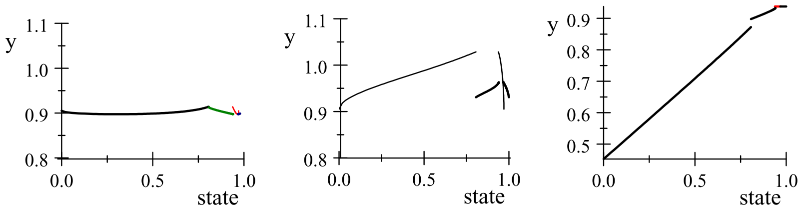
tell the truth, in which case the DM will choose action θ, than mimic his partner θ', for action θ'. In our rescaled action and state space, this reduces to the following requirement[16]:

$$(\theta - \theta - 1)^2 \leqslant \left(\theta' - \theta - 1\right)^2 \quad \Leftrightarrow \quad \theta' \geqslant \theta + 2. \tag{6}$$

That is, each of the pairs $(\theta^I, h(\theta^I))$ and $(\theta^i, g(\theta^i))$ $(i = II, III)$ must be at least 2 units apart.

It turns out to be very tricky to satisfy the global incentive compatibility constraints together with the local constraints. It requires a minimum of two distinct actions prior to the revelation phase (this is why e.g. Group *III* pairs must change their recommendation from u_1 to u_2 at time $2\alpha_a$, even though doing so reveals no further information), and that the type space be partitioned into a minimum of four intervals.[17] Moreover, for any partition into four intervals, we found only *one* partner function $g : [0, \theta_1] \to [\theta_2, \theta_3]$ that satisfied the global incentive requirements at all three interval endpoints. We believe (after extensive efforts to prove otherwise) that there is no partition which would allow for expert-incentivizing action functions which are myopically optimal from the DM's perspective. This is why our construction relies on trigger strategies: the expert *only* reveals the truth if the DM follows all of his advice.

We graph the equilibrium actions u_1, v_1 in the left-most graph, the u_2, v_2 in the middle graph, and the average action for $b = \frac{1}{60.885}$ and $T = 4$:



Actions u_1, v_1 Actions u_2, v_2 Average Action

Incentivizing the DM: Suppose that the expert recommends an action $u_1(\theta)$, which the DM believes could only have come from types $\theta, g(\theta)$. If the DM follows the recommendation, then he expects the expert to switch his recommendation to $u_2(\theta)$ at time $2\alpha_a$, and then recommend the true state θ for the final $T-2$ periods. If the DM assigns probabilities $p_\theta, 1 - p_\theta$ to types $\theta, g(\theta)$, then this yields an expected disutility of

$$p_\theta\left(2\alpha_a\left(u_1(\theta) - \theta\right)^2 + 2(1-\alpha_a)\left(u_2(\theta) - \theta\right)^2\right) + (1 - p_\theta)\left(2\alpha_a\left(u_1(\theta) - g(\theta)\right)^2 + 2(1-\alpha_a)\left(u_2(\theta) - g(\theta)\right)^2\right)$$

(noting that disutility in the final $T-2$ periods is zero). The problem is that the initial recommendations $u_1(\theta), u_2(\theta)$ do not coincide with the DM's myopically optimal action, $y^*(\theta) \equiv p_\theta\theta + (1 - p_\theta)g(\theta)$. We therefore employ *trigger strategies*: the expert only reveals the truth in the final stage if the DM follows his recommendations at the beginning of the game. If the DM

[16] The LHS is type θ's per-period disutility from inducing action θ, and the RHS is the per-period disutility from action θ'. The constraint for type θ' to not mimic θ is immediate from $\theta' > \theta$ and (1).

[17] This is explained briefly in our derivation of the equilibrium strategies, provided in an online appendix. Essentially, we first show that a partition into just two intervals makes it impossible to construct (total) payoff functions which incentivize the expert to follow a truthful strategy; necessarily, either a local or global IC constraint would be violated. Then, we show that the desired total payoff functions cannot be achieved with just one action; with two, we can manipulate both the variance and the expectation of the equilibrium actions.

ever rejects his advice, then the expert babbles for the rest of the game, and so the DM's disutility is at best

$$T \cdot \left[p_\theta \cdot \left(p_\theta\theta + (1-p_\theta)g(\theta) - \theta\right)^2 + (1-p_\theta) \cdot \left(p_\theta\theta + (1-p_\theta)g(\theta) - g(\theta)\right)^2\right].$$

So, for the equilibrium to work for the DM, we need to make sure that the benefit to learning the exact state, rather than just the separable group containing it, is large enough to compensate him for the cost of following the expert's initial recommendations, rather than deviating to the myopically optimal actions. This is what limits the priors for which our construction works, and imposes the upper bound $b \cong \frac{1}{61}$ on the bias. The construction works for the expert $\forall b < \frac{1}{16}$ (see footnote 28 in Appendix C.2).

Beliefs: After each expert recommendation, the DM calculates his posteriors using Bayes' rule. However, this requires some care in our model, since we are explicitly looking for an equilibrium in which finite sets of types pool together (while the prior can be described by a density over the state space $[0,1]$), and so the DM's information sets all contain measure-zero sets of types. If there is an interval $I \subset [0,1]$, a continuous message function $m : I \to \mathbb{R}$, and a continuous partner function $p : I \to ([0,1]\backslash I)$ with the property for all $x \in I$, types x and $p(x)$ pool together to send the message $m \equiv m(x) = m(p(x))$, then Bayes' rule implies that the DM's beliefs satisfy

$$\frac{\Pr(x|m)}{\Pr(p(x)|m)} = \lim_{\varepsilon\to 0} \frac{\Pr(\theta \in [x-\varepsilon, x+\varepsilon])}{\Pr(\theta \in [p(x-\varepsilon), p(x+\varepsilon)])} = \frac{f(x)}{f(p(x))} \cdot \frac{1}{|p'(x)|} \tag{7}$$

where f is the density associated with the DM's prior over the state space. This says that the likelihood of type x relative to $p(x)$ is equal to the unconditional likelihood ratio (determined by the prior), times a term which depends on the shape of the p-function, in particular due to its influence on the size of the interval of partner types $p(x)$ (for all $x \in I$) compared to the interval I.

5.2. Main result

Theorem 1. *Suppose that $\delta^E = \delta^{DM} = 1$ and that the preferences of the expert and of the DM are given by (1). For any bias $b \leqslant \frac{1}{61}$, there is an open set of priors F,*[18] *and a horizon T^*, for which a fully revealing equilibrium exists whenever $T \geqslant T^*$.*

Substantively, this theorem establishes an unexpected finding: even with a forward-looking expert and an infinite state space, there are equilibria in which the truth is revealed in finite time. We initially expected to prove the opposite result. Technically, the construction involves several innovative ideas that we expect to be useful in analyzing many dynamic games with persistent asymmetric information.

In our construction, the true state is revealed at either time 2 or time $T-\tau$, where $T-\tau$ can be chosen to be at most 5 (see (14)). Thus, the DM chooses his best possible action, equal to the true state, in all but the first few periods. It is tempting to conclude that a long horizon means our equilibrium approaches the complete-information one, but unfortunately this is not true when the DM and expert are equally patient. A long horizon also makes it difficult to incentivize the

[18] This is slightly strengthened from previous versions of the paper, which claimed only an infinite (rather than open) set of priors.

expert, requiring a proportionally larger distortion in the initial recommendation functions, and thereby imposing a proportionally larger cost to the DM (from having to follow such bad early advice in order to learn the truth).

It is true, however, that there is an equilibrium with close to the full-information payoffs if the horizon is sufficiently long, *and* if the DM is sufficiently patient compared to the expert. Moreover, for the priors and biases covered by Theorem 1, our construction can be modified (via a trivial rescaling of the timeline) to yield a fully revealing equilibrium for any pair of discount factors, so long as the DM is at least as patient as the expert. This is easiest to describe if we assume that the expert can revise his recommendation at any point in time. Letting r^E, r^{DM} denote the continuous-time discount rates for the expert and the DM (and interpreting the preferences in (1) as flow payoffs), leave all specifications from the proof of Theorem 1 unchanged, except for the timeline shown in Fig. 1: now, let Group I pairs recommend v_1 up to time $t_1(\alpha_0)$, then v_2 up to time t_4, and then reveal the truth, and let Group II, III pairs now recommend u_1 up to time $t_2(\alpha_a)$, u_2 up to time t_3, then reveal the truth, where

$$t_1(\alpha_0) = \frac{\ln(1 - 2\phi\alpha_0 r^E)}{-r^E}, \qquad t_2(\alpha_a) = \frac{\ln(1 - 2\phi\alpha_a r^E)}{-r^E},$$
$$t_3 = \frac{\ln(1 - 2\phi r^E)}{-r^E}, \qquad t_4 = \frac{\ln(1 - (T - \tau)\phi r^E)}{-r^E} \tag{8}$$

with $\phi = \frac{1 - e^{-r^E \widehat{T}}}{T r^E}$, $\widehat{T}$ is the (freely specified) horizon, and the T is the horizon used in our original construction.

By construction, this modification multiplies all expected payoffs from our original construction by a constant, ϕ. It can further be shown that the DM's incentive constraints are relaxed as he grows more patient – intuitively, his incentives to follow the expert's recommendations grow as he puts more weight on learning the truth in the future – and so we obtain a fully revealing equilibrium whenever the DM is more patient than the expert. As $\widehat{T} \to \infty$ and $r^{DM} \to 0$, the times in (8) remain finite, and so the DM (by following the expert's advice) ends up knowing the truth in all but the early stages of a very long game. He will then find it optimal to follow the expert's advice for nearly all priors over the state space, and earns (asymptotically) his full-information payoffs.[19]

6. Concluding remarks

This paper shows that dynamic strategic communication differs from its static counterpart. Our most striking result is that fully revealing equilibria exist. The equilibria are admittedly complex, and we do not suggest that they resemble any communication schemes currently in practice. This was not our goal; rather, we wished to determine whether it is *possible* for a DM to design a questions-and-incentives scheme to elicit the precise truth out of a biased expert, such that the expert would be willing to follow the proposed scheme. Our construction proves that it is

[19] This result was formally included in the previous version of the paper. Specifically, we found that for any bias $b < \frac{1}{16}$ (compared to the cutoff $b < \frac{1}{61}$ required in this paper for equal discount rates), any fixed expert discount rate $r^E > 0$, and any prior with densities that are everywhere bounded away from zero and infinity, one can choose a horizon long enough, and r^{DM} sufficiently close to zero, that the proposed strategies constitute a fully revealing equilibrium.

indeed possible, explains exactly how to do so when the expert has quadratic-loss preferences[20] and the true state is constant over time,[21] and highlights the conditions under which he would indeed desire to do so. In particular, the proposed communication scheme would benefit the DM if he is either more patient than the expert, or if final decisions are relatively more important for the decision-maker compared to earlier ones.

The main novel ingredient of our model is that there are multiple rounds of communication, with a new action chosen after each round. The dynamic incentive considerations for the expert allow us to group together types that are far apart, forming "separable groups", which is the key to obtaining greater information revelation. Our dynamic setup also allows for future communication to be conditioned on past actions (trigger strategies), and we show how information revelation can be facilitated through this channel.

The forces that we identify may be present in many dynamic environments with asymmetric information and limited commitment. In these models as well, past behavior sets the stage for future behavior. And, in contrast to the vast majority of the recent literature on dynamic mechanism design,[22] one needs to worry about both global and local incentive constraints, even with simple stage payoffs that satisfy the single-crossing property.

Lastly, given the important insights from cheap talk literature which have been widely applied in both economics and political science, we hope and expect that the novel aspects of strategic communication emphasized in our analysis will shed light on many interesting dynamic problems.

Appendix A. Proof of Proposition 2

When we restrict attention to monotonic partition equilibria, there will be some point in the game at which the last subdivision of an interval occurs, say period $\widehat{T} \leqslant T$. Assume that some interval is partitioned into two, inducing actions y_1 and y_2, and let $\hat{\theta}$ be the expert type who is indifferent between y_1, y_2. Since no subdivision occurs after period $\widehat{T}$, it follows that type $\hat{\theta}$'s indifference condition in period $\widehat{T}$ is

$$\left(1 + \delta + \cdots + \delta^{T-\widehat{T}-1}\right) u^E(y_1, \hat{\theta}, b) \geqslant \left(1 + \delta + \cdots + \delta^{T-\widehat{T}-1}\right) u^E(y_2, \hat{\theta}, b),$$

which reduces to the static indifference condition. But then, if this subdivision is possible, it cannot be the case that all static equilibria are equivalent to babbling equilibria. This follows by Corollary 1 of Crawford and Sobel [10].

Observe that all the arguments in this proof go through even if we allow for trigger strategies. This is because at the point where the last subdivision occurs, it is impossible to incentivize the DM to choose anything other than his myopic best response: he knows that no further information

[20] It would be interesting to understand more generally the types of expert preferences for which this is possible, but this is beyond the scope of the current paper. The general question is difficult to analyze, given the large class of possible equilibrium structures.

[21] One could presumably apply our construction in a model where the state evolves slowly over time, for example by restricting how frequently the expert can observe state changes, and playing our equilibrium within each "block" between state observations. If the probability of a state change between observations is small, this would lead to an equilibrium where the DM knows the true state most of the time.

[22] In recent years, motivated by the large number of important applications, there has been substantial work on dynamic mechanism design. See, for example, the survey of Bergemann and Said [7] and the references therein, or Pavan, Segal, and Toikka [24].

will be revealed, and so he knows that he cannot be rewarded in the future for choosing a suboptimal action now. So, the above argument applies.

Appendix B. Example 4: more partitions can reduce welfare

The following example demonstrates that equilibria with more partitions can be Pareto inferior to the equilibria with fewer partitions.

Take $\delta^E = \delta^{DM} = 1$ and $b = 0.08$, with the state θ drawn from a uniform distribution on [0, 1]. Consider the most informative static partition equilibrium where the number of partitions is $p = 3$. At this equilibrium the state space is divided into [0, 0.013], [0.013, 0.347] and [0.347, 1]. The corresponding optimal actions of the DM are given by

$$y_1 = 0.0067, \qquad y_2 = 0.18, \qquad y_3 = 0.673$$

from which we can calculate the ex-ante expected utility levels for the expert -0.0325 and for the DM -0.0263. Then, at the equilibrium of the dynamic game where the most informative static equilibrium is played at $t = 1$ and babbling thereafter, the total expected utility is -0.065 for the expert, and -0.053 for the DM.

We now construct a dynamic equilibrium where the type space is subdivided into more subintervals, but both players' ex-ante expected payoffs are lower. We look for an equilibrium with the following signaling rule: types in $[0, \theta_1]$ send message sequence $(m_{1(1)}, m_{2(1)})$, types in $(\theta_1, \theta_2]$ send message sequence $(m_{1(2)}, m_{2(2)})$, types in $(\theta_2, \theta_3]$ send message sequence $(m_{1(2)}, m_{2(3)})$, and types in $(\theta_3, 1]$ send message sequence $(m_{1(3)}, m_{2(4)})$. So types are partitioned into four intervals in stage 2, but in stage 1, the types in $[\theta_1, \theta_2]$ and $[\theta_2, \theta_3]$ pool together to send the same message $m_{1(2)}$. Since the signaling rule does not depend on the DM's action at stage 1, the DM will choose the following myopically optimal actions: $y_{1(1)} = y_{2(1)} = \frac{\theta_1}{2}$, $y_{1(2)} = \frac{\theta_1+\theta_3}{2}$, $y_{2(2)} = \frac{\theta_1+\theta_2}{2}$, $y_{2(3)} = \frac{\theta_2+\theta_3}{2}$, and $y_{1(3)} = y_{2(4)} = \frac{1+\theta_3}{2}$. After any out-of-equilibrium first-period message, the DM assigns probability one to the interval $[0, \theta_1]$, and chooses action $y_{1(1)}$. After any out-of-equilibrium second-period message following initial message $m_{1(1)}$ or $m_{1(3)}$, the DM maintains his first-period beliefs, choosing (respectively) $m_{2(1)}$ or $m_{2(4)}$. After any other out-of-equilibrium sequence of messages, the DM assigns probability one to the interval $[\theta_1, \theta_2]$, and chooses $m_{2(2)}$. With these out-of-equilibrium beliefs it is immediate to see that no type has an incentive to deviate.

At equilibrium, type θ_1 is indifferent between action sequences $(y_{1(1)}, y_{2(1)})$ and $(y_{1(2)}, y_{2(2)})$, type θ_2 is indifferent between 2nd-period actions $y_{2(2)}$ and $y_{2(3)}$, and type θ_3 is indifferent between action sequences $(y_{1(2)}, y_{2(3)})$ and $(y_{1(3)}, y_{2(4)})$. Therefore, equilibrium cutoffs are the solution to the following system of equations[23]:

$$2\left(\frac{\theta_1}{2} - \theta_1 - b\right)^2 - \left(\frac{\theta_1+\theta_3}{2} - b - \theta_1\right)^2 - \left(\frac{\theta_1+\theta_2}{2} - b - \theta_1\right)^2 = 0,$$

$$\left(\frac{\theta_1+\theta_2}{2} - b - \theta_2\right)^2 - \left(\frac{\theta_2+\theta_3}{2} - b - \theta_2\right)^2 = 0,$$

$$2\left(\frac{1+\theta_3}{2} - b - \theta_3\right)^2 - \left(\frac{\theta_1+\theta_3}{2} - b - \theta_3\right)^2 - \left(\frac{\theta_2+\theta_3}{2} - b - \theta_3\right)^2 = 0.$$

[23] It is trivial to check exactly as we did in previous examples that these indifference conditions suffice for global incentive compatibility.

At $b = 0.08$, the only solution that gives numbers in $[0, 1]$ is $\theta_1 = 0.0056$, $\theta_2 = 0.015$, $\theta_3 = 0.345$, and the actions induced for $t = 1$ and for $t = 2$ are respectively given by $y_{1(1)} = y_{2(1)} = 0.00278$, $y_{1(2)} = 0.175$, $y_{2(2)} = 0.0105$, $y_{2(3)} = 0.18$ and $y_{1(3)} = y_{2(4)} = 0.673$. This implies ex-ante expected utility -0.066 for the expert and -0.053 for the DM. Thus, although the interval is subdivided into more subintervals here, both players are strictly worse off than when the best static equilibrium is played in the first period, with players babbling thereafter. This feature is also illustrated in Example 1 of Blume, Board, and Kawamura [8].

Appendix C. Proof of Theorem 1

For brevity of exposition, we will prove Theorem 1 via the "guess-and-verify" method: Appendix C.1 gives the proposed strategies, Appendix C.2 proves that they are optimal from the expert's perspective, Appendix C.3 constructs an open set of priors for which the DM likewise finds it optimal to follow the proposed strategy. We provide the details behind the equilibrium construction in an online appendix. Additionally, for ease of exposition, we assume in Appendices C.1–C.3 that time is continuous, so that messages may be sent and actions may be changed at any point in time. We explain at the end of Appendix C.4 how to modify our timeline for discrete time.

The precise details of our construction differ depending on whether the bias is above or below $\frac{1}{320}$. We provide here the strategies for both cases, but defer some details for the case $b < \frac{1}{320}$ to the online appendix.

C.1. Preliminaries: strategies, timeline, parametrizations

Type parametrizations: For any bias $b < \frac{1}{61}$, partition the (scaled) state space $[0, \frac{1}{b}]$ into four intervals, $[0, \theta_1) \cup [\theta_1, \theta_2] \cup (\theta_2, \theta_3) \cup [\theta_3, \frac{1}{b}]$, with endpoints $\theta_1, \theta_2, \theta_3$ determined by b as follows: first define a parameter $a_b < 0$[24] by

$$\left(a_b - 2 + 2e^{-a_b}\right)e^2 - a_b = \frac{1}{b} \tag{9}$$

and then set

$$\theta_3 = \frac{1}{b} + a_b, \qquad \theta_2 = \theta_3 - 2, \qquad \theta_1 = \theta_2 - \theta_3 e^{-2}. \tag{10}$$

We describe the types in these four intervals parametrically, via functions $x : [-2, 0] \to [0, \theta_1]$, $g : [-2, 0] \to [\theta_2, \theta_3]$, $z : [a_b, 0] \to [\theta_1, \theta_2]$, and $h : [a_b, 0] \to [\theta_3, \frac{1}{b}]$ given by[25]

$$\begin{aligned} x(a) &= \theta_3 + a - \theta_3 e^a, \qquad & g(a) &= \theta_3 + a, \\ z(a) &= \frac{1}{b} + a - 2e^{a-a_b}, \qquad & h(a) &= \frac{1}{b} + a. \end{aligned} \tag{11}$$

Timeline: The timeline involves the following parameters: a horizon T, a time $2 < \tau < T$, and a continuous, weakly decreasing function $\alpha : [-2, 0] \to (0, 1)$; in a slight abuse of notation, we

[24] A straightforward calculation shows that the LHS expression in (9) is strictly decreasing in a_b, and equal to zero at $a_b = 0$; thus, (9) indeed defines a unique value $a_b < 0$ for any $b > 0$.

[25] Observe that x and z are strictly decreasing in a, while h and g are strictly increasing in a, with (by (9) and (10)) $x(0) = 0$, $x(-2) = \theta_1 = z(0)$, $z(a_b) = \theta_2 = g(-2)$, and $g(0) = \theta_3 = h(a_b)$.

define $\alpha_a \equiv \alpha(a)$, $\forall a \in [-2, 0]$. The pair of expert types $(x(a), g(a))$ switches from u_1 to u_2 at time $2\alpha_a$, then reveals the truth at time 2; all pairs $(z(a), h(a))$ switch from v_1 to v_2 at time $2\alpha_0$, and then reveal the truth at time τ. *Importantly, for the case $\frac{1}{320} < b < \frac{1}{61}$ we focus on here, the function α is constant, with $\alpha_a = \alpha_0$ for all $\alpha \in [-2, 0]$.* The time $T - \tau$ at which pairs $(z(a), h(a))$ reveal the truth is determined by the horizon as follows:

$$\frac{\tau}{T-2} = \beta \equiv \frac{(\theta_2-\theta_1)(\theta_2-\theta_1-2)}{(\theta_4-\theta_1)(\theta_4-\theta_1-2)} \tag{12}$$

$$= \frac{(a_b - 2 + 2e^{-a_b})(a_b - 4 + 2e^{-a_b})}{2e^{-a_b}(2e^{-a_b} - 2)} \quad \text{by (10) and (11)} \tag{13}$$

and our proofs for the DM require a horizon $T \in [\underline{T}, \overline{T}]$, where[26,27]

$$\underline{T} = \begin{cases} 7 & \text{if } \beta \in [0.4173, 0.50102), \\ \frac{5-2\beta}{1-\beta} & \text{if } \beta \in [0.50102, 0.79202), \\ \frac{5.4748\beta}{2.7374\beta - 1.7374} & \text{if } \beta \in [0.79202, 0.95203), \\ 6 & \text{if } \beta \geqslant 0.95203, \end{cases}$$

$$\overline{T} = \begin{cases} 7 & \text{if } \beta \in [0.4173, 0.50102), \\ \frac{8-2\beta}{1-\beta} & \text{if } \beta \in [0.50102, 0.79202), \\ \frac{4-2\beta}{1-\beta} & \text{if } \beta \in [0.79202, 0.90913), \\ \frac{12.005\beta}{6.0025\beta - 5.0025} & \text{if } \beta \geqslant 0.90913. \end{cases} \tag{14}$$

Expert's strategy (*on-path*): The expert's strategy along the equilibrium path is as follows: each expert pair $(x(a), g(a))$ with $a \in (-2, 0]$ (covering all types in $[0, \theta_1) \cup (\theta_2, \theta_3]$) recommends an action $u_1(a)$ at time zero, then switches the recommendation to $u_2(a)$ at time $2\alpha_a$, and then reveals the true state at time 2, where

$$u_1(a) = \theta_3 + K - \frac{T-2}{2}a - \sqrt{\frac{1-\alpha_a}{\alpha_a}}\sqrt{\frac{1-\alpha_0}{\alpha_0}K^2 + (T-2)a\left(K - \frac{T}{4}a\right)}, \tag{15}$$

$$u_2(a) = \theta_3 + K - \frac{T-2}{2}a + \sqrt{\frac{\alpha_a}{1-\alpha_a}}\sqrt{\frac{1-\alpha_0}{\alpha_0}K^2 + (T-2)a\left(K - \frac{T}{4}a\right)} \tag{16}$$

and

$$K = \frac{\alpha_0 \tau a_b \left(1 + \sqrt{\frac{(T-2\alpha_0)(T-\tau)}{2\tau\alpha_0}}\right)}{(T - \tau - 2\alpha_0)}. \tag{17}$$

[26] The upper limit on T (relative to our normalization that Group *II*, *III* pairs reveal the truth at time 2) arises for the reasons explained following Eq. (7): a longer horizon makes it more difficult to provide the expert with incentives to reveal the truth, implying that the initial actions must be distorted further away from those which are myopically optimal for the DM. Throughout most of the state space, the DM's IC constrains are nonetheless relaxed as T increases; however, for the closest-together pooled expert pairs (e.g. (θ_2, θ_3), separated by only $2b$ units), knowing the exact state for a larger fraction of the game does not compensate for this distortion, and the DM will deviate to the myopically optimal action if T is too large.

[27] For future reference, note from (13) that $\beta a_b^2 \geqslant 8 \Leftrightarrow a_b \lesssim -3.18 \Leftrightarrow \beta \geqslant 0.79202$, and that in this range, (14) specifies $\overline{T} \leqslant \frac{4-2\beta}{1-\beta}$; using (12), this implies $T - \tau \leqslant 4$. In the range $a_b \in [-2, -3.18) \Leftrightarrow \beta \in [0.50102, 0.79202)$, (14) implies that $T - \tau \in [5, 8]$, noting from (12) that $T = \frac{T-\tau-2\beta}{1-\beta}$.

All expert pairs $(z(a), h(a))$ with $a \in [a_b, 0]$ except for type $h(a_b) = \theta_3$ (covering all expert types in $[\theta_1, \theta_2] \cup (\theta_3, \frac{1}{b}])$ recommend $v_1(a)$ at time zero, switch their recommendation to $v_2(a)$ at time $2\alpha_0$, and then reveal the truth at time τ, where

$$v_1(a) = \theta_3 + \frac{2K - \tau(a - a_b)}{T - \tau} - \frac{\sqrt{\frac{\tau(T-\tau-2\alpha_0)}{\alpha_0}}\sqrt{(\frac{T-\tau-2\alpha_0}{\tau\alpha_0})K^2 + 2K(a-a_b) - \frac{T}{2}(a-a_b)^2}}{T - \tau}, \tag{18}$$

$$v_2(a) = \theta_3 + \frac{2K - \tau(a - a_b)}{T - \tau} + \frac{\sqrt{\frac{4\tau\alpha_0}{T-\tau-2\alpha_0}}\sqrt{(\frac{T-\tau-2\alpha_0}{\tau\alpha_0})K^2 + 2K(a-a_b) - \frac{T}{2}(a-a_b)^2}}{T - \tau}. \tag{19}$$

Note that type $\theta_2 = z(a_b)$ is missing his "partner" $h(a_b)$, as we have specified that type $\theta_3 = g(0) = h(a_b)$ follow the strategy prescribed for type $g(0)$ rather than the one that would be prescribed for type $h(a_b)$. This will not pose any problem for the DM, due to the fact that the two strategies, by construction, are identical.[28]

Expert's strategy (*off-path*): If the DM ever deviates, by choosing a different action than the one recommended by the expert, then (i) if the expert himself has *not* previously deviated, he subsequently babbles; (ii) if the expert *has* observably deviated in the past, he subsequently behaves as if the deviation did not occur.

DM's strategy and beliefs: If there have been no detectable deviations by the expert, then follow all recommendations, using Bayes' rule to assign beliefs at each information set. Following deviations: (i) If the expert observably deviates at time 0 (sending an off-path initial recommendation), subsequently adopt the strategy/beliefs that would follow if the expert had instead sent the recommendation $u_1(0)$ prescribed for types $\{x(0), g(0)\}$; (ii) If the expert observably deviates on his 2nd recommendation, ignore it as an error, and subsequently adopt the strategy/beliefs that would follow had the deviation not occurred; (iii) If the expert deviates observably in the revelation phase, ignore it as an error, assigning probability 1 to the lowest type in the current information set; (iv) If the DM himself deviates, rejecting some expert recommendation, he subsequently maintains his current (time of deviation) beliefs, disregarding any expert messages as uninformative babbling.

C.1.1. Preliminary calculations

Before proceeding with the proof that these strategies indeed constitute a fully revealing equilibrium, we construct a function $\alpha(\cdot)$ with the following properties:

[28] Type $h(a_b)$ would recommend $v_1(a_b)$ initially, $v_2(a_b)$ at time $2\alpha_0$, and then the truth, θ_3, at time $T - \tau$, and by (18) and (19), we have $v_1(a_b) = \theta_3 + \frac{K}{\alpha_0}$, $v_2(a_b) = \theta_3$; on the other hand, type $g(0)$ would recommend $u_1(0)$ initially, then $u_2(0)$ at time $2\alpha_0$, and then θ_3 at time 2, and by (15) and (16), we have $u_1(0) = \theta_3 + \frac{K}{\alpha_0}$, $u_2(0) = \theta_3$. So with either specification, type θ_3 recommends $\theta_3 + \frac{K}{\alpha_0}$ for the first $2\alpha_0$ periods, and θ_3 from then on.

$$\begin{cases} \text{(i)} & \text{if } \beta a_b^2 \leqslant 8, \text{ then } \alpha(\cdot) \text{ is constant, with } \alpha(a) = \alpha_0, \forall a \in [-2,0], \text{ and } \alpha_0 \text{ near } 1, \\ \text{(ii)} & \text{if } \beta a_b^2 > 8, \text{ then } \alpha(\cdot) \text{ is continuous and strictly decreasing, with} \\ & \alpha(0) \equiv \alpha_0 \text{ near } 0, \\ \text{(iii)} & u_1, u_2, v_1, v_2 \text{ are real-valued, and } \forall a \in [a_b, 0], \ \exists a' \in [-2,0] \text{ with} \\ & v_1(a) = u_1(a'). \end{cases} \tag{20}$$

Lemmas A and B below prove that if $\beta a_b^2 \leqslant 8$, then there is a cutoff $\bar{\alpha}_0 < 1$ such that if α is constant, with $\alpha(a) = \alpha_0 \in (\bar{\alpha}_0, 1)$, then property (iii) of (20) is satisfied. We provide details for the case $\beta a_b^2 > 8$ in the online appendix.

Lemma A *(Expert's recommended actions are real-valued). Let $\beta a_b^2 \leqslant 8$, and set $\alpha_a = \alpha_0$, $\forall a \in [-2,0]$. There exists $0 < \overline{\alpha_0} < 1$ such that the action functions u_j, v_j specified in (15)–(19) are real-valued whenever $\alpha_0 > \bar{\alpha}_0$.*

Proof. For $v_j(\cdot)$, we need to prove that the following expression is non-negative for all $a \in [a_b, 0]$:

$$\left(\frac{T-\tau-2\alpha_0}{\tau}\right)\frac{K^2}{\alpha_0} + 2K(a-a_b) - \frac{T}{2}(a-a_b)^2. \tag{21}$$

Since this expression is strictly decreasing in a (noting from (17) that $K < 0$), it is sufficient to prove that the minimum value, at $a = 0$, is non-negative. But this is true by construction: the value of K specified in (17) is precisely the negative root of the equation $(\frac{T-\tau-2\alpha_0}{\tau})\frac{K^2}{\alpha_0} - 2Ka_b - \frac{T}{2}a_b^2$.

For $u_j(\cdot)$ we need to prove that the following expression is non-negative $\forall a \in [-2,0]$:

$$\frac{1-\alpha_0}{\alpha_0}K^2 + (T-2)a\left(K - \frac{T}{4}a\right). \tag{22}$$

This expression is strictly concave (the second derivative w.r.t. a is $-\frac{T(T-2)}{2}$), hence minimized at one of the endpoints of the interval $[-2,0]$. At $a = 0$, it is trivially non-negative, and so we need only show that it is non-negative at $a = -2$, requiring:

$$\frac{1-\alpha_0}{\alpha_0}K^2 - (T-2)(2K+T) \geqslant 0. \tag{23}$$

For this, it suffices (by continuity in α_0) to show that $\lim_{\alpha_0 \to 1}(2K+T) < 0$. But by (17) and the first expression in (12),

$$\left(\frac{1-\beta}{2\beta}\right)\lim_{\alpha_0 \to 1}(2K+T) = a_b\left(1 + \sqrt{\frac{T-\tau}{2\beta}}\right) + \frac{T-\tau}{2\beta} - 1.$$

This is negative for all $\sqrt{\frac{T-\tau}{2\beta}} < 1 - a_b$, which is implied by (14): If $\beta \in [0.4172, 0.50102]$, then $1 - a_b > 2.7726$, while $\sqrt{\frac{T-\tau}{2\beta}} \leqslant \sqrt{\frac{7-5(0.4172)}{2(0.4172)}} = 2.4268$; and if $\beta \in [0.50102, 0.79202]$, then $1 - a_b > 3$, while $T - \tau \leqslant 8$ and $\beta > 0.50102$ imply $\sqrt{\frac{T-\tau}{2\beta}} \leqslant \sqrt{\frac{8}{2(0.50102)}} < 3$. □

Lemma B *(Group I, II recommendations overlap). Let $\beta a_b^2 \leqslant 8$ and set $\alpha_a = \alpha_0$, $\forall a \in [-2,0]$. There exists numbers $0 < \overline{\alpha_0} < 1$ such that for all $\alpha_0 > \bar{\alpha}_0$, and for any $a \in [a_b, 0]$, there exists $a' \in [-2,0]$ such that $v_1(a) = u_1(a')$.*

Proof. By (18), we have

$$v_1''(a) = \frac{\sqrt{\frac{\tau(T-\tau-2\alpha_0)}{\alpha_0}}(\frac{K^2}{\tau\alpha_0})(\frac{T}{2}-\alpha_0)}{((\frac{T-\tau-2\alpha_0}{\tau\alpha_0})K^2+2K(a-a_b)-\frac{T}{2}(a-a_b)^2)^{\frac{3}{2}}} > 0,$$

noting from (14) that $\frac{T}{2} > \alpha_0$, so that $\min_{a\in[a_b,0]} v_1'(a) = v_1'(a_b) = 0$. Therefore v_1 is strictly increasing, with range (using (18))

$$a \in [a_b, 0] \quad \Rightarrow \quad v_1(a) \in \left[v_1(a_b), v_1(0)\right] = \left[\theta_3 + \frac{K}{\alpha_0}, \theta_3 + \frac{2K+\tau a_b}{T-\tau}\right]. \tag{24}$$

By (15) evaluated at $a = 0$, we have $u_1(0) = \theta_3 + \frac{K}{\alpha_0}$, precisely the minimum value of $v_1(\cdot)$ by (24). It remains to prove that for α_0 sufficiently near 1, $u_1(-2)$ exceeds the maximum value of $v_1(a)$. Evaluating (15) at $a = -2$ and taking limits,

$$\begin{aligned}\lim_{\alpha_0\to 1}\left(u_1(-2) - \left(\theta_3 + \frac{2K+\tau a_b}{T-\tau}\right)\right) &= \lim_{\alpha_0\to 1}\left(K + T - 2 - \frac{2K+\tau a_b}{T-\tau}\right)\\ &= (T-2)\left(1 - \sqrt{\frac{\beta a_b^2}{2(T-\tau)}}\right) \quad \text{by (12), (17).}\end{aligned}$$

Since $\beta a_b^2 \leqslant 8$, and $T - \tau \geqslant 4$ in this range by (14), we have $\frac{\beta a_b^2}{2(T-\tau)} < 1$, thus the expression is strictly positive, as desired. □

C.2. Optimality for the expert

Proposition C1 *(Expert optimality: off-path behavior). The expert has no incentive to choose an off-path recommendation sequence.*

Proof. Immediate from the DM strategy and beliefs specified in Appendix C.1: a deviation at time zero is equivalent to mimicking type $x(0)$, deviations between time $t = 0$ and the revelation phase are ignored (the DM behaves as if he had instead received the anticipated recommendation sequence), and a deviation in the revelation phase is equivalent to mimicking the lowest type in the DM's current information set. □

Proposition C2 *(Expert optimality: truth revelation phase). In the prescribed revelation phase,* (i) *if there have been no previous deviations by the DM, then the expert finds it optimal to reveal the truth*; (ii) *if the DM has ever deviated, then the expert finds it optimal to babble* (*e.g. by repeating his last recommendation*).

Proof. Part (ii) follows from Appendix C.1 specification that if the DM himself ever deviates, then he will subsequently choose whichever action was myopically optimal at the time of deviation. The expert therefore cannot influence the DM's behavior, and so babbling (in particular) is optimal. For part (i): by (6), we just need to make sure that all pairs $(x(a), g(a))$, along with all pairs $(z(a), h(a))$, are separated by at least 2 units. By (10), $\min_{a\in[a_b,0]} |h(a) - z(a)| =$

$\theta_3 - \theta_2 = 2$; and by (10) and (11), $\min_{a\in[-2,0]} |g(a) - x(a)| = a_b - 2 + 2e^{-a_b}$, which exceeds 2 whenever $a_b < -0.895 \Leftrightarrow b < \frac{1}{15.76}$.[29] □

It remains to show that each type $\theta \in [0, \frac{1}{b}]$ would rather send the message sequence designated for his type, than mimic the sequence of any other type $\theta \neq \theta' \in [0, \frac{1}{b}]$. Throughout this section, we refer to our four intervals $[0, \theta_1), [\theta_1, \theta_2], (\theta_2, \theta_3], (\theta_3, \frac{1}{b}]$ as (respectively) I_1, I_2, I_3, I_4, and define $I(\theta) \in \{I_1, I_2, I_3, I_4\}$ as the interval containing type θ. Let $D(\theta'|\theta)$ denote the disutility to type $\theta \in [0, \frac{1}{b}]$ from following the strategy prescribed for type $\theta' \in [0, \frac{1}{b}]$ (recalling that we defined disutility as the negative of the payoff, so the expert's goal is to minimize disutility). We also define $D(\theta_i^+|\theta) \equiv \lim_{\theta' \downarrow \theta_i^+} D(\theta'|\theta)$ as the limit of $D(\theta'|\theta)$ as θ' approaches θ_i from the right, and $D(\theta_i^-|\theta) \equiv \lim_{\theta' \uparrow \theta_i^-} D(\theta'|\theta)$ as the limit from the left, for $i = 1, 2, 3$. We begin with two preliminary lemmas, before completing the proof in Proposition C3.

Lemma C3.1 *(Expert optimality: local IC). For all $\theta, \theta' \in [0, \frac{1}{b}]$, $D(\theta'|\theta)$ is strictly increasing in $|\theta' - \theta|$ if $\theta' \in I_1 \cup I_2$, and constant if $\theta' \in I_3 \cup I_4$.*

Proof. First consider type θ's disutility from following the strategy prescribed for a type $\theta' = x(a) \in [0, \theta_1]$. By (15) and (16),

$$D\big(x(a)|\theta\big) = 2\left(\theta_3 + K - \frac{T-2}{2}a - \theta - 1\right)^2 + 2\left(\frac{1-\alpha_0}{\alpha_0}K^2 + (T-2)a\left(K - \frac{T}{4}a\right)\right) + (T-2)\big(x(a) - \theta - 1\big)^2. \tag{25}$$

Differentiating with respect to $x(a)$, using $x'(a) = 1 - \theta_3 e^a$ (by (11)) and simplifying,

$$\begin{aligned}\frac{\partial D(x(a)|\theta)}{\partial x(a)} &= 2(T-2)\left(-\frac{\theta_3 + a - \theta - 1}{x'(a)} + \big(x(a) - \theta - 1\big)\right) \\ &= 2(T-2)\left(\frac{\theta_3 e^a}{\theta_3 e^a - 1}\right)\big(x(a) - \theta\big).\end{aligned} \tag{26}$$

As desired, this is positive if $x(a) > \theta$ (so expert type θ's disutility increases – making him worse off – if he mimics types $x(a)$ further above him), and negative if $x(a) < \theta$, establishing (i).

If type θ follows the strategy prescribed for type $z(a) \in I_2$, then, using (18) and (19), his disutility is given by

$$\begin{aligned}D\big(z(a)|\theta\big) &= 2\alpha_0\big(v_1(a) - \theta - 1\big)^2 + (T - t - 2\alpha_0)\big(v_2(a) - \theta - 1\big)^2 + \tau\big(z(a) - \theta - 1\big)^2 \\ &= \tau\big(z(a) - \theta - 1\big)^2 + (T-\tau)\left(\theta_3 + \frac{2K - \tau(a - a_b)}{T - \tau} - \theta - 1\right)^2 \\ &\quad + \frac{2\tau}{T-\tau}\left(\left(\frac{T - \tau - 2\alpha_0}{\tau\alpha_0}\right)K^2 + 2K(a - a_b) - \frac{T}{2}(a - a_b)^2\right).\end{aligned} \tag{27}$$

[29] This is in fact all that is needed for the construction to work for the expert, but we specify $b < \frac{1}{61}$ in (9) to make the construction work for the DM.

The derivative w.r.t. $z(a)$ is

$$\frac{\partial D(z(a)|\theta)}{\partial z(a)} = -2\tau \frac{(\theta_3 + a - a_b - \theta - 1)}{z'(a)} + 2\tau\big(z(a) - \theta - 1\big)$$
$$= 4e^{a-a_b}\left(\frac{\tau}{2e^{a-a_b} - 1}\right)\big(z(a) - \theta\big) \quad \text{by (11) and (10).} \tag{28}$$

This is positive iff $z(a) > \theta$, establishing that $D(\theta'|\theta)$ is strictly increasing in $|\theta' - \theta|$ if $\theta' \in (\theta_1, \theta_2)$. Part (ii) then follows immediately.

Next, if type θ follows the strategy prescribed for $g(a) \in [\theta_2, \theta_3]$, his disutility is as given by (25), just replacing $x(a)$ with $g(a)$ in the final term; the derivative w.r.t. $g(a)$ is then as in (26), just replacing $x'(a)$ with $g'(a)$, and $x(a)$ with $g(a)$. Therefore, we have

$$\frac{\partial D(g(a)|\theta)}{\partial g(a)} = 2(T-2)\left(-\frac{\theta_3 + a - \theta - 1}{g'(a)} + g(a) - \theta - 1\right)$$
$$= 0 \quad \text{by (11), since } g'(a) = 1 \text{ and } g(a) = \theta_3 + a.$$

Therefore, as desired, the disutility to any type $\theta \in [0, \frac{1}{b}]$ from mimicking type $g(a) \in (\theta_2, \theta_3)$ is a constant, independent of the particular type $g(a)$ chosen. And finally, type θ's disutility from mimicking type $\theta' \in (\theta_3, \frac{1}{b}]$ is as in (27), just replacing $z(a)$ with $h(a)$. Using $h(a) = \theta_4 + a$, this yields $\frac{\partial D(h(a)|\theta)}{\partial h(a)} = 0$, so that $D(\theta'|\theta)$ is constant for $\theta' \in [\theta_3, \frac{1}{b}]$, thus proving (iii). □

Lemma C3.2 *(Expert optimality: endpoints). Payoffs at the endpoints $\theta_1, \theta_2, \theta_3$ satisfy*

$$D\big(\theta_1^-|\theta\big) < D\big(\theta_1^+|\theta\big) \quad \Leftrightarrow \quad \theta < \theta_1, \tag{29}$$
$$D\big(\theta_2^-|\theta\big) < D\big(\theta_2^+|\theta\big) \quad \Leftrightarrow \quad \theta < \theta_2, \tag{30}$$
$$D\big(\theta_3^-|\theta\big) = D\big(\theta_3^+|\theta\big), \quad \forall\theta. \tag{31}$$

Proof. For (29), evaluate (27) at $a = 0$ to obtain an expression for $D(\theta_1^+|\theta)$, and evaluate (25) at $a = -2$ to obtain an expression for $D(\theta_1^-|\theta)$. Subtracting, using (17) to replace $(\frac{T-\tau-2\alpha_0}{\tau\alpha_0})K^2$ with $2Ka_b + \frac{T}{2}a_b^2$, (12) to replace τ with $\frac{(\theta_2-\theta_1)(\theta_2-\theta_1-2)}{(\theta_4-\theta_1)(\theta_4-\theta_1-2)}(T-2)$, and (10) to replace $\theta_3 - a_b$ with θ_4, this yields

$$\frac{D(\theta_1^-|\theta) - D(\theta_1^+|\theta)}{T-2} = \left(\frac{2(\theta_2 - \theta_1)(\theta_4 - \theta_2)}{(\theta_4 - \theta_1 - 2)}\right) \cdot (\theta - \theta_1).$$

This is negative iff $\theta < \theta_1$, thus proving (29).

To prove (30), obtain an expression for $D(\theta_2^-|\theta)$ by evaluating (27) at $a = a_b$, and obtain an expression for $D(\theta_2^+|\theta)$ by evaluating (25) at $a = -2$ (this gives $D(x(-2)|\theta)$), and then replacing $x(-2)$ with $g(-2) = \theta_2$. Subtract the expressions, and note that K cancels out (by construction). Then, using $\tau = \beta(T-2)$ and $\theta_3 = \theta_2 + 2$ (from (12) and (10)), we are left with the following expression, negative (as desired) iff $\theta < \theta_2$:

$$\frac{D(\theta_2^-|\theta) - D(\theta_2^+|\theta)}{(T-2)} = 4\beta(\theta - \theta_2) \quad \text{by (11).} \tag{32}$$

Finally, for (31), recall from Lemma C3.1 that $D(g(0)|\theta) = D(g(-2)|\theta)$, and that $D(h(a_b)|\theta)$ equals $D(z(a_b)|\theta) + \tau(\theta_3 - \theta_2)(\theta_2 + \theta_3 - 2\theta - 2)$. Subtracting,

$$D(\theta_3^-|\theta) - D(\theta_3^+|\theta) = D(g(-2)|\theta) - D(z(a_b)|\theta) - \tau(\theta_3 - \theta_2)(\theta_2 + \theta_3 - 2\theta - 2)$$
$$= 4\tau(\theta_2 - \theta) - \tau(\theta_3 - \theta_2)(\theta_2 + \theta_3 - 2\theta - 2) \quad \text{by (32), (12).}$$

By (10) this reduces to zero, as desired to establish that all types $\theta \in [0, \frac{1}{b}]$ are indifferent between the strategies prescribed for types $g(0), h(a_b)$. □

Proposition C3 *(Expert optimality: global IC). For all $\theta \in [0, \frac{1}{b}]$, the disutility to expert type θ from following the strategy prescribed for type $\theta' \in [0, \frac{1}{b}]$ is minimized at the truth, $\theta' = \theta$.*

Proof. This is almost immediate from Lemmas C3.1 and C3.2, which established that type θ's disutility from mimicking any type θ' is increasing in $|\theta' - \theta|$, and thus minimized (so utility is maximized) by following the truthful strategy. Consider, in particular, a type $\theta \in I_2$. By Lemma C3.1, truth-telling is better than mimicking any other type $\theta' \in I_2$, and so in particular, earns a disutility which is weakly below both $D(\theta_1^+|\theta)$ and $D(\theta_2^-|\theta)$ (the disutilities from following the strategies at the left and right endpoints of interval I_2). By (29), our type $\theta \in [\theta_1, \theta_2]$ prefers the strategy at the left endpoint of I_2 to the one at the one at the right endpoint of I_1, which in turn is preferred to the strategy of any other type in I_1; that is, $D(\theta|\theta) \leqslant D(\theta_1^+|\theta) < D(\theta_1^-|\theta) = \min_{\theta' \in I_1} D(\theta'|\theta)$, and so type θ will not mimic any type $\theta' \in I_1$. And (30), our type $\theta \in I_2$ prefers the strategy at the right endpoint of I_2 to the one at the left endpoint of I_3, which, by Lemma C3.1 and (31) yields the same disutility as mimicking any other type $\theta' \in I_3 \cup I_4$; therefore, mimicking such a type is not optimal. Altogether, this establishes that type θ_2 would rather follow the truthful strategy than mimic any other type.

The proofs for the remaining three intervals are nearly identical. □

C.3. Optimality for the DM

Throughout this section, assume that the expert's strategy is as specified in Appendix C.1, with α_0 chosen according to (20). The DM's off-path strategies and beliefs (specified in Appendix C.1) trivially satisfy all PBE requirements. Therefore, we need only prove that along the equilibrium path, for some open set of priors over the state space, a Bayesian DM will find it optimal to follow all of the expert's recommended actions.

First, some notation. Let $[v_j^{\min}, v_j^{\max}]$, $[u_j^{\min}, u_j^{\max}]$ $(j = 1, 2)$ denote the ranges of the functions u_j, v_j. By property (20), we have $v_1^{\min} = u_1^{\min}$, and $v_1^{\max} \leqslant u_1^{\max}$. Therefore, if the DM receives an initial recommendation $v_1(a) \in [v_1^{\min}, v_1^{\max}]$, he believes that it was sent by a type in $\{z(a), h(a), x(a'), g(a')\}$, where $a' = u_1^{-1}(v_1(a))$; let $(r_1(a), r_2(a), p_1(a), p_2(a))$ denote the DM's posterior probabilities (summing to 1) on these four types in his information set, and recall that $r_2(a_b) = 0$.[30] If the DM receives an initial recommendation $u_1(a) \in [v_1^{\max}, u_1^{\max}]$, his information set contains only the pair $(x(a), g(a))$; let $(p_1(a), p_2(a))$ (with $p_1(a) + p_2(a) = 1$) denote his posteriors on these two types.

The structure of the proof is as follows: Proposition C4 shows that it is sufficient to rule out profitable deviations at times $t \in \{0, 2\alpha_0, 2\}$, to the action which is myopically optimal at time t. Proposition C5 constructs a set of posteriors which rule out profitable deviations at $t = 0$, Propositions C6.1, C6.2 construct posteriors which rule out profitable deviations at times $t \in \{2\alpha_0, 2\}$,

[30] The recommendation $v_1(a_b) = u_1(0)$ is sent only by the three types $\{x(0), z(a_b), g(0)\}$; type $\theta_3 = h(a_b) = g(0)$ follows the strategy prescribed for type $g(0)$ rather than $h(a_b)$.

and Proposition C7 proves that there is an open set of priors over the state space which generate, as Bayesian posteriors, the beliefs required by Propositions C5, C6.1, C6.2.

Proposition C4 *(Strongest incentives to deviate). If the DM cannot gain by deviating to the myopically optimal action in any of the following three scenarios, then there are no profitable deviations*: (i) *at time $t=0$ after a recommendation in* $[u_1^{\min}, u_1^{\max}]$; (ii) *at time $t=2\alpha_0$, if the DM does not change his recommendation, or recommends an action in* $[v_2^{\min}, v_2^{\max}]$; (iii) *at time $t=2$, if the expert recommends θ_3.*

Proof. Prior to the revelation phase, types $\theta \notin \{0, \theta_2, \theta_3\}$ reveal information only at times $t=0$ (with the initial recommendation) and $t=2\alpha_0$ (when Group I pairs separate out from any Group II types who sent the same initial recommendation).[31] Types $\theta \in \{0, \theta_2, \theta_3\}$ all pool together until time $t=2$ (recommending $u_1(0)=v_1(a_b)$ initially, then $u_2(0)=v_2(a_b)=\theta_3$ at time $2\alpha_0$), at which point type 0 separates out by revealing the truth. The result then follows immediately from two observations. First, recall that the expert's strategy is to babble until the end of the game if the DM ever deviates. Then, at any time $t>0$, the best possible deviation is to choose the myopically optimal action (given the time t information) until the end of the game. Second, the incentive to deviate is strongest at the *earliest* times that new information is revealed, when the "reward" phase – revelation of the truth – is furthest away. As desired, this establishes that it is sufficient to rule out profitable deviations in scenarios (i)–(iii) of the proposition. □

Proposition C5 *(DM deviations at $t=0$). Let $\beta a_b^2 \leqslant 8$. There exists a continuous function $p_a^*: [-2,0] \to (0,1)$, numbers $\varepsilon, \gamma > 0$, and $0 < \alpha'' < 1$ such that if the DM receives an initial recommendation $u_1(a) \in [u_1^{\min}, u_1^{\max}]$, his gain to deviating is strictly negative whenever the following three conditions hold*: (i) $\alpha > \max\{\bar{\alpha}_0, \alpha''\}$, *with $\bar{\alpha}_0$ as in Lemma B*; (ii) $\frac{p_1(a)}{p_1(a)+p_2(a)} \in (p_a^* - \varepsilon, p_a^* + \varepsilon)$, $\forall a \in [-2,0]$; (iii) *for all recommendations* $u_1(a) \in [v_1^{\min}, v_1^{\max}]$, $r_1(a) + r_2(a) < \gamma$.

Proof. Let $G(u_1(a)|\{x(a), g(a)\})$ denote the DM's expected gain from deviating to the myopically optimal action conditional on knowing $\theta \in \{x(a), g(a)\}$. The proof proceeds in 5 steps:

Step 1. It suffices to prove existence of $\varepsilon > 0$, $0 < \alpha'' < 1$, and a continuous function $p_a^*: [-2,0] \to (0,1)$ such that $\forall u_1(a) \in [u_1^{\min}, u_1^{\max}]$, conditions (i)–(ii) imply $G(u_1(a)|\{x(a), g(a)\}) < 0$.

Proof. The statement is trivially true for recommendations $u_1(a) \in [v_1^{\max}, u_1^{\max}]$ sent only by types $x(a)$ and $g(a)$, so that the DM's maximum gain to deviating is precisely $G(u_1(a)|\{x(a), g(a)\})$. So, consider a recommendation $u_1(a) \in [u_1^{\min}, v_1^{\max}]$ sent by four types, $\{z(a'), h(a'), x(a), g(a)\}$, with $a' = v_1^{-1}(u_1(a))$. To show that the DM does not want to deviate, it suffices to ensure that the following upper bound on the gain to deviating is weakly negative:

$$\begin{aligned}&\big(p_1(a)+p_2(a)\big)\cdot G\big(u_1(a)|\{x(a), g(a)\}\big)\\&\quad+\big(r_1(a)+r_2(a)\big)\cdot G\big(u_1(a)|\{z(a'), h(a')\}\big)\end{aligned} \tag{33}$$

[31] Note that the time $t=2\alpha_a$ recommendations sent by Group II, III pairs do not convey any new information, since the DM would already have inferred the true separable group at time $2\alpha_0$.

where $G(u_1(a)|\{z(a'), h(a')\})$ is the maximal gain to deviating at information set $\{z(a'), h(a')\}$.[32] For this, it suffices to choose posteriors for which $G(u_1(a)|\{x(a), g(a)\})$ is strictly negative $\forall u_1(a) \in [u_1^{\min}, u_1^{\max}]$: since the actions u_1, u_2, v_1, v_2 are all bounded, and so $G(u_1(a)|\{z(a'), h(a')\})$ is likewise bounded, it will then follow immediately from continuity that we can choose a small enough weight γ to guarantee that the expression in (33) is negative whenever $r_1(a) + r_2(a) < \gamma$. □

Step 2. $G(u_1(a)|\{x(a), g(a)\}) < 0$ only if $p_a \in (p_a^* - \sqrt{\varepsilon_a}, p_a^* + \sqrt{\varepsilon_a})$, where $p_a \equiv \frac{p_1(a)}{p_1(a)+p_2(a)}$, and

$$p_a^* = \frac{1}{2} + \frac{a}{\theta_3 e^a} - \frac{1 + \frac{2K}{\theta_3 e^a}}{T}, \tag{34}$$

$$\varepsilon_a = \left(\frac{T-2}{T}\right)\left(\frac{1}{4} + \frac{a}{\theta_3 e^a} - \frac{(1 + \frac{2K}{\theta_3 e^a})^2}{2T}\right) - \left(\frac{1-\alpha_0}{\alpha_0}\right)\frac{2K^2}{T(\theta_3 e^a)^2}. \tag{35}$$

Proof. If the DM receives recommendation $u_1(a)$ and assigns probabilities $p_a, 1 - p_a$ to types $x(a), g(a)$ (with $p_a \equiv \frac{p_1(a)}{p_1(a)+p_2(a)}$), then his expected disutility is

$$\begin{aligned} &p_a\bigl(2\alpha_a\bigl(u_1(a) - x(a)\bigr)^2 + 2(1-\alpha_a)\bigl(u_2(a) - x(a)\bigr)^2 + (T-2)(0)\bigr) \\ &\quad + (1-p_a)\bigl(2\alpha_a\bigl(u_1(a) - g(a)\bigr)^2 + 2(1-\alpha_a)\bigl(u_2(a) - g(a)\bigr)^2 + (T-2)(0)\bigr). \end{aligned}$$

Using (15) and (16), and substituting in $x(a) = \theta_3 + a - \theta_3 e^a$, $g(a) = \theta_3 + a$ (from (11)), this simplifies to

$$\frac{2K^2}{\alpha_0} + Ta^2 - 4Ka + 2p_a\theta_3 e^a\bigl(2K - Ta + \theta_3 e^a\bigr). \tag{36}$$

If he instead chooses the myopically optimal action $p_a x(a) + (1-p_a)g(a)$ for the remaining T periods of the game, he earns disutility

$$\begin{aligned} &T \cdot \bigl(p_a\bigl(p_a x(a) + (1-p_a)g(a) - x(a)\bigr)^2 + (1-p_a)\bigl(p_a x(a) + (1-p_a)g(a) - g(a)\bigr)^2\bigr) \\ &\quad = Tp_a(1-p_a)\bigl(\theta_3 e^a\bigr)^2, \quad \text{using } g(a) - x(a) = \theta_3 e^a, \text{ by (11)}. \end{aligned} \tag{37}$$

Subtracting (37) from (36), we obtain the following expression for $G(u_1(a)|\{x(a), g(a)\})$:

$$\frac{2K^2}{\alpha_0} + Ta^2 - 4Ka + \bigl(4K - 2Ta - (T-2)\theta_3 e^a\bigr)\bigl(\theta_3 e^a\bigr)p_a + Tp_a^2\bigl(\theta_3 e^a\bigr)^2.$$

This expression is negative if and only if $p_a \in (p_a^* - \sqrt{\varepsilon_a}, p_a^* + \sqrt{\varepsilon_a})$, where p_a^*, ε_a are as given by (34) and (35). □

Step 3. To complete the proof for case $\beta a_b^2 \leqslant 8$, it suffices to show that Step 2 expressions p_a^*, ε_a satisfy $\lim_{\alpha_0 \to 1} p_a^* \in (0, 1)$, $\forall \in [-2, 0]$, and $\varepsilon \equiv \min_{a\in[-2,0]}(\lim_{\alpha_0\to 1} \varepsilon_a) > 0$, where

[32] This expression describes the amount that the DM could gain, *if* he were able to learn which of the sets $\{x(a), g(a)\}$, $\{z(a'), h(a')\}$ contained the true state θ prior to choosing his deviation. Since this information is in fact not available at time $t = 0$ – he knows only that $\theta \in \{x(a), g(a), z(a'), h(a')\}$ – the expression in (33) is an upper bound on the gain to deviating.

$$\lim_{\alpha_0\to 1}\varepsilon_a = \left(\frac{T-2}{T}\right)\left(\frac{1}{4}+\frac{a}{\theta_3 e^a}-\frac{1}{2T}+\frac{1-(1+\frac{2K}{\theta_3 e^a})^2}{2T}\right), \tag{38}$$

$$\lim_{\alpha_0\to 1} p_a^* = \frac{1}{2}+\frac{a}{\theta_3 e^a}-\frac{1+\frac{2K}{\theta_3 e^a}}{T}, \tag{39}$$

with

$$K = \frac{\beta a_b}{1-\beta}\left(1+\sqrt{\frac{T-\tau}{2\beta}}\right). \tag{40}$$

Proof. Expressions (38) and (39) are the limits of (34) and (35) as $\alpha_0 \to 1$, using the value for K obtained by taking limits (as $\alpha_0 \to 1$) in (17). Since all expressions are continuous in both a and α_0, Steps 2 and 3 together imply existence of $\alpha'' < 1$, $\varepsilon \equiv \min_{a\in[-2,0]}\varepsilon_a > 0$, and a continuous function $p_a^* : [-2,0] \to (0,1)$ such that $\alpha_0 > \alpha''$ and $\frac{p_1(a)}{p_1(a)+p_2(a)} \in (p_a^* - \varepsilon, p_a^* + \varepsilon)$ imply $G(u_1(a)|\{x(a), g(a)\}) < 0$. For recommendations $u_1(a)$ sent only by Group *III* pairs $(x(a), g(a))$, this is precisely a restatement of Proposition C5. For recommendations $u_1(a)$ sent by Group *I* and *II* pairs, Step 1 established existence of $\gamma > 0$ such that whenever $r_1(a) + r_2(a) < \gamma$, the condition $G(u_1(a)|\{x(a), g(a)\}) < 0$ is sufficient to rule out profitable DM deviations at time $t = 0$. □

Step 4. Completing the proof if $a_b \in [-3.18, -2) \Leftrightarrow \beta \in (0.50102, 0.79202]$. By Step 3, we need only show that the expression $\lim_{\alpha_0\to 1}\varepsilon_a$ in (38) is strictly positive $\forall a \in [-2, 0]$, and that the expression $\lim_{\alpha_0\to 1} p_a^*$ in (39) is strictly positive and below 1.

We first prove that $\lim_{\alpha_0\to 1}\varepsilon_a > 0$, $\forall a \in [-2, 0]$. For this, we first show that there exists $\varepsilon' > 0$ with $1 + \frac{K}{\theta_3 e^{-2}} > \varepsilon'$: Substituting (40) and the relationship $\theta_3 e^{-2} = a_b - 2 + 2e^{-a_b}$ (from (10)) into this inequality, we find that it holds iff $\sqrt{\beta(\frac{T-\tau}{2})} < (\frac{2(1-\beta)(e^{-a_b}-1)+a_b}{-a_b})$; the RHS of this equation, using (13), is greater than 2 for $a_b \in [-3.18, -2]$, while the LHS is strictly below $2\sqrt{\frac{4}{5}}$ by (14), in particular $T - \tau \leqslant 8$, and by the fact that $\beta < \frac{4}{5}$ for the range under consideration. This establishes existence of the desired ε'. But then since $K < 0 \Rightarrow \frac{d}{da}\frac{K}{\theta_3 e^a} > 0$, it then follows that

$$\begin{aligned}
\min_{a\in[-2,0]}\left(1-\left(1+\frac{2K}{\theta_3 e^a}\right)^2\right) &= \min_{a\in[-2,0]}\frac{-4K}{\theta_3 e^a}\left(1+\frac{K}{\theta_3 e^a}\right) \\
&\geqslant \frac{-4K}{\theta_3 e^0}\left(1+\frac{K}{\theta_3 e^{-2}}\right) > -\frac{4K}{\theta_3}\varepsilon' \\
\Rightarrow \quad \max_{a\in[-2,0]}\left(1+\frac{2K}{\theta_3 e^a}\right)^2 &< 1-\varepsilon, \quad \text{where } \varepsilon = -\frac{4K}{\theta_3}\varepsilon' > 0.
\end{aligned} \tag{41}$$

Substituting this into the final term in (38), also noting that $a/\theta_3 e^a$ is increasing in a, we obtain

$$\varepsilon \equiv \min_{a\in[-2,0]}\left(\lim_{\alpha_0\to 1}\varepsilon_a\right) > \left(\frac{T-2}{T}\right)\left(\frac{1}{4}-\frac{2}{\theta_3 e^{-2}}-\frac{1}{2T}+\frac{\varepsilon}{2T}\right).$$

This is strictly positive, as desired, by the fact (using (14) and (10)) that $a_b \in [-3.18, -2)$ implies $\theta_3 e^{-2} > 10.778$ and $T \geqslant \frac{5-2\beta}{1-\beta} \geqslant \frac{5-2(0.50102)}{1-0.50102}$, so that

$$\frac{1}{4} - \frac{2}{\theta_3 e^{-2}} - \frac{1}{2T} > 0. \tag{42}$$

Finally, we prove $p_a^* \in (0,1)$. By (41), $(1 + \frac{2K}{\theta_3 e^a}) \in (-1,1)$; substituting this into (34), we obtain

$$\min_{a\in[-2,0]}\left(\frac{1}{2} + \frac{a}{\theta_3 e^a} - \frac{1}{T}\right) \leqslant \lim_{\alpha_0\to 1} p_a^* \leqslant \max_{a\in[-2,0]}\left(\frac{1}{2} + \frac{a}{\theta_3 e^a} + \frac{1}{T}\right).$$

Since $a/\theta_3 e^a$ is increasing in a, the lower bound is $2(\frac{1}{4} - \frac{1}{\theta_3 e^2} - \frac{1}{2T})$, which is positive by (42), and the upper bound is less than 1 by $T > 2$. □

Step 5. Completing the proof if $a_b \in [-2, -1.775]$: see online appendix. □

Proposition C6.1 *(No DM deviations at $t = 2$). If $(p_1(a_b), p_2(a_b), r_1(a_b))$ satisfy $r_1(a_b) \leqslant \frac{\beta}{1-\beta} p_2(a_b)$ and the conditions in Proposition C5, then there are no profitable deviations at $t = 2$.*

Proof. By Proposition C4, we need only worry about deviations at time $t = 2$ if $\theta \in \{0, \theta_2, \theta_3\}$. Recall that at time $t = 2$, type $x(0) = 0$ separates by recommending the true state, 0, while types $g(0), z(a_b)$ continue to follow observationally equivalent strategies: type $g(0) = \theta_3$ recommends the true state, and type $z(a_b) = \theta_2$ continues to recommend $v_2(a_b) = \theta_3$ until revealing the truth at time $T - \tau$. It is clear that the DM cannot gain by deviating if the expert recommends zero, so consider a recommendation θ_3 at time $t = 2$. In this case, Bayesian updating implies posterior probability $\frac{p_2(a_b)}{p_2(a_b)+r_1(a_b)}$ on type θ_3, and the residual probability on type θ_2. If the DM follows the recommendation, he earns expected disutility 0 if he is in fact facing a type θ_3, and $(T - \tau - 2)(\theta_3 - \theta_2)^2$ if he is in fact facing type θ_2 (who recommends θ_3 until revealing the truth at time $T - \tau$); using (12) and (10), the expected disutility is then $(1-\beta)(T-2)(4)\frac{r_1(a_b)}{p_2(a_b)+r_1(a_b)}$. The best deviation is to choose the myopically optimal action, $\frac{p_2(a_b)\theta_3 + r_1(a_b)\theta_2}{p_2(a_b)+r_1(a_b)}$, in all $T - 2$ remaining periods, for disutility $4(T-2)\frac{p_2(a_b)r_1(a_b)}{(p_2(a_b)+r_1(a_b))^2}$. Comparing the two payoffs, we find that deviations are unprofitable whenever the following condition holds

$$(1-\beta)(T-2)(4)\frac{r_1(a_b)}{p_2(a_b)+r_1(a_b)} \leqslant 4(T-2)\frac{p_2(a_b)r_1(a_b)}{(p_2(a_b)+r_1(a_b))^2}$$

which rearranges to the desired condition, $(1-\beta)r_1(a_b) \leqslant \beta p_2(a_b)$. □

Proposition C6.2 *(No deviations at $t = 2\alpha_0$). Let $\beta a_b^2 \leqslant 8$. For all $a \in (a_b, 0]$, let $q_a \equiv \frac{r_1(a)}{r_1(a)+r_2(a)}$, so that $(q_a, 1 - q_a)$ are the DM's posteriors on types $(z(a), h(a))$ after recommendation $v_2(a)$. If Proposition C5 conditions hold, then there exist numbers $\alpha^{**} < 1$ and $\varepsilon \geqslant 0.145$, and a continuous function $q_a^* : [a_b, 0] \to (0.2, 0.7)$ such that the DM's gain to deviating at $t = 2\alpha_0$ is strictly negative whenever $\alpha_0 > \alpha^{**}$ and $q_a \in (q_a^* - \varepsilon, q_a^* + \varepsilon)$.*

Proof. Recall that time $t = 2\alpha_0$, Group *II* pairs separate out from Group *I* pairs who sent the same initial recommendation. If the DM learns that he is facing a Group *II* pair $\{x(a), g(a)\}$, profitable deviations are ruled out by posteriors specified in Proposition C5: The gain to deviating at time $t = 2\alpha_0$ is smaller than $G(u_1(a)|\{x(a), g(a)\})$ (defined following Proposition C5 statement), which was shown to be negative. So, it remains to rule out profitable deviations following Group I recommendations $v_2(a)$. We prove this in three steps:

Step 1. The DM's maximum gain to deviating after recommendation $v_2(a)$ is negative whenever $q_a \in (q_a^* - \sqrt{\varepsilon_a}, q_a^* + \sqrt{\varepsilon_a})$, where

$$q_a^* = \frac{\phi^2 - 1 - 2(\frac{v_2(a)-h(a)}{h(a)-z(a)})}{2\phi^2}, \tag{43}$$

$$\varepsilon_a = \frac{\phi^2-1}{(2\phi^2)^2}\left(\phi - 1 - 2\left(\frac{v_2(a)-h(a)}{h(a)-z(a)}\right)\right)\left(\phi + 1 + 2\left(\frac{v_2(a)-h(a)}{h(a)-z(a)}\right)\right), \tag{44}$$

$$\phi^2 \equiv \frac{T-2\alpha_0}{T-\tau-2\alpha_0}. \tag{45}$$

Proof. If the DM follows recommendation $v_2(a)$ at time $2\alpha_0$ (expecting to choose this action until learning the truth at time $T-\tau$), his expected disutility is

$$(T-\tau-2\alpha_0)\left(q_a\left(v_2(a)-z(a)\right)^2 + (1-q_a)\left(v_2(a)-h(a)\right)^2\right) + \tau(0).$$

Choosing the myopically optimal action $q_a z(a) + (1-q_a)h(a)$ in all remaining $T-2\alpha_0$ periods yields disutility

$$(T-2\alpha_0)q_a(1-q_a)\left(h(a)-z(a)\right)^2.$$

So, the gain to deviating is negative at any belief q_a satisfying the following inequality:

$$\begin{aligned}0 &> \left(q_a\left(v_2(a)-z(a)\right)^2 + (1-q_a)\left(v_2(a)-h(a)\right)^2\right)\\ &\quad - \frac{(T-2\alpha_0)}{(T-\tau-2\alpha_0)}q_a(1-q_a)\left(h(a)-z(a)\right)^2\\ &= q_a\left(2\left(\frac{v_2(a)-h(a)}{h(a)-z(a)}\right)+1\right) + \left(\frac{v_2(a)-h(a)}{h(a)-z(a)}\right)^2 - \phi^2 q_a(1-q_a),\end{aligned}$$

with ϕ^2 as defined in (45). Solving for the values of q_a for which this (quadratic) expression is negative, we obtain the desired Step 1 conditions. □

Step 2. Step 1 functions q_a^*, ε_a satisfy the following bounds:

$$q_a^* \in \left[\frac{\phi^2-1}{2\phi^2}, \frac{\phi^2 - 1 + \frac{\phi^2+\phi\sqrt{\phi^2-t}}{e}}{2\phi^2}\right], \tag{46}$$

$$\min_{a\in[a_b,0]} \varepsilon_a \geqslant \frac{\phi^2-1}{(2\phi^2)^2}(\phi-1-0)\left(\phi+1-\frac{\phi^2+\phi\sqrt{\phi^2-t}}{e}\right). \tag{47}$$

Proof. We first calculate bounds on $2(\frac{v_2(a)-h(a)}{h(a)-z(a)})$. By (19) and (11), we have

$$\begin{aligned}&v_2(a) - h(a)\\ &\quad = \frac{2K - T(a-a_b)}{T-\tau} + \frac{\sqrt{\frac{4\tau\alpha_0}{T-\tau-2\alpha_0}}\sqrt{(\frac{T-\tau-2\alpha_0}{\tau\alpha_0})K^2 + 2K(a-a_b) - \frac{T}{2}(a-a_b)^2}}{T-\tau}\end{aligned}$$

$$= \frac{2K - T(a-a_b)}{T-\tau} + \sqrt{\left(\frac{2K}{T-\tau}\right)^2 + \frac{2\tau\alpha_0}{(T-\tau-2\alpha_0)(T-\tau)}\left(\frac{4K}{T-\tau}(a-a_b) - \frac{T}{(T-\tau)}(a-a_b)^2\right)}.$$

Setting $k \equiv \frac{2K}{T-\tau}$, $t \equiv \frac{T}{T-\tau}$, and $y \equiv a - a_b$, noting (using (45)) that $\frac{2\tau\alpha_0}{(T-\tau)(T-\tau-2\alpha_0)} = \phi^2 - t$, and multiplying by $\frac{2}{h(a)-z(a)} = \frac{1}{e^y}$ (by (11) with $y = a - a_b$), we then obtain

$$2\left(\frac{v_2(a,\alpha_0) - h(a)}{h(a) - z(a)}\right) = \frac{k - ty + \sqrt{k^2 + (\phi^2 - t)(2ky - ty^2)}}{e^y} \equiv \frac{\xi(y)}{e^y}. \tag{48}$$

So we wish to obtain upper and lower bounds on the expression $\frac{\xi(y)}{e^y}$ in (48), for $a \in [a_b, 0] \Leftrightarrow y \in [0, -a_b]$. By construction, the value of K specified in (17) sets the square rooted portion of $v_2(\cdot)$ is equal to zero at $a = 0 \Leftrightarrow y = -a_b$, so we have

$$k^2 + (\phi^2 - t)(-2ka_b - ta_b^2) = 0 \quad \Leftrightarrow \quad k = a_b\left(\phi^2 - t + \phi\sqrt{\phi^2 - t}\right). \tag{49}$$

Next, differentiate (48) to obtain

$$\xi'(y) = -t + \frac{(\phi^2 - t)(k - ty)}{\sqrt{k^2 + (\phi^2 - t)(2ky - ty^2)}}$$

and

$$\xi''(y) = \frac{-k^2\phi^2(\phi^2 - t)}{(k - ty + \sqrt{k^2 + (\phi^2 - t)(2ky - ty^2)})^{\frac{3}{2}}};$$

both are strictly negative, by $\phi^2 > t$, $k > 0$, and $y \geqslant 0$, so we conclude that $\xi(\cdot)$ is strictly decreasing and concave. Therefore, $\xi(\cdot)$ reaches a maximum over the interval $y \in [0, -a_b]$ at $y = 0$, and lies above the straight line connecting the points $(0, \xi(0))$ and $(-a_b, \xi(-a_b))$: since we have $\xi(-a_b) = k + ta_b$ and $\xi(0) = k + \sqrt{k^2} = 0$ (by (48) and (49)), this line $\tilde{\xi}$ is given by

$$\tilde{\xi}(y) - \tilde{\xi}(0) = \frac{\tilde{\xi}(-a_\gamma) - \tilde{\xi}(0)}{-a_b}(y - 0) \quad \Rightarrow \quad \tilde{\xi}(y) = \frac{k + ta_b}{-a_b}y.$$

Substituting in (49), we then obtain the following bounds:

$$\begin{aligned}
\min_{y\in[0,-a_b]} \frac{\xi(y)}{e^y} &\geqslant \min_{y\in[0,-a_b]} \frac{\tilde{\xi}(y)}{e^y} = \left(-\phi^2 - \phi\sqrt{\phi^2 - t}\right)\left(\max_{y\in[0,-a_b]} \frac{y}{e^y}\right) \\
&= \frac{-\phi^2 - \phi\sqrt{\phi^2 - t}}{e}, \\
\max_{y\in[0,-a_\gamma]} \frac{\xi(y)}{e^y} &\leqslant \frac{\max_{y\in[0,-a_\gamma]} \xi(y)}{\min_{y\in[0,-a_\gamma]} e^y} = \frac{\xi(0)}{e^0} = 0.
\end{aligned}$$

Finally, the desired expressions (46) and (47) follow immediately by substituting $2(\frac{v_2(a)-h(a)}{h(a)-z(a)}) \in [\frac{-\phi^2-\phi\sqrt{\phi^2-t}}{e}, 0]$ into (44) and (43). □

Step 3. To complete the proof, it suffices to prove that in the limit $\alpha_0 \to 1$, the expression in (47) exceeds $(0.145)^2$, and the interval in (46) is contained in the interval $(0.2, 0.7)$.

Proof. Immediate by Steps 1, 2, and the obvious continuity of q_a^*, ε_a in both a, α_0. $\square$

Step 4. Completing the proof. Consider the limit as $\alpha_0 \to 1$. Then, $\phi^2 \to \frac{T-2}{T-\tau-2} = \frac{1}{1-\beta}$ (using (12), in particular $\tau = \beta(T-2)$), and $t = \frac{T}{T-\tau} = (1 - \frac{2\beta}{T-\tau})/(1-\beta)$; substituting into (46) and (47), we obtain

$$q_a^* \in \left[\beta/2, \beta/2 + \left(1 + \sqrt{\frac{2\beta}{T-\tau}}\right)/2e\right], \tag{50}$$

$$\min_{a\in[a_\gamma,0]} \varepsilon_a \geqslant \frac{\beta}{2}\left(\beta - \left(\frac{1-\sqrt{1-\beta}}{\sqrt{1-\beta}}\right)\left(\frac{1+\sqrt{\frac{2\beta}{T-\tau}}}{e}\right)\right). \tag{51}$$

For the range $\beta \in [0.4173, 0.50102)$, (14) specifies $T = 7$, so that (by (12)) $T - \tau = 7 - 5\beta$; in this case, it may easily be verified numerically that our lower bound on $\sqrt{\varepsilon_a}$ in (51) reaches a minimum (at $\beta = 0.4172$) of 0.163, our lower bound on q_a^* in (50) is at least $\frac{\beta}{2} \geqslant \frac{0.4172}{2}$, and our upper bound on q_a^* in (50) is at most $\max_{\beta\in[0.4172,0.50102]}(\frac{\beta}{2} + \frac{1+\sqrt{\frac{2\beta}{7-5\beta}}}{2e}) = 0.521\,30$. For the range $\beta \in [0.50102, 0.79202)$, (14) specifies $T - \tau \in [5, 8]$. Over this range, it may easily be verified numerically that our lower bound on $\sqrt{\varepsilon_a}$ in (51) is minimized at $\beta = 0.79202$, and is increasing in $T - \tau$, with a minimum value (at $\beta = 0.79202, T - \tau = 5$) of 0.14505 (and any $T - \tau \in [6, 8]$ guarantees $\varepsilon_a > 0.15$). Our lower bound on q_a^* in (50) is at least $\frac{0.50102}{2} > 0.25$, and the upper bound is at most $\max_{\beta\in[0.50102,0.79202]}(\frac{\beta}{2} + \frac{1+\sqrt{\frac{\beta}{2.5}}}{2e}) = 0.7$. As desired, this establishes that if we choose a horizon T satisfying (14) and an α_0 sufficiently near 1, then $q_a^* \in (0.2, 0.7)$ and $\varepsilon_a > 0.145$. $\square$

Bayesian beliefs: Our incentive constraints for the DM were specified in terms of his posteriors, which in turn depend both on his prior F, and on the precise details of our construction. We now show in Proposition C7 (via a preliminary result in Lemma C7.1) that the posteriors satisfying the conditions in Propositions C5, C6.1, C6.2 are the Bayesian posteriors corresponding to some prior over the state space.

Lemma C7.1 *(Constructing priors). Let strategies be as specified in Appendix C.1, with $\alpha(\cdot)$ satisfying (20). For any continuous functions $p : [-2, 0] \to (0, 1)$, $q : [a_b, 0] \to (0, 1)$, and $r : [a_b, 0] \to [0, 1]$ such that $p(\cdot), q(\cdot)$ are bounded away from 0 and 1, there exists a density f over the state space such that, in our construction, a Bayesian DM will hold the following posterior beliefs*: (i) $\forall a \in [-2, 0]$, $\frac{\Pr(\theta=x(a)|u_1(a))}{\Pr(\theta=g(a)|u_1(a))} = \frac{p(a)}{1-p(a)}$; (ii) $\forall a \in (a_b, 0]$, $\frac{\Pr(\theta=z(a)|v_1(a))}{\Pr(\theta=h(a)|v_1(a))} = \frac{q(a)}{1-q(a)}$; (iii) $\forall a \in (a_b, 0]$, $\frac{\Pr(\theta\in\{z(a),h(a)\}|v_1(a))}{\Pr(\theta\in\{x(u_1^{-1}(v_1(a))),g(u_1^{-1}(v_1(a)))\}|v_1(a))} < \frac{\gamma}{1-\gamma}$, *for any* $\gamma > 0$; *and* (iv) $\frac{\Pr(\theta=z(a_b)|v_2(a_b))}{\Pr(\theta=g(0)|v_2(a_b))} < \frac{\beta}{1-\beta}$.

Proof. As explained in Section 5, we assume that the DM is Bayesian. For our construction, (7) then implies the following relationships between priors and posteriors:

$$\frac{\Pr(\theta = x(a)|u_1(a))}{\Pr(\theta = g(a)|u_1(a))} = \frac{f(x(a))}{f(g(a))}\left|\frac{x'(a)}{g'(a)}\right| = \frac{f(x(a))}{f(g(a))}\left(\theta_3 e^a - 1\right), \quad \forall a \in [-2, 0], \tag{52}$$

$$\frac{\Pr(\theta = z(a)|v_2(a))}{\Pr(\theta = h(a)|v_2(a))} = \frac{f(z(a))}{f(h(a))}\left|\frac{z'(a)}{h'(a)}\right| = \frac{f(z(a))}{f(h(a))}\left(2e^{a-a_b} - 1\right), \quad \forall a \in [a_b, 0]. \tag{53}$$

And, for recommendations $v \in (v_1^{\min}, v_1^{\max}]$ sent by the four types $\{z(a), h(a), x(a'), g(a')\}$ with $a = v_1^{-1}(v) > a_b$ and $a' = u_1^{-1}(v) < 0$, we use[33]

$$\frac{\Pr(\theta = z(a)|v_1(a))}{\Pr(\theta = x(a')|v_1(a))} = \lim_{\varepsilon \to 0} \frac{F(z(v_1^{-1}(v+\varepsilon))) - F(z(v_1^{-1}(v-\varepsilon)))}{F(x(u_1^{-1}(v+\varepsilon))) - F(x(u_1^{-1}(v-\varepsilon)))} = \frac{dF(z(a))}{dF(x(a'))}. \tag{54}$$

Consider first the expression in (54). If $\beta a_b^2 \leqslant 8$, so $\alpha_a = \alpha_0$ (constant), solving (15), (18) for $u_1^{-1}(v)$, $v_1^{-1}(v)$ and differentiating yields[34]

$$\frac{dF(z(a))}{dF(x(a'))} = \frac{f(z(a))}{f(x(a'))}\left(\frac{2e^{a-a_b}-1}{\theta_3 e^{a'}-1}\right)\left|\frac{u_1'(a')}{v_1'(a)}\right|,$$

with

$$\left|\frac{u_1'(a')}{v_1'(a)}\right| = \left(\frac{(\sqrt{\frac{T-\tau-2\alpha_0}{2\beta}})(-T(v-v_1^{\min}) - \frac{(T-2\alpha_0)K}{\alpha_0}) - \sqrt{(T-2)(v-v_1^{\min})(-2K(T-2\alpha_0) - T\alpha_0(v-v_1^{\min}))}}{\sqrt{1-\alpha_0}(-T(v-v_1^{\min}) - \frac{(T-2\alpha_0)K}{\alpha_0}) + \sqrt{(T-2)(v-v_1^{\min})(-2K(T-2\alpha_0) - T\alpha_0(v-v_1^{\min}))}}\right).$$

This expression for $|\frac{u_1'(a')}{v_1'(a)}|$ is strictly decreasing in $(v - v_1^{\min})$, and therefore reaches a maximum value, at $v = v_1^{\min}$, of $\sqrt{\frac{T-\tau-2\alpha_0}{2\beta(1-\alpha_0)}}$; in fact, it can be shown that this same upper bound obtains in the case $\beta a_b^2 > 8$, where α is a continuous decreasing function which reaches a minimum value, α_0, at $a = 0 = u_1^{-1}(v_1^{\min})$. Therefore, for any $\alpha_0 < 1$, there exists a finite number λ such that

$$\frac{\Pr(\theta = z(a)|v_1(a))}{\Pr(\theta = x(a')|v_1(a))} \leqslant \lambda\left(\frac{f(z(a))}{f(x(a'))}\right), \quad \forall a \in [a_b, 0]. \tag{55}$$

It is now straightforward to construct a density with the desired properties. Here is one such construction: begin by setting

$$f(x(a)) = \frac{1}{M}, \quad \forall a \in [-2, 0] \tag{56}$$

and with M a constant to be determined (this specifies a density over $[0, \theta_1]$). Next, to specify a density over $[\theta_2, \theta_3]$ such that condition (i) holds, substitute (56) into (52) and condition (i), to obtain

$$f\big(g(a)\big) = \left(\frac{\theta_3 e^a - 1}{M}\right)\left(\frac{1-p(a)}{p(a)}\right), \quad \forall a \in [-2, 0]. \tag{57}$$

For (iv), use (55) and condition (i) to obtain

$$\frac{\Pr(\theta = z(a_b)|v_1(a))}{\Pr(\theta = g(0)|v_1(a))} = \frac{\Pr(\theta = z(a_b)|v_1(a))}{\Pr(\theta = x(0)|v_1(a))}\frac{p(0)}{1-p(0)} \leqslant \lambda\left(\frac{f(z(a_b))}{f(x(0))}\right)\left(\frac{p(0)}{1-p(0)}\right).$$

[33] In the limit expression, replace $(v+\varepsilon)$ with v if $v = v^{\max}$, and replace $(v-\varepsilon)$ with v if $v = v^{\min}$.

[34] The point of this paragraph is to show that $|\frac{u_1'(a')}{v_1'(a)}|$ is bounded. If this were not the case, then it would not be possible (via a suitable choice of prior) to ensure posterior beliefs satisfying Proposition C5 requirement that the weight on pair $\{z(a), h(a)\}$ be below some cutoff λ. This takes some care for recommendations near $u_1(0) = v_1(a_b)$, noting from (15) and (18) that $u_1'(0) = v_1'(a_b) = 0$.

So, by (56), we can satisfy condition (iv) by setting

$$\lambda\left(\frac{f(z(a_b))}{f(x(0))}\right)\left(\frac{p(0)}{1-p(0)}\right) \leqslant \frac{\beta}{1-\beta}$$
$$\Leftrightarrow \quad f\big(z(a_b)\big) \leqslant \frac{1}{\lambda}\left(\frac{\beta}{1-\beta}\right)\left(\frac{1}{M}\right)\left(\frac{1-p(0)}{p(0)}\right). \tag{58}$$

For (iii), use (with $a' \equiv u_1^{-1}(v_1(a))$)

$$\frac{\Pr(\theta \in \{z(a), h(a)\}|v_1(a))}{\Pr(\theta \in \{x(a'), g(a')\}|v_1(a))} = \frac{\Pr(\theta = z(a)|v_1(a))}{\Pr(\theta = x(a')|v_1(a))}\left(\frac{1+\frac{\Pr(\theta=h(a)|v_1(a))}{\Pr(\theta=z(a)|v_1(a))}}{1+\frac{\Pr(\theta=g(a')|u_1(a'))}{\Pr(\theta=x(a')|u_1(a'))}}\right)$$
$$< \lambda\left(\frac{f(z(a))}{f(x(a'))}\right)\left(\frac{p(a')}{q(a)}\right) \quad \text{by (55) and condition (ii).}$$

So, if condition (ii) is satisfied and (56) holds, then we can satisfy condition (iii) by setting

$$f\big(z(a)\big) < \frac{1}{\lambda}\left(\frac{\gamma}{1-\gamma}\right)\left(\frac{1}{M}\right)\left(\frac{q(a)}{p(u_1^{-1}(v_1(a)))}\right), \qquad \forall a \in [a_b, 0]. \tag{59}$$

Together with (58), this specifies a density over $[\theta_1, \theta_2]$ which satisfies properties (iii) and (iv) of the proposition. Lastly, to construct a density over $(\theta_3, \frac{1}{b}]$, substitute (53) into condition (ii), to obtain

$$f\big(h(a)\big) = f\big(z(a)\big)\big(2e^{a-a_b}-1\big)\left(\frac{1-q(a)}{q(a)}\right). \tag{60}$$

Thus, it suffices to set (59) to hold with equality, and substitute into (60) to obtain

$$f(h(a)) = \frac{1}{\lambda}\left(\frac{\gamma}{1-\gamma}\right)\left(\frac{1}{M}\right)\left(\frac{1-q(a)}{p(u_1^{-1}(v_1(a)))}\right)\big(2e^{a-a_b}-1\big). \tag{61}$$

Finally, choose M so that the total measure of the type space integrates to 1: this is possible since the densities in (56), (57), (59), and (60) are all finite. So, integrating over the state space yields a finite number divided by M; choose M so that this number equals 1. □

Proposition C7 *(Priors for Theorem 1). There is an open set of priors F over the state space which yield Bayesian posterior satisfying the conditions in Propositions C5, C6.1, C6.2.*

Proof. After any recommendation $u_1(a) \in [u_1^{\min}, u_1^{\max}]$, let $(r_1(a'), r_2(a'), p_1(a), p_2(a))$ denote the DM's posteriors on types $(z(a'), h(a'), x(a), g(a))$, with $a' = v_1^{-1}(u_1(a))$ if also $u_1(a) \in [v_1^{\min}, v_1^{\max}]$, and defining $r_1(a') = r_2(a') = 0$ otherwise. Proposition C5 requires two conditions: (i) that $\frac{p_1(a)}{p_1(a)+p_2(a)}$ lie with in an ε-interval of a number p_a^* (for some strictly positive ε, and with p_a^* a continuous function bounded away from 0 and 1); that is, $\frac{p_1(a)}{p_2(a)} \equiv \frac{\Pr(x(a))}{\Pr(g(a))}$ must be sufficiently close to $\frac{p_a^*}{1-p_a^*}$; (ii) that $\frac{r_1(a')+r_2(a')}{p_1(a)+p_2(a)}$ be below $\frac{\gamma}{1-\gamma}$, for some $\gamma > 0$. Proposition C6.2 requires additionally that $\forall a \in [a_b, 0]$, $\frac{r_1(a)}{r_2(a)} \equiv \frac{q_a}{1-q_a}$ be sufficiently close to a number $\frac{q_a^*}{1-q_a^*}$, with q_a^* a continuous function bounded away from 0 and 1. We showed in Lemma C7.1 that for such functions p_a^*, q_a^*, and such a number $\gamma > 0$, there is a prior F generating Bayesian posteriors $\frac{p_1(a)}{p_2(a)} = \frac{p_a^*}{1-p_a^*}$ (condition (i)), $\frac{r_1(a')+r_2(a')}{p_1(a)+p_2(a)} < \frac{\gamma}{1-\gamma}$ (condition (iii)), $\frac{r_1(a)}{r_2(a)} \equiv \frac{q_a^*}{1-q_a^*}$ (condition (ii)), and which satisfies the condition in Proposition C6.1 (condition (iv)). It follows immediately that any density which is sufficiently close to F will generate posteriors satisfying the conditions in Propositions C5, C6.1, C6.2, thus yielding the desired open set. □

C.4. Completing the proof of Theorem 1

Appendix C.2 proved that the expert cannot gain by deviating from the strategy specified in Appendix C.1. Appendix C.3 established (Proposition C4) that it is sufficient to rule out profitable DM deviations at times $t \in \{0, 2\alpha_0, 2\}$, and then proved that there is an open set of priors (Proposition C7) which generate, as Bayesian posteriors, beliefs at which the DM will find it optimal to follow all expert recommendations at time $t = 0$ (Proposition C5), and at times $t = 2\alpha_0, 2$ (Propositions C6.1, C6.2). Since the off-path strategies specified in Appendix C.1 are trivially optimal, and the DM's beliefs Bayesian, it follows that our strategies indeed constitute a fully revealing equilibrium under Proposition C7 priors. This completes the proof if time is continuous. With integer constraints, our timeline is most easily modified via public randomization and "scaling up".[35,36]

Appendix D. Supplementary material

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.jet.2013.12.012.

References

[1] A. Ambrus, S. Lu, Robust almost fully revealing equilibria in multi-sender cheap talk, Mimeo, 2010.
[2] L. Anderlini, D. Gerardi, R. Lagunoff, Communication and learning, Rev. Econ. Stud. 79 (2) (2012) 419–450.
[3] S. Athey, K. Bagwell, Collusion with persistent cost shocks, Econometrica 76 (3) (2008) 493–549.
[4] R. Aumann, S. Hart, Long cheap talk, Econometrica 71 (6) (2003) 1619–1660.
[5] A. Aumann, M. Maschler, Repeated Games with Incomplete Information, MIT Press, 1995 (Chapter 1).
[6] M. Battaglini, Multiple referrals and multi-dimensional cheap talk, Econometrica 70 (4) (2002) 1379–1401.
[7] D. Bergemann, M. Said, Dynamic auctions, in: Wiley Encyclopedia of Operations Research and Management Science, vol. 2, 2011, pp. 1511–1522.
[8] A. Blume, O. Board, K. Kawamura, Noisy talk, Theoretical Econ. 2 (4) (2007) 395–440.
[9] Y. Chen, N. Kartik, J. Sobel, Selecting cheap talk equilibria, Econometrica 76 (1) (2008) 117–136.
[10] V. Crawford, J. Sobel, Strategic information transmission, Econometrica 50 (6) (1982) 1431–1451.
[11] P. Eso, Y.-F. Fong, Wait and see, Mimeo, 2008.

[35] Public randomization does not play any substantive role here: the sole purpose is to facilitate the description of a fully revealing equilibrium, via allowing the relative lengths of the different recommendation phases to be set (in expectation) to any desired levels.

[36] A sketch of the modification is as follows. If $\beta a_b^2 \leqslant 8$, so that the function α specified in (20) is constant, $\alpha(a) = \alpha_0$, then choose an integer T satisfying (14), and an integer λ large enough that setting $\alpha_0 = \frac{\lambda-1}{\lambda}$ satisfies the bounds in Lemma B and Propositions C5, C6.1, C6.2. Multiply all time parameters by λ, so that the expert switches to recommendation functions u_2, v_2 at time $2\lambda\alpha_0 = 2(\lambda - 1)$, Group *II*, *III* pairs reveal the truth at time 2λ, and the game lasts for $T\lambda$ periods, all integer times. The only (potential) non-integer is the time, now $(T - \tau)\lambda$, at which Group *I* pairs reveal the truth; if this is not an integer, choose the two nearest integers $t_1 < (T - \tau)\lambda < t_2$, and use public randomization to determine whether the expert reveals the truth at time t_1 or t_2 (such that the expected revelation time is $(T - \tau)\lambda$). This modification simply scales up all expected payoff expressions by a factor of λ, and so our analysis goes through unchanged. The timeline modification is slightly more complicated in the range $\beta a_b^2 > 8$, but can again be achieved via public randomization along a grid of discrete times.

[12] F. Forges, F. Koessler, Multistage communication with and without verifiable types, Int. Game Theory Rev. 10 (2) (2008) 145–164.
[13] F. Forges, F. Koessler, Long persuasion games, J. Econ. Theory 143 (1) (2008) 1–35.
[14] M. Goltsman, J. Hőrner, G. Pavlov, F. Squintani, Mediation, arbitration and negotiation, J. Econ. Theory 144 (4) (2009) 1397–1420.
[15] M. Goltsman, G. Pavlov, How to talk to multiple audiences, Games Econ. Behav. 72 (1) (2011) 100–122.
[16] M. Ivanov, Dynamic informational control, Mimeo, 2011.
[17] N. Kartik, M. Ottaviani, F. Squintani, Credulity, lies, and costly talk, J. Econ. Theory 134 (1) (2007) 93–116.
[18] V. Krishna, J. Morgan, A model of expertise, Quart. J. Econ. 116 (2) (2001) 747–775.
[19] V. Krishna, J. Morgan, The art of conversation: eliciting information from experts through multi-stage communication, J. Econ. Theory 117 (2) (2004) 147–179.
[20] J. Mathis, Full revelation of information in sender – receiver games of persuasion, J. Econ. Theory 143 (1) (2008) 571–584.
[21] S. Morris, Political correctness, J. Polit. Economy 109 (2) (2001) 231–265.
[22] M. Ottaviani, P. Sorensen, Professional advice, J. Econ. Theory 126 (1) (2006) 120–142.
[23] M. Ottaviani, P. Sorensen, Reputational cheap talk, RAND J. Econ. 37 (1) (2006) 155–175.
[24] A. Pavan, I. Segal, J. Toikka, Dynamic mechanism design: revenue equivalence, profit maximization and information disclosure, Discussion Paper, MIT, Northwestern University, Stanford University, 2011.
[25] M. Peski, Repeated games with incomplete information, Theoretical Econ. 3 (1) (2008) 29–84.
[26] J. Renault, E. Solan, N. Vieille, Dynamic sender-receiver games, Mimeo, 2011.
[27] J. Sobel, A theory of credibility, Rev. Econ. Stud. 52 (4) (1985) 557–573.
[28] J. Sobel, Signalling games, in: R. Meyers (Ed.), Encyclopedia of Complexity and Systems Science, Springer, 2009, pp. 8125–8139.