Professional advice[*]

Marco Ottaviani[a,†] and Peter Norman Sørensen[b]

[a] Economics Subject Area, London Business School, Sussex Place, Regent's Park, London NW1 4SA, England

[b] Institute of Economics, University of Copenhagen, Studiestræde 6, DK–1455 Copenhagen K, Denmark

Abstract

This paper studies strategic communication by an expert who is concerned about appearing to be well informed. The expert is assumed to observe a private signal with a simple and particularly tractable ("multiplicative linear") structure. The quality of the expert's information is evaluated on the basis of the advice given and the realized state of the world. In equilibrium of this reputational cheap-talk game, no more than two messages are effectively reported. The model is extended to consider sequential communication by experts with conditionally independent signals. In the long run, learning is incomplete and herd behavior arises.

JEL classification: D72; D82; D83; G20; J30.

Keywords: Reputation; Cheap talk; Advice; Herding.

"Much has been written about the doubtful accuracy of economists' predictions. ...they are better at predicting the direction than the actual magnitude of events. ...This is disappointing, but it does not mean that economics is not a science." ('Economics', Encyclopaedia Britannica Online).

1. Introduction

Professional advisers are often concerned with their reputation for being well informed, rather than with the decisions made on the basis of their recommendations. For example, business consultants care about the perceived quality of their information services, judges intend to promote their career, and politicians are interested in being appointed. A number of recent empirical studies have indeed argued that reputational concerns of this sort play an important role in the financial industry.[1] In this context, it is an open question whether these implicit market incentives guarantee truthful reporting of information. If not, what biases are induced? What form will communication assume? Will experts be driven to engage in herd behavior?

This paper addresses these questions within a simple and tractable model of strategic information transmission by a purely professional *expert.* We assume that the expert privately observes a *signal* generated by a *multiplicative linear experiment,* a natural generalization of the binary signal experiment that allows for continuously varying intensity. The expert's unobserved *ability* parametrizes the amount of information about the *state* of the world contained in the expert's signal. After observing the signal, the expert is free to publicly report any *message.* An *evaluator* then combines the expert's message with the ex-post realization of the state to update the belief regarding the expert's ability. This posterior belief about ability is the expert's *reputation* and determines her payoff.

The first point we investigate is whether this reputational concern gives the right incentives to truthfully report information. We show that this is not the case. The signal that gives rise to the highest expected reputational payoff is typically different from the signal actually observed. To check this, suppose that the evaluator conjectures that the expert truthfully reports the signal. Consider a situation with a prior belief strongly in favor of a high state. In this case, the expert will maximize the chance of being right by reporting a high signal even when observing a low signal. If the prior belief is biased enough in favor of high states, the expert wants to pretend to have received a higher signal than the one

---

[1]See e.g. Lamont [16] on macroeconomic forecasters, Ehrbeck and Waldmann [9] on three-month U.S. Treasury Bills interest rate forecasters, Graham [11], Hong, Kubik and Solomon [15], Welch [36], and Zitzewitz [37] on security analysts, and Chevalier and Ellison [6] on mutual fund managers.

actually observed. Thus, truthtelling cannot be sustained. Similarly, if the prior belief is biased in favor of low states, the adviser wants to pretend to have received a signal below the actual one, again giving an incentive to deviate from truthtelling. Intuitively, in these two cases more extreme signals (i.e., signals resulting in posterior beliefs about the state that are more biased than the prior) are more likely to be associated with high ability and so result in higher expected reputational payoff. If instead the prior belief on the state is intermediate, the expert wishes to deviate to a more moderate signal, which lies between an interior signal and the one actually received.

Secondly, we provide a complete characterization of the set of equilibria in this model. We show that at most two messages can be transmitted in equilibrium. The most informative equilibrium is either binary or completely uninformative. Roughly speaking, the expert can at best communicate the direction of her information, but cannot convey its intensity. Further, if the prior is already very concentrated on any particular state, the expert always wishes to bias her report in the same direction. The adviser is then unable to communicate any information at all in equilibrium.

Thirdly, we extend this model to a dynamic setting, in which different experts give advice one after the other, holding fixed the state. Experts who report early in the sequence are able to learn about the state by listening to the messages sent by the preceding experts. As more experts speak informatively, the beliefs of later experts become ever more concentrated on the true state. However, in finite time the belief on the state becomes so concentrated that experts stop being informative and learning ceases at an inconclusive belief.[2]

1.1. Related Literature

We build on previous work in the area of reputation, cheap-talk communication, and herd behavior. In our model, reputation is about the quality of the expert's information. The first model of this sort of reputation is contained in Section 3.2 of Holmström [13].[3] In his setting of delegated reputational signalling, the expert is directly acting (rather than sending a message) and can affect whether the state of the world is observable or not. We

---

[2]As noted by Welch [36], if advisers have a continuous message space it should be possible to invert their (supposedly separable) strategies, thereby uncovering their private signals. If reports are released sequentially and fully incorporate previous reports, the most recently issued report should efficiently aggregate the private information held by all advisers. For this reason, Welch studies financial analysts with an exogenously coarse message space (consisting of recommendations like "sell", "hold", and "buy") in order to obtain herding. Since the message space is endogenously coarse in reputational cheap talk equilibrium, there is no need to impose exogenous restrictions on the message space to obtain herding.

[3]Others have considered reputational concerns in settings in which ability adds directly to the value of the project undertaken (e.g., Holmström and Ricart i Costa [14]).

instead assume that the state is eventually observed by the evaluator, regardless of the message sent.[4] We directly formulate a static game with exogenously given reputational concerns. As in [13], this reputational payoff can be easily derived from a second period continuation game.[5]

Our model can be seen as one of reputational cheap talk, in which the expert is a sender and the evaluator is a privately informed receiver. The message sent by the expert is *per se* costless, and affects the reputational payoff only through the inference made by the evaluator about the expert's signal. While Crawford and Sobel [8] focused on strategic information transmission by a *partisan* expert who is interested in biasing the action of a decision maker, we consider the case of a *professional* expert concerned with her reputation for ability. Note that our receiver observes not only the message sent, but also the ex-post realization of the state. As a result, experts with different signals have different preferences for different messages.[6]

More directly, we follow the lead of Scharfstein and Stein [29], who analyzed a reputational cheap talk game with two experts giving sequential recommendations. In their model, signals, states and ability types are all binary, and a high ability expert observes a signal conditionally more correlated with the state than a low ability type. In order to guarantee existence of an informative equilibrium for the first expert, [29] fixed the initial prior and focused on the incentives of the second expert. We instead characterize the set of equilibria in the static communication game as a function of the prior belief on the state. The qualitative properties of the equilibrium uncovered in our continuous signal model could not be detected in a binary signal model.

In our dynamic extension, we depart from Scharfstein and Stein by assuming that signals are conditionally independent across experts. Our reputational herding model can then be solved forward, like the statistical herding models of Banerjee [2] and Bikhchandani, Hirshleifer and Welch [3]. The herding result obtained in our reputational model parallels that obtained in the context of statistical herding.

The paper is organized as follows: Section 2 sets up the model. Section 3 contains

---

[4]In Brandenburger and Polak's [4] model, the state realization is never observed and the expert cares about the stock market's assessment of the quality of the decision made. In our model instead, the expert does not care about the quality of the decision made.

[5]Others have considered fully-fledged repeated games in which players differ in preferences, rather than in quality of information possessed (e.g., Sobel [33]).

[6]As noticed more generally by Seidmann [31] and Watson [35], information can sometimes be transmitted even if all sender types share the same ordinal preferences, provided that the receiver's decision is based on some information not available to the sender. Olszewski [21] shows that honest reporting is the *unique* equilibrium when the sender is exclusively motivated by the reputation concern for honesty, and if the receiver observes a signal informative about the sender's signal. His result does not apply to our model, because our expert is concerned about her reputation for accuracy, not honesty.
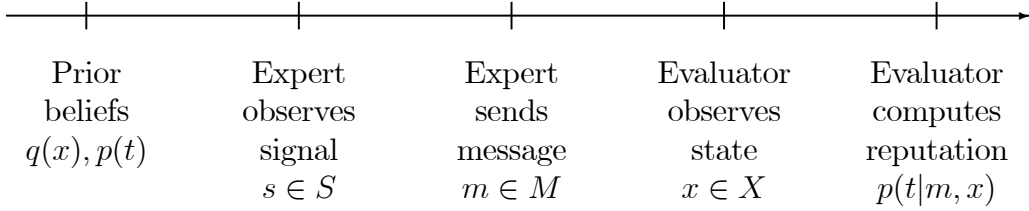
Fig. 1. Timeline for the model. The conditional density of the signal is $f(s|x,t)$.

some preliminary analysis. Section 4 characterizes the optimal deviation if the evaluator expects truthful reporting. Section 5 analyzes the equilibrium and derives comparative statics predictions. Section 6 considers the dynamic extension. Section 7 concludes. All proofs are collected in the Appendix.

2. Model

Setup. We consider the following reputational cheap talk game between an expert and an evaluator. The expert's *ability* $t \in T = [0,1]$ and the *state* of the world $x \in X = [-1,1]$ are given exogenously and so are not affected by the strategic choices of the players. Ability and state are assumed to be statistically independent, with non-degenerate probability density functions (p.d.f.) $p(t)$ and $q(x)$. These common prior beliefs on state and ability are updated by the players during the game. The timing of the model is summarized in Fig. 1.

First, the expert observes a *signal* $s \in S = [-1,1]$ about the state of the world x. This signal is drawn from the conditional probability density $f(s|x,t)$ specified below. The expert's ability t parametrizes the amount of information about the state contained in this signal. We assume that the expert does not know her own ability type t (cf. the discussion in Section 5.5).

Second, the expert decides which *message* $m \in M$ to send. The messages that can be sent do not depend on the signal observed by the expert. As explained below, the message space is determined only as part of an equilibrium. A strategy for the expert prescribes a probability distribution $\mu(\cdot|s)$ over messages m for each signal s.

Third, the evaluator observes the message m and the realization of the state x and forms the posterior belief $p(t|m,x)$ about ability. This belief affects the expert's payoff.
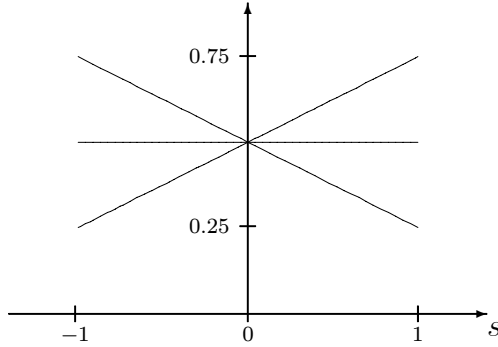
4

Fig. 2. Graphs of the conditional densities $f(s|x,t) = (1 + stx)/2$ for fixed $t = 1/2$ and three values of $x = -1, 0, 1$. The downward sloping line corresponds to the case with $x = -1$, the flat one to $x = 0$, and the increasing one to $x = 1$. Intermediate values of x would give intermediate lines. Each line (other than the one corresponding to $x = 0$) becomes steeper as t increases.

Expert's Information. The expert's signal has conditional probability density function

$$f(s|x,t) = tg(s|x) + (1-t)h(s) = t\frac{1+sx}{2} + (1-t)\frac{1}{2} = \frac{1}{2}(1+stx), \qquad (2.1)$$

as illustrated in Fig. 2. This *multiplicative linear experiment* is a mixture between an informative and an uninformative experiment.[7] It satisfies the monotone likelihood ratio property (MLRP) in s, x for any value of $t > 0$: the likelihood ratio $f(s|x,t)/f(s|x',t)$ is increasing in s for $x > x'$. Clearly, also $g(s|x) = f(s|x,1)$ and $f(s|x) = \int_T f(s|x,t)p(t)\,dt = f(s|x, Et)$ satisfy the MLRP.

Better experts are more likely to receive a signal drawn from the informative $g(s|x) = (1 + sx)/2$ rather than the uninformative $h(s) = 1/2$. Naturally, a more talented expert receives better information in the sense of Blackwell. To see this, consider the garbling of s into $\widetilde{s}$ whereby $\widetilde{s} = s$ with probability $\tau < 1$, and otherwise $\widetilde{s}$ is independently redrawn from $h(s)$. Then $\widetilde{s}$ has conditional p.d.f. $\tau f(\widetilde{s}|x,t) + (1-\tau)h(\widetilde{s}) = f(\widetilde{s}|x, \tau t)$, so that the garbled signal to an expert of ability $t > 0$ is distributed as the non-garbled signal to an expert of ability $\tau t < t$.

---

[7]Note the similarity of our linear model $tg(s|x) + (1-t)h(s)$ with Green and Stokey's [12] success-enhancing model. In the success-enhancing model, the experiment fails with positive probability, in which case the signal is uninformative about the state. While in the success-enhancing model the expert observes whether the experiment failed or not, in our linear model the expert only knows the probability with which the experiment is contaminated.

Special versions of our multiplicative linear model have been extensively used in economics. See e.g. Lohmann's [19] generalization of the binary model and Piccione and Tan's [25] example of a signal structure with an uninformative signal (page 504).

The widely used symmetric binary model has the same generalized p.d.f. (2.1), with $S = X = \{-1, 1\}$ and $T = \{\underline{t}, \overline{t}\}$, where $0 \leq \underline{t} < \overline{t} \leq 1$ (see e.g. Scharfstein and Stein [29]). It is thus useful to think of a signal satisfying (2.1) as being binary, but of a continuously varying intensity level. The multiplicative linear experiment is then a natural generalization of the binary experiment to allow for a continuum of states, signals, and ability types in a tractable way. By normalizing the support of S and X to the unit interval and taking the prior on the state to be uniform, it is easily verified that this is the Farlie-Gumbel-Morgenstern distribution with uniform marginals (cf. Conway [7]).

Evaluation. The evaluator is rewarded for predicting as accurately as possible (in probability distributions) the ability t of the expert, based on the message m and the ex-post realization of the state x. The evaluator forms a conjecture $\hat{\mu}$ on the strategy used by the expert, and computes the chances of the evidence, $\hat{f}(m|x, t) = \int_S \hat{\mu}(m|s) f(s|x, t) \, ds$ and $\hat{f}(m|x) = \int_T \hat{f}(m|x, t) p(t) \, dt$. The posterior reputation is then calculated by Bayes' rule, $p(t|m, x) = p(t) \hat{f}(m|x, t) / \hat{f}(m|x)$.

Expert's Preferences. The expert's preferences over posterior reputations are assumed to be represented by the strictly increasing von Neumann-Morgenstern *utility function* $v(t)$. The expert aims at maximizing her expected reputational value, where the expectation is taken with respect to the posterior reputation $p(t|m, x)$. The reputational payoff of message m in state x is

$$W(m|x) \equiv \int_T v(t) p(t|m, x) \, dt, \tag{2.2}$$

so that the expected payoff for an expert with signal s who sends message m is

$$V(m|s) = \int_X W(m|x) q(x|s) \, dx, \tag{2.3}$$

where the expert's posterior belief about the state x conditional on receiving signal s is given by Bayes' rule as $q(x|s) = f(s|x) q(x) / f(s)$, with $f(s|x) = \int_T f(s|x, t) p(t) \, dt$ and $f(s) = \int_X f(s|x) q(x) \, dx$.

This reduced-form formulation of the reputational payoff captures the fact that experts with better information tend to be more valuable to future decision makers. This reputational payoff can be derived from the value of the services provided in a second and last period by the expert, as in Holmström [13]. In the second period, there are no strict incentives and so truthful revelation can be supported as an equilibrium outcome. Since experts with higher ability have more informative signals, they provide information services of higher value in the second period and so should receive higher rewards. In

addition, if information about ability has no value in the second period, the reputational payoff is linear in the posterior belief about ability, as specified in (2.3).[8]

Equilibrium. By definition, a weak perfect Bayesian Nash equilibrium (henceforth, *equilibrium*) consists of a message space M, an expert's strategy μ, and an evaluator's conjecture $\hat{\mu}$, such that $\hat{\mu} = \mu$ and $\mu(\cdot|s)$-almost all m maximize $V(\cdot|s)$ defined through $\hat{\mu}$. This equilibrium is weak because no restrictions are imposed on the evaluator's beliefs when observing off-path messages.

3. Preliminary Analysis

Regardless of the privately observed signal, the expert wishes to induce the evaluator to have the most favorable posterior beliefs about ability. But these beliefs depend on the realization of the state, which is not known to the expert when the message is sent. Each message then induces a lottery over posterior reputations, with final payoffs that depend on realized states. Holding fixed the evaluator's rule for computing posterior reputations, different messages result in different lotteries over these posterior reputations. An expert can then find the lotteries corresponding to different messages differently appealing depending on her posterior beliefs about the state, as determined by the signal observed.

In order to address whether full sorting of experts with different signals into different messages is possible and compatible with rationality of the evaluator, it is worth noting some properties of the expert's signal (2.1). First, the likelihood ratio $f(s|x,t')/\,f(s|x,t)$ depends on s and x only through $g(s|x)$, and it is strictly increasing in $g(s|x)$ for $t' > t$. This means that it is good news for the expert's reputation to receive a combination of signal and state that corresponds to a higher density conditional on the signal being informative. Since the uninformative part $h(s)$ is uniform, there is no concern about biasing the report away from signals more typically associated with the uninformed expert. Second, $g(s|x)$ is a strictly increasing function of sx. This implies that it is good for the expert's reputation to convey the impression that s is high whenever the state is known to be $x > 0$, while lower signals are more attractive when $x < 0$. As established in the

---

[8]Here are two scenarios in which information about ability has no value. First, this is the case if in the second period the decision marker's mapping from signal to action does not depend on the expert's ability (e.g., if the decision maker's objective function is $u(a,x) = ax$, with $a \in \{0,1\}$ and the prior is uniform, $q(x) = 1/2$). Second, suppose alternatively that the expert's ability is publicly revealed in the second period after the expert is hired, but before the decision is taken. When hired, the expert is rewarded on the basis of the expected reputational value (2.3).

More generally, information about ability can have positive value, as noticed by Li [18].

next section, these facts have implications for the signal that experts would like to convey depending on their posterior belief about the state.

These key features of the signal structure allow us to greatly simplify the expert's preferences and to show that they satisfy the following properties:

Proposition 1. *It is without loss of generality to represent the expert's preferences by*

$$W(m|x) = \frac{mx}{1 + mxEt}, \qquad (3.1)$$

where each message m is identified with the corresponding expected signal $E[s|m] = \int_S s\hat{\mu}(m|s)\,ds / \int_S \hat{\mu}(m|s)\,ds$, calculated using the evaluator's conjecture $\hat{\mu}$. $W(m|x)$ is strictly supermodular in (m, x), and strictly concave in m when $x \neq 0$. In turn, $V(m|s)$ is (i) strictly concave in m, and (ii) strictly supermodular in (m, s). Finally, (iii) $V_m(s|s)$ is continuous in s, and $V_m(s|s) = 0$ implies that $V_m(s'|s') < 0 < V_m(s''|s'')$ for all $s' > s > s''$.

The expert's payoff from message m thus depends only on the evaluator's beliefs about the average signal $E[s|m]$ of an expert sending message m. Intuitively, this property derives from the fact that the expert's reputation is an increasing function of sx, which is linear in s. Of course, $W(m|x)$ inherits the property of being an increasing function of mx.

Note that it is technically possible that two separate messages m and m' result in $E[s|m] = E[s|m']$. According to Proposition 1, in this case the expert must be indifferent between these two messages. Moreover, the two messages convey identical information, not only about t, but also about x, since they give the same interaction term $E[s|m]tx$. It is then without loss of generality to regard them as one. Of course, the "merged" message $m'' = m \vee m'$ retains the common $E[s|m''] = E[s|m] = E[s|m']$.

Note that $W(m|x)$ in (3.1), and thus $V(m|s)$, depend on the prior reputation $p(t)$ only through the expected value Et. This is a consequence of the linearity in t of the signal's density. Moreover, in our model (2.1) higher mx is better news about t, so that the posterior reputations $p(t|x, m)$ are unambiguously ranked in the first-order stochastic order, depending on the pair (x, m). Our results then hold for any strictly increasing utility function $v(t)$.

Proposition 1 allows us to identify the message m with the evaluator's conjectured $E[s|m]$. Our game can then be seen as a version of Crawford and Sobel's [8] sender-receiver game, in which s is the sender's private information, m is the receiver's decision, and the receiver has quadratic payoff $-(m - s)^2$. The payoff function $V(m|s)$ endogenously derived for our sender is fairly simple and satisfies some of the properties (strict concavity and increasing differences) assumed by [8]. Differently from their setting, we show below

that the sender does not always want to bias the receiver's decision in one particular direction.

4. Optimal Deviation from Truthtelling

In order to check for the existence of fully revealing equilibria in cheap talk games, it is without loss of generality to restrict the expert to send truthful signals. This is because in cheap talk games the language is arbitrary and incentives are driven by the evaluator's understanding of the meaning of the messages sent in equilibrium. The evaluator uncovers correctly the expert's signal from the message by inverting the expert's strategy. By ruling out the existence of a truthtelling equilibrium, we will conclude that our cheap talk game does not admit any fully revealing equilibrium.

We now show that if the evaluator conjectures that the expert is truthfully reporting the signal, the expert gains from deviating in certain directions:[9]

Proposition 2. *Assume that the evaluator conjectures truthful reporting of all s. Then there are three mutually exclusive cases: (a) If $\int_X \left(x / \left(1 - xEt \right) \right) q(x) \, dx \leq 0$, all experts with $s > -1$ have $V_m(s|s) < 0$ and wish to deviate to a lower signal $s' \in [-1, s)$. (b) If $\int_X \left(x / \left(1 + xEt \right) \right) q(x) \, dx \geq 0$, all experts with $s < 1$ have $V_m(s|s) > 0$ and wish to deviate from truthtelling to a higher signal $s' \in (s, 1]$. (c) Otherwise, there exists a unique signal $\hat{s} \in (-1, 1)$ such that all experts with $s \neq \hat{s}$ have $V_m(s|s) \neq 0$ and wish to deviate from truthtelling to a signal s' strictly between s and $\hat{s}$.*

This result provides a complete characterization of the direction of the expert's bias. Since $x / \left(1 - xEt \right)$ and $x / \left(1 + xEt \right)$ are both increasing functions of x, the condition $\int_X \left(x / \left(1 - xEt \right) \right) q(x) \, dx \leq 0$ for case (a) arises when the prior $q(x)$ puts enough weight on low realizations of x, while case (b) results when high states are likely. In order to uncover the logic of the model, we now discuss the different deviation incentives in the three cases defined by this proposition.

In both cases (a) and (b), the expert desires messages m which are ex-ante likely to be of the same sign as the state x and therefore result in high reputations. The condition for case (a), $\int_X \left(x / \left(1 - xEt \right) \right) q(x) \, dx \leq 0$, is equivalent to $V_m(-1|-1) \leq 0$, meaning that even at the left-most signal realization there is a weak desire to bias the report downwards. The proof employs the properties of V established in Proposition 1 to show that this bias then holds strictly at all higher signals. The result derives from the fact that the expert

---

[9]For another similar but independently derived result see Campbell's [5] Proposition 3.1.

desires a high value of mx. When high states are sufficiently likely, the expert wishes to convey a signal m higher than the s actually possessed.

Case (c) arises for a middle range of prior beliefs. When the prior on the state x is not too biased in either direction, any given message m is likely to result in both positive and negative values of mx. In this case, the expert has an incentive to bias the report in a different direction depending on the signal observed. From $V_m(-1|-1) > 0$ and $V_m(1|1) < 0$, the bias is towards the middle.[10]

Due to these incentives to deviate, the expert's behavior is not compatible with the evaluator's conjecture that the expert is telling the truth. Rationality of the evaluator makes this incentive self defeating. This is a generic property of the reputational cheap talk model, well beyond the specification of $f(s|x,t)$ imposed in (2.1). As shown by Ottaviani and Sørensen [24], truthtelling cannot be an equilibrium in any generic smooth model in which the combined observation of the signal and the state contains information about the expert's ability. With the signal structure considered in this paper, we show below that equilibria with partial revelation arise only when the bias is towards the middle (case (c) above).

5. Equilibrium

We now derive our main characterization results. As is typical of cheap-talk games, there is always a pooling (or "babbling") equilibrium in which no information is transmitted. If the evaluator expects all messages to be uninformative, the experts have no choice but to babble. Rather than discussing equilibrium selection, we investigate all equilibria. Informative equilibria must be binary (Section 5.1), and there is no informative equilibrium at all when the prior belief on the state is sufficiently concentrated (Section 5.2). After presenting an extended example with binary state (Section 5.3), we briefly discuss some comparative statics results (Section 5.4). Finally, we outline some broader implications and discuss the importance of our assumptions (Section 5.5).

---

[10] Although driven by different forces, this last case is reminiscent of the "yes-men" effect analyzed by Prendergast [27]. He considers an agent with access to two private signals, one on the state of the world and the other on the principal's private signal on the state. The principal commits to a reward scheme based on the difference between the agent's report and the principal's signal. As a result, the agent honestly reports her best estimate of the principal's signal. But the principal cannot fully infer the agent's direct signal about the state, so information transmission is inefficient. Ewerhart and Schmitz [10] have argued that efficiency can be restored by asking the agent to also report her private information.

5.1. Binary Equilibria

Equilibria have a partition structure whereby connected sets of signals are pooled. We noticed in (ii) of Proposition 1 that $V(m|s)$ is strictly supermodular in the pair (m, s), i.e., $V(m'|s) - V(m|s)$ is strictly increasing in s whenever $m' > m$. Incentive compatibility implies that all expert types sending message m' have higher signals than those sending m. This implies that each message m sent in equilibrium corresponds to signals that belong to some interval subset of S. By Proposition 2 we know that there cannot be truthful reporting in any subinterval of S, for this would imply $V_m(s|s) = 0$ on that interval. Hence, the typical message interval has a non-empty interior, although there might be occasional isolated one-point intervals:[11]

Proposition 3. *All equilibria have interval messages.*

We now establish that all informative equilibria are binary: at most two messages are sent, one for $s \geq a$ and the other for $s < a$, where $a \in (-1, 1)$.[12]

Proposition 4. *All informative equilibria are binary.*

Some intuition for this result can be gained by considering the special case with $Ex = 0$. In this case, simple algebra yields $V(m|s) = (m - s)k(m)$ where the function k is continuously differentiable, strictly decreasing, and satisfies $k(0) = 0$. Suppose that $m' > m$ form an arbitrary pair of equilibrium messages. Any equilibrium message expresses the average signal in the interval where it is optimal. By continuity of V, we can then find some signal $a \in (m, m')$ with $V(m|a) = V(m'|a)$. This indifference equation implies that $a = (mk(m) - m'k(m'))/(k(m) - k(m'))$. The fact that $a < m'$ then reduces to $k(m) > 0$, which implies $m < 0$. We have concluded that the lower message must be negative. Symmetrically, the higher message must be positive. Since the message pair was arbitrarily selected, we cannot have three or more messages in equilibrium. The proof of Proposition 4 shows that this logic extends beyond the case with $Ex = 0$.

---

[11]It is technically possible to construct non-interval equilibria in which two messages m and m' sent in equilibrium have $E[s|m] = E[s|m']$. As explained above in the discussion of Proposition 1, without loss of generality we collapse such messages into one with the same expected signal and the same informational content.

[12]When $Ex = 0$, there exists an equilibrium with messages -1, 0, and 1. We consider this equilibrium uninformative, since messages -1 and 1 are only sent with probability zero (i.e., when the realized signals are -1 and 1).

5.2. Further Characterization of Equilibria

We now characterize more precisely the conditions for the existence of binary equilibria. Let m denote the message sent for $s \in [-1, a]$ and m' the message sent for $s \in (a, 1]$, with $-1 < a < 1$. The low message has average signal $E[s|m] = (a-1)/2$, the high has $E[s|m'] = (a+1)/2$. Notice that the distance between those averages is always $1 = (a+1)/2 - (a-1)/2$. The indifference condition $V(m|a) = V(m'|a)$ is

$$\int_X \frac{x(1 + axEt)}{[2 + (a-1)xEt][2 + (a+1)xEt]} q(x)\, dx = 0. \tag{5.1}$$

Messages $[-1, a]$ and $(a, 1]$ constitute a binary equilibrium if and only if $a \in (-1, 1)$ solves this equation.

Proposition 5. *If a binary equilibrium exists, then $\int_X (x/(1 + xEt))\, q(x)\, dx < 0 < \int_X (x/(1 - xEt))\, q(x)\, dx$.*

Combining Propositions 2 and 5, we observe that the binary equilibrium exists only when the bias from truthtelling is towards the middle. In Crawford and Sobel's [8] model, informative equilibria can instead arise when the sender wants to bias the report in a uniform direction regardless of his signal realization (i.e., with $V_m(s|s)$ of constant sign).[13]

We now investigate three properties of the equilibria of this game that are particularly useful for the dynamic extension contained in Section 6. First, we obtain a sufficient condition for the existence of informative equilibria. Second, we characterize whether there are multiple informative equilibria. Third, we consider what happens when the prior belief is concentrated on a particular state.

First, we identify a prominent instance in which a binary equilibrium exists:

Proposition 6. *If the prior on x is symmetric around 0, there exists a symmetric binary equilibrium with messages $[-1, 0]$ and $(0, 1]$.*

The expert learns nothing about the state when receiving the signal $s = 0$, and with the symmetric prior (and posterior) on the state it is equally attractive to send either the high or the low message.

---

[13] Admati and Pfleiderer [1] consider a cheap talk game where the receiver has utility $-(m-s)^2$ and the sender has utility $-(m - \lambda s)^2$ where $\lambda > 0$. They focus on the case of an overconfident sender with $\lambda \geq 1$, in which case the sender has a bias away from the middle, since $V_m(s|s)$ has the same sign as s. It follows from their analysis that there exist equilibria with any finite number of messages. It can be proved that at most two messages can be sent in equilibrium in a version of their model with $\lambda \leq 1/2$, representing strong underconfidence of the sender. The bias towards the middle induced by underconfidence has a similar effect to the reputational bias resulting in our model.

Second, if there is an informative equilibrium in this model, then there is typically another informative equilibrium:

Proposition 7. *For a generic prior belief on the state, the number of binary equilibria is even.*

This multiplicity can be better understood by considering the equilibrium condition, $V(m'(a)|a) - V(m(a)|a) = 0$, where $m'(a)$ denotes the high message $m'(a) = (1+a)/2$ and $m(a) = (a-1)/2$ the low message. On the one hand, supermodularity of V implies that $V(m'(a)|s) - V(m(a)|s)$ is increasing in s for any a. This means that holding fixed the evaluator's beliefs, the expert's preference for the high message increases with the level of the signal. On the other hand, $V(m'(a)|s) - V(m(a)|s)$ decreases in a for any s. This means that holding fixed the expert's signal, the higher message becomes less appealing when the evaluator's belief increases. The balance between these two opposite effects determines whether $V(m'(a)|a) - V(m(a)|a)$ is increasing or decreasing in a. The multiplicity of equilibria is due to the fact that $V(m'(a)|a) - V(m(a)|a)$ is equal to zero for several values of a.

Third, if the prior is highly concentrated on some particular state $x > 0$, the posterior $q(x|s)$ is still concentrated around x, due to the fact that the signal s is of bounded informativeness about the state. Whenever the state turns out positive, it is favorable to the expert's reputation to report a message with $m > 0$. Thus, in this case, if there exists a message with $E[s|m] > 0$, the expert would want to send this message given any realization of the signal s. This cannot hold in equilibrium, since signals are equal to zero on average.

Proposition 8. *If the prior distribution on the state is sufficiently concentrated on any $x \neq 0$, there exists no informative equilibrium.*

Note that this result does not hold for $x = 0$, since Proposition 6 guarantees the existence of an informative equilibrium for any (symmetric) prior arbitrarily concentrated on 0.

5.3. Binary State Example

The equilibria of this game are well illustrated in a variation of the model with prior $q(x)$ concentrated on states -1 and $+1$. Denote the prior probability of state $+1$ by q. In this example, the continuous private signal implies a continuously varying posterior probability
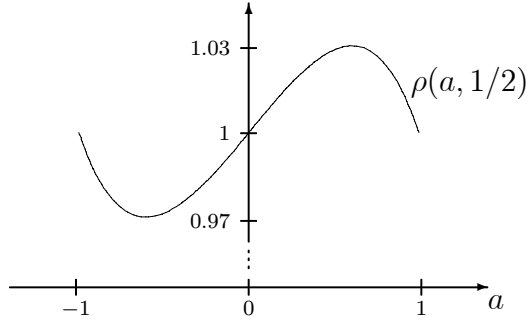
Fig. 3. Graph of the right-hand side of (5.2) when $Et = 1/2$.

of state $+1$. Our game then concerns communication of this posterior probability. In this special case, (5.1) can be re-written as

$$\frac{q}{1-q} = \frac{1-aEt}{1+aEt}\frac{2+(a-1)\,Et}{2-(a-1)\,Et}\frac{2+(a+1)\,Et}{2-(a+1)\,Et}, \tag{5.2}$$

a third-order polynomial equation in a. Denote the right-hand side of (5.2) by $\rho(a, Et)$, plotted in Fig. 3 for $Et = 1/2$. For $q = 1/2$ the equation has one solution in $(-1, 1)$, so that $a = 0$ is an equilibrium. The equation has two solutions in $(-1, 1)$ when $q \in (1 - \bar{q}, \bar{q})$ with $\bar{q} > 1/2$, one solution when q is equal to $1 - \bar{q}$ and $\bar{q}$, and no solution in $(-1, 1)$ whenever $q > \bar{q}$ and $q < 1 - \bar{q}$. For the special case with $Et = 1/2$ illustrated in the figure, we have $\bar{q} = \frac{1}{2} + \left(\frac{4-\sqrt{11}}{60}\right)\sqrt{\left(7 - 2\sqrt{11}\right)} = .5069$. It is easy to show that there can be no solution to (5.2) for $q < 1/3$ or $q > 2/3$, no matter how good is the prior reputation Et. No expert can speak credibly for these prior beliefs.

5.4. Comparative Statics

We now perform two natural comparative statics exercises.

Do Better Reputed Experts Send More Informative Messages? By assumption, ex-ante better experts are Blackwell-better informed about x. We now show that better experts' messages need not be Blackwell-more informative. This effect is due to the equilibrium garbling of the signal. Take $q = .505$ in the binary-state illustration above and consider two experts, the first with $Et = .49$ and the second with $Et = .5$. Assume that the informative equilibrium with a threshold nearest to 0 has been selected. When message $s \in [-1, a]$ is observed, the posterior belief is

$$q(x = 1|m = [-1, a]) = \frac{[2 + (a-1)Et]\,q}{[2 + (a-1)Et]\,q + [2 - (a-1)Et]\,(1-q)}.$$

14

A similar expression defines $q(x = 1|m = (a, 1])$. The following is based on numerical solution of equation (5.2). The expert with $Et = .5$ has an equilibrium with $a = .329$ yielding $q(x = 1|m = [-1, a]) = .42101$ and $q(x = 1|m = (a, 1]) = .67058$. The expert with $Et = .49$ has an equilibrium with $a = .357$ resulting in $q(x = 1|m = [-1, a]) = .42617$ and $q(x = 1|m = (a, 1]) = .67072$. In a decision problem with two actions and indifference at a belief in the interval $(.67058, .67072)$, the expert with $Et = .5$ is of no value while the expert with $Et = .49$ transmits valuable information.

According to this counter-example, the message reported by a more informed expert need not be more informative. Note that this result does not invalidate our assumption that experts who are believed to be better informed have higher rewards. As explained in Section 2, an increasing reputational payoff function can be derived from the truthtelling equilibrium resulting in a second and last period in which the expert is consulted. Nevertheless, this perverse comparative statics result suggests that it would be difficult to endogenously derive a monotonic reputational payoff in a dynamic version of the model with more than two periods.

Are Better Reputed Experts Credible for a Larger Set of Priors? Consider an expert B with a prior reputation $p(t)$ which is higher than the prior reputation of expert A in the sense of first-order stochastic dominance (FSD). Is B credible for a wider set of priors on the state of the world than A? As seen from equation (5.1), credibility depends on the prior reputation only through Et. With two states of the world, the right-hand side of (5.2) is increasing in Et for any $a \in (0, 1)$. In this binary state example, the credibility region increases in the prior reputation.

5.5. Discussion

We have found that the most informative equilibrium is either binary or completely uninformative. Since reported messages pool many signals, they are far less precise than the expert's true signal. The expert can communicate at most the direction of her information, but is unable to credibly convey its intensity. The equilibrium loss of information typically results in a welfare loss for those who listen to the expert's advice (decision makers), compared to the case of truthful revelation.[14] In addition, future decision makers will have less information about the expert's true ability. If they value information about the expert's

---

[14]If the utility function $v(t)$ is linear, the expert's ex-ante expected reputational value of sending any message profile is equal to its prior value. In this case, the expert is therefore indifferent in ex-ante terms between the different equilibria.

ability, they would prefer the signal not to be garbled.[15] Overall, the loss of information resulting in equilibrium is typically undesirable.

We now report on some robustness checks. First, our results extend to cases in which the expert also cares about the accuracy of the decision made on the basis of the message reported. This can be done by adding a second receiver who uses the expert's message to make a decision under uncertainty about the state of the world. The expert's payoff is assumed to depend on the reputational payoff as well as on the payoff of this second receiver. It can be shown that all informative equilibria are binary, provided that the expert puts enough weight on the reputational payoff.

Second, the result that all informative equilibria must be binary (Proposition 4) is striking, but also quite special to the multiplicative linear experiment (2.1). To see this, we now construct an example of a signal structure resulting in an informative equilibrium with three messages. Let $f(s|x,t) = tg(s|x) + (1-t)h(s)$, with $g(s|x) = (1+sx)/2$ and $h(s) = \gamma + \delta s^2$, where $\gamma = 5/12$, $\delta = 1/4$. Compared to (2.1), we have made the middle message more attractive by requiring the uninformed $h(s)$ to put greater weight on extreme signals. Setting the priors to $\Pr(x=1) = \Pr(x=-1) = 1/2$ and $Et = 1/2$, it is possible to verify numerically that the three messages $\{[-1, -a], (-a, a], (a, 1]\}$ with $a \approx .80218$ constitute an equilibrium.

Third, Proposition 8 relies on the fact that the multiplicative linear signal is boundedly informative about the state. This result does not hold in the normal learning model, for which Ottaviani and Sørensen [23] have shown that there is always a binary equilibrium.

Finally, since experts specialize in giving information, it is natural to consider what happens when the *same expert* repeatedly reports information about *different states* of the world. Since the equilibrium message contains only part of the expert's information, the expert typically acquires private information about her own ability over time. In order to check the robustness of our results, the model can be extended to allow for the expert to have private information about her own ability, as first posited by Trueman [34] in a discrete signal model. It can be shown that there is then an additional exaggeration incentive, akin to the one discussed by Prendergast and Stole [28] in a context in which the evaluator does not observe ex post the state realization. The coarseness result is robust, but with *perfect* knowledge on own ability there is always an informative equilibrium and so Proposition 8 does not hold.

---

[15]This would clearly result if future employers are short-run players. Given that information is valuable to these employers, information on the quality of information received is also valuable, because it allows them to make better decisions. This information is also valuable in some dynamic settings with competition. For example, Prat [26] develops two simple two-period settings in which future employers necessarily benefit from information about ability.

6. Reputational Herding

We now turn to the analysis of sequential reporting by *different experts* about the *same state* of the world. The model of Section 2 is extended as follows. Experts are indexed by $i \in \mathbb{N}$ and arrive in an exogenously given sequence. Let $h_1 = \varnothing$ and $h_i = \{m_1, ..., m_{i-1}\}$ for $i > 1$ denote the public history of messages reported. Expert i of unknown ability $t_i \in [0, 1]$ observes the private signal $s_i \in [-1, 1]$ with conditional p.d.f. $f(s_i|x, t_i)$ given by (2.1) as well as the public history h_i, and then publicly sends a message m_i. After all experts have spoken, the evaluator observes the state of the world $x \in [-1, 1]$ and computes the posterior reputation. There is a common prior distribution on state and abilities, such that $\{x, t_1, t_2, ...\}$ are independent.

We make the natural assumption that the experts' private signals are independent conditionally on $\{x, t_1, t_2, ...\}$. Scharfstein and Stein [29] instead assumed that more able experts received signals more positively correlated, conditionally on the state of the world. Conditional independence implies that once x has been observed, signal s_i is informative about t_i but independent of t_j for all $j \neq i$. Each expert's message therefore carries information about the expert's own ability, but not about the ability of others. This greatly simplifies the dynamic strategic analysis, because expert i has no incentive to manipulate the reports of experts who speak later. A contingent strategy for expert i is a random map from s_i to m_i (as before), with the added feature that it depends on the public history of messages.

Since the history is common knowledge among the expert and the evaluator, in any contingency, the expert and evaluator face the same problem as in our static analysis. Our characterization of the single expert problem directly carries over to this dynamic setting. The only difference is that the prior belief $q(x)$ about the state of the world depends on the public history of messages reported.

Specifically, when an expert of expected ability Et_i issues message m_i, identified as before with the average of his signals leading to this report, the belief on the state is updated from the prior $q_i(x)$ to the posterior

$$q_{i+1}(x) = \frac{\hat{f}(m_i|x)}{\hat{f}(m_i)} q_i(x) = \frac{1 + m_i x E t_i}{1 + m_i E_i x E t_i} q_i(x),$$

where $E_i x$ denotes the expectation of x under the belief q_i. The possible message values m_i depend on the static equilibrium played with the i-th expert. If the expert is uninformative ($m_i = 0$), the posterior q_{i+1} is equal to the prior q_i. If the equilibrium is binary, then the high message is some $m \in (0, 1)$ and the low message is $m - 1$. By the MLRP of $f(s_i|x, t_i)$ in (s_i, x), the high message is good news about x, while the low message is bad news.

Thus, when the high message is observed, q_{i+1} is FSD-higher than q_i, while the opposite conclusion holds after the low message.

Proposition 8 is crucial for the outcomes of this herding model. Experts learn about the state by listening to the messages reported by the previous experts. As long as experts speak informatively, the law of large numbers implies that the beliefs of later experts become ever more concentrated on the true state x. According to Proposition 8, once the belief becomes sufficiently concentrated, experts can no longer be informative, and so learning stops completely before the true state is revealed (except if beliefs are concentrated on $x = 0$). All later experts engage in herd behavior, as in the statistical herding model (Banerjee [2], Bikhchandani, Hirshleifer and Welch [3], and Smith and Sørensen [32]).[16]

Proposition 9. *Suppose that there exists an $\varepsilon > 0$ such that (i) all experts satisfy $Et_i > \varepsilon$, and (ii) whenever expert i speaks informatively, the high equilibrium message satisfies $m_i \in (\varepsilon, 1 - \varepsilon)$. Contingent on any $x \neq 0$, with probability one there is only a finite number of experts who speak informatively in equilibrium.*

Assumptions (i) and (ii) require that an expert who speaks informatively reports an amount of information bounded away from zero. The first assumption imposes a lower bound on the exogenous expected talent of the experts, while the second assumption is about the amount of information contained in the (endogenously determined) equilibrium messages sent by each expert.[17] This proposition leaves open the possibility that all experts speak informatively forever, but in this case the amount of information revealed in each round must necessarily converge to zero. According to this result, reputational herding results with conditionally independent signals.

6.1. Binary State Example

As a more detailed illustration of the working of the dynamic model, consider the case of two experts with prior expectations on abilities $Et_1 = Et_2 = 1/2$. Assume that the state is binary, with initial fair prior $q_1(1) = q_1(-1) = 1/2$, so that the prior expectation is $E_1 x = 0$. It follows from the analysis of the one-period problem that in the most informative equilibrium the first expert reports the high message $1/2$ when observing $s_1 \in (0, 1]$ and the low message $-1/2$ when observing $s_1 \in [-1, 0]$.

---

[16] As observed by Lee [17] in the context of statistical herding, for herding to arise it is necessary that equilibrium actions do not perfectly reveal signals. See Ottaviani and Sørensen [22] for more on the connection between statistical and reputational herding models.

[17] The binary state example illustrated in Figure 3 provides us with an example of an informative equilibrium close to pooling that contains increasingly less information as q converges to $1/2$.

In the second period, the belief on the state upon observation of message $1/2$ is $q_2(1) = \Pr(x = 1|m_1 = 1/2) = (1 + 1/4)/2 = 5/8$. The equilibrium for the second expert is then determined by Fig. 3. Herding results since $q_2(1) = 5/8 > \bar{q} = .5069$ for $Et_2 = 1/2$. When expert 1 sends a high message, so does expert 2. Similarly, when 1 sends the low message, 2 also sends a low message.

In conclusion, the fact that differential conditional correlation is necessary to obtain herding is not a robust finding of the binary signal model, even when Scharfstein and Stein's [30] strong definition is used: here "2 *always* ignores his own information and follows agent 1" under conditional independence.[18]

7. Conclusion

Increasing specialization of labor suggests that information should be collected by professional experts. These experts are interested in their reputation for being well informed. In our model, experts are evaluated on the basis of all ex-post available information, in the absence of explicit monetary rewards.

Compared to the partisan expert case, the analysis is complicated by the necessary presence of simultaneous learning on state and ability. We have managed to make the problem tractable by focusing on a special but natural information structure, which extends the binary signal model to allow for varying intensity.

When the signal and the state can be cross-checked to update beliefs about the expert's ability, the equilibrium cannot be fully separating. In a putative fully separating equilibrium, the signal and the realized state are informative about ability, giving an incentive to the expert to manipulate the report of the signal in order to generate a better reputation. If the evaluator naively believes that the expert is telling the truth, the expert has an incentive to deviate towards an intermediate signal, provided that the prior on the state is not too extreme. If the prior on the state is strongly biased in one direction, the expert has an incentive to report even more extreme posteriors. These deviation incentives are clearly incompatible with equilibrium behavior.

Equilibria can be taken to have a partition structure. In this model, the most informative equilibria are typically multiple and do not involve more than two messages. When the prior is sufficiently concentrated on any state, the only equilibrium is pooling. These results could not be easily foreseen in a two-signal model, but hold in this natural

---

[18]Note that this conclusion is valid regardless of whether Scharfstein and Stein's [29] non-informativeness condition is satisfied or not. The non-informativeness condition is satisfied in this example, but would fail for a slightly different prior on the state.

generalization.

These results have implications for dynamic extensions of the model in which experts sequentially report information for fixed state. If experts have conditionally independent signals, the model can be solved forward. Reputational herding results eventually provided that experts do not have perfect knowledge about their own ability, even in the absence of differential conditional correlation.

Appendix A: Proofs

Proof of Proposition 1. Using $\hat{\mu}$, we first compute $\hat{g}(m|x) = \int_S \hat{\mu}(m|s)g(s|x)\,ds$, $\hat{h}(m) = \int_S \hat{\mu}(m|s)h(s)\,ds$, $\hat{f}(m|x,t) = t\hat{g}(m|x) + (1-t)\hat{h}(m)$ and $\hat{f}(m|x) = \int_T \hat{f}(m|x,t)p(t)\,dt$. Substitution in (2.3) gives, after simple algebraic manipulation,

$$
\begin{aligned}
V(m|s) &= \int_X \int_T v(t)\frac{t\hat{g}(m|x) + (1-t)\hat{h}(m)}{E[t]\hat{g}(m|x) + (1 - E[t])\hat{h}(m)}p(t)\,dt\,q(x|s)\,dx \\
&= E[v(t)] + (E[tv(t)] - E[t]E[v(t)])\int_X \frac{\hat{g}(m|x) - \hat{h}(m)}{\hat{f}(m|x)}q(x|s)\,dx, \quad \text{(A.1)}
\end{aligned}
$$

Notice that the constant $E[tv(t)] - E[t]E[v(t)]$ is strictly positive when v is strictly increasing and t does not have a degenerate prior distribution. Applying a positive affine transformation to (A.1), the expert's payoff (2.3) can therefore be rewritten as

$$
V(m|s) = \int_X \frac{\hat{g}(m|x) - \hat{h}(m)}{\hat{f}(m|x)}q(x|s)\,dx.
$$

Using again (2.1), we have $\hat{f}(m|x,t) = (1 + E[s|m]tx)\int_S \hat{\mu}(m|s)ds/2$, and we arrive at (3.1),

$$
W(m|x) = \frac{\hat{g}(m|x) - \hat{h}(m)}{\hat{f}(m|x)} = \frac{E[s|m]x}{1 + E[s|m]xEt} \equiv \frac{mx}{1 + mxEt}.
$$

Notice that $Et < 1$ since the distribution of t is non-degenerate, so the denominator is bounded away from zero.

Differentiation of $W(m|x)$ gives $W_m(m|x) = \frac{x}{(1+mxEt)^2}$. Then we have $W_{mx}(m|x) = \frac{1-mxEt}{(1+mxEt)^3} > 0$, verifying strict supermodularity. Also $W_{mm}(m|x) = -\frac{2x^2Et}{(1+mxEt)^3} < 0$ for $x \neq 0$, verifying strict concavity.

Using (2.3) we have $V_m(m|s) = \int_X W_m(m|x)q(x|s)\,dx$. Property (i) then follows from $V_{mm}(m|s) = \int_X W_{mm}(m|x)q(x|s)\,dx < 0$, which holds because $W_{mm}(m|x) < 0$ except at $x = 0$ and the posterior $q(x|s)$ is not degenerate at $x = 0$, since the prior $q(x)$ is non-degenerate and the signal cannot reveal $x = 0$.

To obtain (ii), note that $V_{ms}(m|s) > 0$ follows by Milgrom's [20] Proposition 1 from the facts that $W_{mx}(m|x) > 0$ and $f(s|x)$ satisfies the MLRP.

To obtain (iii), observe that

$$
V_m(s|s) = \frac{1}{1 + sExEt}\int_X \frac{x}{1 + sxEt}q(x)\,dx. \quad \text{(A.2)}
$$

Thus $V_m(s'|s') \gtrless 0$ if and only if $\int_X (x/(1 + s'xEt))\,q(x)\,dx \gtrless 0$. Note that this expression is strictly decreasing in s' since the derivative is $\int_X -(x^2Et/(1 + s'xEt)^2)\,q(x)\,dx < 0$. Using the assumption that $\int_X (x/(1 + sxEt))\,q(x)\,dx = 0$, we obtain the desired signs of $V_m(s'|s')$ and $V_m(s''|s'')$. $\quad\square$

Proof of Proposition 2. We apply the properties (i)–(iii) from Proposition 1. By (iii), if $V_m(-1|-1) \leq 0$, then any $s > -1$ satisfy $V_m(s|s) < 0$. By (i), $V(m|s)$ is concave in m, so the optimal message at s is in the interval $[-1, s]$. Likewise, if $V_m(1|1) \geq 0$ then all $s < 1$ have $V_m(s|s) > 0$ and an optimal response in $(s, 1]$. Finally, if $V_m(1|1) < 0 < V_m(-1|-1)$, the continuity of $V_m(s|s)$ guarantees the existence of $\hat{s} \in (-1, 1)$ such that $V_m(\hat{s}|\hat{s}) = 0$. By (iii), $\hat{s}$ is unique, and now $V_m(s'|s') < 0 < V_m(s''|s'')$ for all $s' > \hat{s} > s''$. By (ii), $V_m(\hat{s}|\hat{s}) = 0$ implies that $V_m(\hat{s}|s'') < 0 < V_m(\hat{s}|s')$. Since $V_m(s'|s') < 0 < V_m(\hat{s}|s')$, the concavity of V in m implies that the optimal response to s' is in $(\hat{s}, s')$. A similar argument applies to s''. Finally, insertion into (A.2) shows that $V_m(-1|-1) \leq 0$ if and only if $\int_X (x/(1 - xEt)) q(x) \, dx \leq 0$, and that $V_m(1|1) \geq 0$ if and only if $\int_X (x/(1 + xEt)) q(x) \, dx \geq 0$.

Cases (a) and (b) are mutually exclusive as a consequence of Proposition 1, (iii). Indeed, it can also be verified directly that $\int_X (x/(1 - xEt)) q(x) \, dx - \int_X (x/(1 + xEt)) q(x) \, dx = \int_X \left(x^2 Et / \left(1 - x^2 (Et)^2 \right) \right) q(x) \, dx > 0$. $\square$

Proof of Proposition 3. Assume to the contrary, that message m is sent at signal s while message $m' \neq m$ is sent at signals s', s'' where $s' < s < s''$. Optimizing behavior of the expert implies that $V(m|s) \geq V(m'|s)$ and $V(m'|s') \geq V(m|s')$ and $V(m'|s'') \geq V(m|s'')$. In case $m' > m$, the relation $V(m'|s') - V(m|s') \geq 0 \geq V(m'|s) - V(m|s)$ implies with strict supermodularity of V that $s' \geq s$, a contradiction. In case $m' < m$, the relation $V(m|s) - V(m'|s) \geq 0 \geq V(m|s'') - V(m'|s'')$ likewise implies that $s'' \leq s$, another contradiction. In any case, a contradiction arises, so the contrary assumption must be false. $\square$

Proof of Proposition 4. Suppose that three or more distinct messages were sent in equilibrium. Propositions 2 and 3 imply that there must be at least two message intervals with a non-empty interior, except if the equilibrium has message intervals $[-1, -1], (-1, 1), [1, 1]$, but this latter strategy is almost surely equivalent to a one-message strategy. For the remainder of the proof we can therefore assume that three signal intervals $[a, b], [b, c], (c, d]$ define adjacent equilibrium messages, with $a < b \leq c < d$. Letting $m = E[s|s \in [a, b)]$, $m' = E[s|s \in [b, c]]$, and $m'' = E[s|s \in (c, d]]$, we have then $m < b \leq m' \leq c < m''$.

By incentive compatibility and payoff continuity, an individual with signal $s = b$ must satisfy indifference $V(m|b) = V(m'|b)$. Using (3.1), the difference $V(m'|b) - V(m|b)$ is

$$\int_X \left(\frac{m'x}{1 + m'xEt} - \frac{mx}{1 + mxEt} \right) q(x|b) \, dx = \frac{(m' - m)}{1 + bExEt} \int_X \frac{x(1 + bxEt)}{(1 + m'xEt)(1 + mxEt)} q(x) \, dx.$$

Eliminating the strictly positive factors $m' - m$ and $(1 + bExEt)$, the condition $V(m|b) = V(m'|b)$ reduces to

$$\int_X \frac{x\,(1 + bxEt)}{(1 + mxEt)\,(1 + m'xEt)} q\,(x)\ dx = 0. \tag{A.3}$$

Proceeding similarly, we have that $V(m''|c) \lesseqgtr V(m'|c)$ whenever

$$\int_X \frac{x\,(1 + cxEt)}{(1 + m'xEt)\,(1 + m''xEt)} q\,(x)\ dx \lesseqgtr 0. \tag{A.4}$$

Subtracting this condition (A.4) from (A.3), $V(m''|c) \lesseqgtr V(m'|c)$ and $V(m|b) = V(m'|b)$ hold whenever

$$\int_X \left[\frac{1 + bxEt}{1 + mxEt} - \frac{1 + cxEt}{1 + m''xEt} \right] \frac{x}{1 + m'xEt} q(x)\,dx \gtreqless 0. \tag{A.5}$$

We now verify that the integral on the left-hand side of (A.5) is necessarily positive, showing that the expert with $s = c$ has a strict incentive to deviate to message m', implying the failure of this putative equilibrium. This verification relies on three steps. First, we show that the factor in square brackets has the same sign as x. It certainly vanishes at $x = 0$. The MLRP of $f\,(s|x,t) = (1 + stx)/2$ implies that $(1 + bxEt)/(1 + mxEt)$ is strictly increasing in x since $b > m$. Likewise, $(1 + cxEt)/(1 + m''xEt)$ is strictly decreasing in x since $c > m''$. We conclude that the difference is strictly increasing, and therefore has the same sign as x. Second, the other integrand factor $x/(1 + m'xEt)$ is immediately seen to have the same sign as x. Finally, since both factors in the integrand have the same sign as x, the integrand must be strictly positive at $x \neq 0$. Since q is not degenerate by assumption, the integral in (A.5) is strictly positive. $\qquad\square$

Proof of Proposition 5. We have noticed that the distance between the messages is 1. Using (3.1), observe that for all $m \in [-1, 0]$,

$$\begin{aligned}
V\,(m + 1|s) - V\,(m|s) &= \int_X \frac{x}{[1 + (m + 1)\,xEt]\,[1 + mxEt]} q(x|s)\,dx \\
&= \frac{1}{1 + sExEt} \int_X \frac{x\,(1 + sxEt)}{[1 + (m + 1)\,xEt]\,[1 + mxEt]} q(x)\,dx.
\end{aligned}$$

At $s = m$, we find $V\,(m+1|m) - V\,(m|m) > 0$ if and only if $\int_X (x/[1+(m+1)\,xEt])\,q(x)\,dx > 0$. The latter expression is a strictly decreasing function of m, since its derivative is $\int_X -\left(x^2 Et/[1+(m+1)\,xEt]^2\right) q(x)\,dx < 0$. Likewise, at $s = m+1$ we find $V\,(m+1|m+1) - V\,(m|m+1) < 0$ if and only if $\int_X (x/[1 + mxEt])\,q(x)\,dx < 0$, which is again a strictly decreasing function of m.

Assume now that $\int_X (x/(1 - xEt)) q(x)\, dx \le 0$. The monotonicity in m implies that $\int_X (x/[1 + mxEt]) q(x)\, dx < 0$ for all $m > -1$. Thus, $V(m + 1|m + 1) - V(m|m + 1) < 0$. Supermodularity implies that $V(m + 1|s) - V(m|s) < V(m + 1|m + 1) - V(m|m + 1) < 0$ whenever $s < m + 1$. We can conclude that there exists no solution $a \in (-1, 1)$ to the necessary arbitrage equation $V((a + 1)/2|a) - V((a - 1)/2|a) = 0$.

Analogously, if $\int_X (x/(1 + xEt)) q(x)\, dx \ge 0$ then $\int_X (x/[1 + (m + 1) xEt]) q(x)\, dx > 0$ for all $m < 0$, which in turn implies $V(m + 1|m) - V(m|m) > 0$ and therefore $V(m + 1|s) - V(m|s) > 0$ for all $s > m$. Again, there is no solution to the necessary arbitrage equation. $\qquad\square$

Proof of Proposition 6. Inserting $a = 0$ in (5.1) we find

$$\int_X \frac{x}{(2 - xEt)(2 + xEt)} q(x)\, dx = 0.$$

As the function $x/[4 - x^2(Et)^2]$ is anti-symmetric around $x = 0$, we conclude that the symmetric equilibrium exists when the distribution of x is symmetric around 0. $\qquad\square$

Proof of Proposition 7. At the ends of the interval, $a = -1$ and $a = 1$, the left-hand side of (5.1) is equal to $Ex/4$. The integral varies continuously with a, showing that the number of binary equilibria must be generically even. $\qquad\square$

Proof of Proposition 8. Consider any $\hat{x} < 0$. Since $\hat{x}/(1 - \hat{x}Et) < 0$, by continuity we have $\int_X (x/(1 - xEt)) q(x)\, dx < 0$ once q is sufficiently concentrated around $\hat{x}$. By Proposition 5 no binary equilibrium exists. A symmetric argument applies to any $\hat{x} > 0$. $\square$

Proof of Proposition 9. Since beliefs cannot eventually be concentrated on a false outcome, when $x \ne 0$ there exists some $\delta > 0$ such that with probability one the state interval $(-\delta, \delta)$ is assigned a mass less than $1 - \delta$ infinitely often.

Assume now that the history has resulted in a belief $q(x)$ according to which the state interval $[\delta, 1]$ has chance at least $\delta/2$. We show that then there is an $N \in \mathbb{N}$ such that if N credible experts in a row send the high message, all subsequent experts are unable to speak credibly. To see this, note that these high messages satisfy $m_i Et_i > \varepsilon^2$ by assumption, and so the state posterior will be FSD-higher than if the message ε of strength ε were observed N times. But, as $N \to \infty$, we have $(1 + \varepsilon^2 x)^N \to 0, 1, +\infty$ as respectively $x < 0, x = 0, x > 0$. Thus, eventually all weight is concentrated on states $x > 0$. In that limit, we have $\int_X (x/(1 + x)) q(x)\, dx > 0$ with this strict inequality holding already when N is sufficiently large. By Proposition 5, the subsequent expert is not credible and

leaves the belief unchanged. By induction, we conclude that no subsequent expert can be credible.

A symmetric argument applies to the case where the state interval $[-1, -\delta]$ has chance at least $\delta/2$. For N sufficiently large, after N low messages in a row credible communication is halted. With a slight abuse of notation, let N be so large as to work in both this and the previous paragraph.

We conclude that infinitely often there exists a continuation history of N identical messages leading to a termination of communication. Moreover, the chance of such termination is bounded below because the required messages (high or low) are issued for private signals in a range of at least 2ε. With a lower bound on the infinitely recurring termination chance, termination must occur almost surely in finite time. $\square$

References

[1] A.R. Admati, P. Pfleiderer, Broadcasting opinions with an overconfident sender, Int. Econ. Rev. 45 (2004), 467–498.

[2] A.V. Banerjee, A simple model of herd behavior, Quart. J. Econ. 107 (1992), 797–817.

[3] S. Bikhchandani, D. Hirshleifer, I. Welch, A theory of fads, fashion, custom, and cultural change as informational cascades, J. Pol. Econ. 100 (1992), 992–1026.

[4] A. Brandenburger, B. Polak, When managers cover their posteriors: Making the decisions the market wants to see, RAND J. Econ. 27 (1996), 523–541.

[5] C. Campbell, Learning and the market for information, Mimeo, Ohio State University, 1998.

[6] J. Chevalier, G. Ellison, Career concerns of mutual fund managers, Quart. J. Econ. 114 (1999), 389–432.

[7] D. Conway, Farlie-Gumbel-Morgenstern distributions, *in* "Encyclopedia of Statistical Sciences" (S. Kotz, N.L. Johnson, Eds.), pp. 28–31, Wiley-Interscience, New York, 1983.

[8] V. Crawford, J. Sobel, Strategic information transmission, Econometrica 50 (1982), 1431–1452.

[9] T. Ehrbeck, R. Waldmann, Why are professional forecasters biased? Agency versus behavioral explanations, Quart. J. Econ. 111 (1996), 21–40.

[10] C. Ewerhart, P.W. Schmitz, "Yes men," integrity, and the optimal design of incentive contracts, J. Econ. Behav. Organ. 43 (2000), 115–125.

[11] J. Graham, Herding among investment newsletters: Theory and evidence, J. Finance 54 (1999), 231–268.

[12] J.R. Green, N.L. Stokey, A two-person game of information transmission, Harvard Institute of Economic Research, Discussion Paper 751, 1980.

[13] B. Holmström, Managerial incentive problems: A dynamic perspective, Rev. Econ. Stud. 66 (1999), 169–182.

[14] B. Holmström, J. Ricart i Costa, Managerial incentives and capital management, Quart. J. Econ. 101 (1986), 835–860.

[15] H. Hong, J.D. Kubik, A. Solomon, Security analysts' career concerns and herding of earning forecasts, RAND J. Econ. 31 (2000), 121–144.

[16] O. Lamont, Macroeconomic forecasts and microeconomic forecasters, J. Econ. Behav. Organ. 48 (2002), 265–280.

[17] I.H. Lee, On the convergence of informational cascades, J. of Econ. Theory 61 (1993), 395–411.

[18] W. Li, Mind changes and the design of reporting protocols, Mimeo, University of California at Riverside, 2004.

[19] S. Lohmann, Information aggregation through costly political action, Amer. Econ. Rev. 84 (1994), 518–530.

[20] P. Milgrom, Good news and bad news: Representation theorems and applications, Bell J. Econ. 12 (1981), 380–391.

[21] W. Olszewski, Informal communication, J. Econ. Theory 117 (2004), 180-200.

[22] M. Ottaviani, P. Sørensen, Herd behavior and investment: Comment, Amer. Econ. Rev. 99 (2000), 695–704.

[23] M. Ottaviani, P. Sørensen, The strategy of professional forecasting, Mimeo, London Business School, 2004.

[24] M. Ottaviani, P. Sørensen, Reputational cheap talk, Mimeo, London Business School, 2004.

[25] M. Piccione, G. Tan, A simple model of expert and non-expert bidding in first-price auctions, J. Econ. Theory 70 (1996), 501–515.

[26] A. Prat, The wrong kind of transparency, Mimeo, London School of Economics, 2003.

[27] C. Prendergast, A theory of "yes men," Amer. Econ. Rev. 83 (1993), 757–770.

[28] C. Prendergast, L. Stole, Impetuous youngsters and jaded oldtimers: Acquiring a reputation for learning, J. Pol. Econ. 104 (1996), 1105–1134.

[29] D. Scharfstein, J. Stein, Herd behavior and investment, Amer. Econ. Rev. 80 (1990), 465–479.

[30] D. Scharfstein, J. Stein, Herd behavior and investment: Reply, Amer. Econ. Rev. 90 (2000), 705–706.

[31] D. Seidmann, Effective cheap talk with conflicting interests, J. Econ. Theory 50 (1990), 445–458.

[32] L. Smith, P. Sørensen, Pathological outcomes of observational learning, Econometrica 68 (2000), 371–398.

[33] J. Sobel, A theory of credibility, Rev. Econ. Stud. 52 (1985), 557–573.

[34] B. Trueman, Analyst forecasts and herding behavior, Rev. Finan. Stud. 7 (1994), 97–124.

[35] J. Watson, Information transmission when the informed party is confused, Games Econ. Behav. 12 (1996), 143–161.

[36] I. Welch, Herding among security analysts, J. Finan. Econ. 58 (2000), 369–396.

[37] E. Zitzewitz, Measuring herding and exaggeration by equity analysts, Mimeo, Stanford GSB, 2001.