Economics Letters 1 (2002) 000–000

economics letters

www.elsevier.com/locate/econbase

Prevention of herding by experts

Vladislav Kargin*

Cornerstone Research, 599 Lexington Avenue, New York, NY 10022, USA

Received 27 August 2001; received in revised form 23 May 2002; accepted 5 September 2002

Abstract

Experts tend to herd if they can communicate among themselves. I show that every expert can be induced to do independent research and report truthful results if and only if payments to the expert depend on other experts' reports. If the experts are risk-averse then the prevention of herding is costly.

Keywords: Herding; Information cascades; Communication; Multiple agents

JEL classification: D82

1. Introduction

Consider the following situation. For predicting market returns, a financial company uses the services of several experts who can communicate and share their results. The first problem is that an expert may shirk responsibilities if the client does not offer sufficient incentives. The second problem arises when these incentives are offered because in these cases the expert may choose to bias his own results towards results of the other experts. The difficulty of contract design in this situation is that inducing research may simultaneously create stimuli for misrepresenting the results obtained.

For example, reports of financial analysts who track a company for a corporate client are often biased by prior reports.[1] It is also well known that the analysts are paid on the basis of their relative performance. *Institutional Investor* publishes a ranking of the analysts and they get paid according to their position in the ranking. Do relative performance payments cause herding, or do they help reduce herding? And, do analysts enjoy any rent from the possibility of herding?

This paper derives two results about contracting in this situation. (1) It proves that the client can

*Tel.: +1-212-605-5018; fax: +1-212-759-3045.
E-mail address: slavakargin@yahoo.com (V. Kargin).

[1]See Trueman (1994), Graham (1999), and Welch (2000).

0165-1765/02/$ – see front matter
PII: S0165-1765(02)00252-5

induce research and truthful reporting by all experts only if she can make the payment to an expert a function of all experts' reports. (2) The expected payment to an expert is greater if the client uses the herding prevention scheme, if we compare it to a case in which the experts cannot communicate. These results suggest that relative performance payments are a useful but costly tool in reducing herding.

Several studies examined similar situations where attempts to solve problems of asymmetric information creates new incentive problems. Eichberger et al. (1999), Bergemann and Valimaki (2000), and Osband (1989) are all recent examples. These studies, however, are not concerned with the problem of herding, which is central for my work. On the other hand, there is a large literature about herding in finance and economics applications (e.g. Banerjee, 1992; Bikhchandani et al., 1992; Devenow and Welch, 1996) but it has not asked the question of herding prevention. Maskin and Riley (1984) and Cremer and McLean (1988) are also related studies. My model is, however, about the simultaneous inducement of research and truthful reporting, which is different from the focus of these papers.

The remainder is organized as follows. Section 2 describes the model, Section 3 formulates auxiliary results for the case with only one expert, Section 4 presents the main results, and Section 5 concludes.

2. Model

- Timing: the first one of a finite sequence of experts decides whether to conduct the research or not. Then he makes the alleged results of the research available to all other experts, but not to the client. Next, the second expert decides whether to conduct research or not, and what to report, and so on. The order of research is exogenous and not known to the client. This assumption is meant to capture the real-life difficulties in resolving priority questions. Next, all the experts simultaneously submit their reports to the client. The results in the reports are not necessarily the same as the results that were made available to other experts, although it will be shown that they are in the equilibrium. Next, the client takes an action based on the reported research results; then the true state is revealed. Finally, the experts are paid.
- Preferences: the preferences of the client are assumed lexicographic in that her main goal is to insure that the maximal amount of research is undertaken and that truthful results are reported to her. Given that this goal is assured, the client wants to minimize the expected amount of money to be paid to the experts. This simplifying assumption is taken to isolate the problem of herding prevention from the trade-off between costs of implementing this goal and benefits of knowing truthful results. In reality, high costs can force the client to abandon the goal of herding prevention. This possibility is left outside of the model.

  Each expert has a vNM utility function of the form $-\gamma + u(t)$, where t is the payment to the expert, γ is $c > 0$ if an expert does research and 0 otherwise, and $u(\cdot)$ is an increasing and concave function. For simplicity, I assume that experts always resolve indifferences in favor of the client.
- Research: there is a finite number of the states of the world, and every expert who undertakes research gets a signal from a finite set. The signals are i.i.d. conditional on the state of the world. The prior probability of state i is $\pi_i > 0$. The probability of obtaining signal s conditional on the realization of state i is $p(s|i)$, the joint probability of signal s and state i is $p(s, i) = p(s|i)\pi_i$, and the

conditional probability of state i given signal s is $p(i|s)$. Similar notation is used for the joint probabilities of several signals. For example, $p(s_1, s_2|i)$ is the conditional probability that the first expert gets signal s_1, and the second expert gets signal s_2 given state i. Likelihood ratio of states i and j given signal s is $l_{ij}(s) \equiv p(s|i)/p(s|j)$. States are identifiable by signals, that is for any i and j there are such s and s' that $l_{ij}(s) \neq l_{ij}(s')$.

- Communication: an expert makes a report available to all other experts. I assume that if the experts are indifferent then they will report truthful results.
- Contracting: the client can pay an expert depending on the realization of the state of the world and on the set of all experts' reports. The function $t(s, S, i)$ represents the payment to an expert if s is his own report and if S is the set of the other experts' reports. The model assumes limited liability of an expert: $t(s, S, i) \geq 0$. In the case of only one expert, I abuse notation and denote the payment as $t(s, i)$.

3. Results for one expert

This section considers the case with only one expert. There is no communication in this case but I anticipate its appearance in the next section by focusing on the dependence of optimal schemes on the expert's prior. Indeed, the main role of communication is to change the experts' priors.

Doing research and reporting truthfully requires that the following set of conditions be satisfied: for each pair of true signal s and reported signal $\hat{s}$

$$\sum_i p(i|s)u(t(s, i)) \geq \sum_i p(i|s)u(t(\hat{s}, i)) \tag{3.1}$$

$$\sum_s \sum_i p(i, s)u(t(s, i)) \geq \sum_i \pi_i u(t(\hat{s}, i)) + c. \tag{3.2}$$

The first equation is the truthtelling condition, the second guarantees doing research.

Proposition 1. *The client can always induce the expert to do the research and report his results truthfully.*

Outline of proof. We take two states (0 and 1) and order the signals so that a measure of the signals' information supporting state 1 against state 0 declines. Then we construct a payment scheme with the following properties: $t(s, 1)$ is decreasing and $t(s, 0)$ is increasing. It is easy to fine-tune the payments so that the experts with the true signal s will never want to switch to a neighboring signal s'. The crucial point is that this local property implies the global property: the expert will never want to switch to *any* false signal. As a result, the truthtelling and research can be strictly encouraged. An appropriate scaling can amplify this incentive so that research is profitable for the expert. The size of the scaling, of course, will depend on the cost of research and on the informational structure of the signals. The general case can be easily reduced to the case with only two states.

Proof. If payments $t(s, i)$ satisfy (3.1) and (3.2) for prior π, then payments

$$t'(s,i) =: u^{-1}\left(\frac{\pi_i}{\pi'_i} u(t(s,i))\right) \tag{3.3}$$

will satisfy (3.1) and (3.2) for prior π'. Thus, without loss of generality, we can assume a uniform prior on the states. Firstly, consider the case where the number of states is 2. Let

$$\begin{aligned} x_{ss'} &= u(t(s,0)) - u(t(s',0)), \\ z_{ss'} &= u(t(s,1)) - u(t(s',1)). \end{aligned} \tag{3.4}$$

The problem is to define $x_{ss'}$ and $z_{ss'}$ in such a way that they are consistent, that they satisfy truthtelling constraints[2] for any s, s':

$$x_{ss'} + l_{10}(s) z_{ss'} \geq 0, \tag{3.5}$$

and that for each s' there is at least one s such that the constraint (3.5) is strict. This third condition ensures that knowledge of true signal provides some positive benefit to the expert. By scaling the payments appropriately, the client can make this expected benefit greater than the cost of research.

Without loss of generality, assume that

$$l_{10}(s_1) \geq l_{10}(s_2) \geq \cdots \geq l_{10}(s_n). \tag{3.6}$$

Define

$$z_{s_i s_j} = j - i, \forall i, j, \tag{3.7}$$

choose $x_{s_i, s_{i+1}}$ in a way such that

$$-l_{10}(s_{i+1}) \geq x_{s_i, s_{i+1}} \geq -l_{10}(s_i), \tag{3.8}$$

and define $x_{s_i s_j}$ (with $j > i+1$) by

$$x_{s_i s_j} = x_{s_i s_{i+1}} + x_{s_{i+1}, s_{i+2}} + \cdots + x_{s_{j-1}, s_j}. \tag{3.9}$$

It is easy to check that the conditions are satisfied. Since the client can arbitrarily choose two states of the world and restrict herself to schemes that pay only if one of the two states is realized, she can induce doing research in the general finite-state case. □.

The minimal expected payment to the expert depends on his prior. This dependence is convex as the following proposition shows.

Proposition 2. *For the case of a risk-averse expert, the minimal expected payment* $R(\pi)$ *is convex.*

Proof. Let $x_i(s) = u(t(s,i))\pi_i$. Then the client's expected payment is

---

[2]Indeed, the truthtelling constraint is $p(s,0)u(t(s,0)) + p(s,1)u(t(s,1)) \geq p(s,0)u(t(s',0)) + p(s,1)u(t(s',1))$, which is the same as $p(s|0)\pi_0 x_{ss'} + p(s|1)\pi_1 z_{ss'} \geq 0$, which is equivalent to (3.5) because of the uniform prior assumption.

$$f(x, \pi) = \sum_{s} \sum_{i} p(s|i) u^{-1}\left(\frac{x_i(s)}{\pi_i}\right) \pi_i. \tag{3.10}$$

The client chooses $\{x_i(s)\}_{i,s}$ by solving

$$R(\pi) = \inf_{x_i(s)} f(x,\pi) \tag{3.11}$$

subject to linear constraints on $x_i(s)$. The conclusion of the proposition follows from the following lemma and from elementary calculations that show that $f(x, \pi)$ is convex in both arguments.

Lemma 1. *Suppose $f(x, \pi)$ is convex in both arguments, X is a convex set, and*

$$F(\pi) = \inf_{x \in X} f(x, \pi).$$

Then $F(\pi)$ is convex.

The proof of Lemma 1 can be found in Boyd and Vandenberghe (2003; Section 2.2.5).

4. Results for multiple experts

This section applies the results of the previous section to the situation with multiple experts and communication. Firstly, the following impossibility result holds.

Theorem 1. *If the client can only pay each expert according to his own report and the realized state of the world, she cannot guarantee that all experts always conduct research and report truthfully.*

Proof. Suppose instead that such a scheme exists, and the second expert reports the results of his own research independently of what was reported by the first expert. The second expert's ex-ante expected payoff is

$$\sum_{s_1,s,i} p(s_1, s, i) t(s, i) - c = \sum_{s,i} p(s, i) t(s, i) - c, \tag{4.1}$$

where s_1 is the signal of the first expert. The ex-ante expected payoff from the strategy of doing no research and repeating the results of the first expert is

$$\sum_{s_1,i} p(s_1, i) t(s_1, i) = \sum_{s,i} p(s, i) t(s, i), \tag{4.2}$$

which is higher. This is a contradiction with incentive compatibility. □

Of course, this theorem does not imply that the optimal strategy of the second expert is to mimic the first expert independently of the result of the first expert. The claim is only that the client cannot avoid cases where the second expert finds this imitation beneficial.

The conclusion is reversed if the client can condition the payment on the reports of the other experts. Then the following theorem holds.

Theorem 2. *If the payments to each expert can depend on the reports of the other experts, then the client can design a scheme such that doing research and reporting the results truthfully is a Nash equilibrium for the experts.*

Proof. Suppose that the client knows the order of research and an expert is the last in the sequence. From the results of the one-expert case, we know that there is such a scheme that

$$E(u(t(s, S, i))|S, s) - E(u(t(\hat{s}, S, i))|S, s) \geq 0 \tag{4.3}$$

for every S, s, $\hat{s}$, where S is the set of reports by the previous experts and $E(\cdot|S, s)$ is the expectation over the distribution of states conditioned on the realization of signals $\{S, s\}$. Consider a scheme where every expert is paid as if he were the last in the sequence of the experts. We need to check only the incentives of an expert who is not the last in the sequence. Let S^- and S^+ denote the sets of reports before and after the expert has made his own report. His expected utility is then $E(u(t(\hat{s}, S, i))|S^-, s)$. This is an expectation of $E(u(t(\hat{s}, S, i))|\{S^-, S^+\}, s)$ over the distribution of S^+. By taking the expectation of inequality (4.3) over S^+ it follows then that

$$E(u(t(s, S, i))|S^-, s) - E(u(t(\hat{s}, S, i))|S^-, s) \geq 0 \tag{4.4}$$

for every S^-, s, $\hat{s}$; and, given S^-, $\hat{s}$, inequality (4.4) is strict for some choice of s. It follows that this scheme guarantees that doing research and reporting truthfully to the client is an equilibrium. □

The intuition behind the result is clear: under the proposed scheme an expert knows that he would do research and tell the truth if all information about other experts' results were available to him. So he conducts research and reports truthfully, even though he does not know all reports at the moment when he is obliged to make his decisions.

Note that this payment scheme is a sophisticated relative performance scheme, in which an expert's payment depends on his incremental contribution to the final posterior belief. This is the reason why an expert prefers not to herd: even if he gets only a weak or unprobable signal the contribution of this signal to the final posterior is positive while the incremental contribution of a repeated signal is zero.

Concerning the question of the cost of communication for the client, the following result holds.

Theorem 3. *If the experts are risk-averse, the minimal expected payment is greater in the situation with communicating experts and herding-preventing schemes comparing to the case when the experts cannot communicate.*

Proof. The expected payment to an expert without communication is $R(\pi)$. In the case with communication, the ex-ante expected payment $\Sigma p(S)R(\pi_S)$ where $\pi_S = p(\cdot|S)$ is the posterior of the expert after the set of other experts' signals S is communicated to him and $p(S)$ is the ex-ante probability of this set of signals. $R(\cdot)$ is a convex function from Proposition 2 and $\pi = \Sigma p(s)\pi_s$. It follows that $R(\pi) < \Sigma p(s)R(\pi_s)$. □

Thus, the client is always expected to pay more if she is going to prevent herding among experts.

5. Conclusion

I have shown that the client cannot induce both research and truthful reporting by all experts if the payments only depend on the expert's own report and the realized state of the world. The client can always achieve this goal, however, if the payments depend on the reports of other experts. If experts are risk-averse, then communication increases the cost of research for the client.

Acknowledgements

I dedicate this paper to the memory of Robert W. Rosenthal. I am grateful to Dilip Mookherjee and Zvika Neeman for helpful discussions, and to anonymous referees for their comments.

References

Banerjee, A.V., 1992. A simple model of herd behavior. Quarterly Journal of Economics 107 (3), 797–817.

Bergemann, D., Valimaki, J., 2000. Information Acquisition and Efficient Mechanism Design, Working paper.

Bikhchandani, S., Hirshleifer, D., Welch, I., 1992. A theory of fads, fashion, custom, and cultural change in informational cascades. Journal of Political Economy 100 (5), 992–1026.

Boyd, S., Vandenberghe, L., 2003. Convex optimization. Forthcoming. Internet version at http://www.stanford.edu/class/ee364/.

Cremer, J., McLean, R.P., 1988. Full extraction of the surplus in bayesian and dominant strategy auctions. Econometrica 56 (6), 1247–1257.

Devenow, A., Welch, I., 1996. Rational herding in financial economics. European Economic Review 40, 603–615.

Eichberger, J., Grant, S., King, S.P., 1999. On relative performance contracts and fund managers incentives. European Economic Review 43, 135–161.

Graham, J.R., 1999. Herding among investment newsletters: theory and evidence. Journal of Finance 54 (1), 237–268.

Maskin, E., Riley, J., 1984. Monopoly with incomplete information. Rand Journal of Economics 15 (2), 171–196.

Osband, K., 1989. Optimal forecasting incentives. Journal of Political Economy 97 (5), 1091–1112.

Trueman, B., 1994. Analysts forecasts and herding behavior. Review of Financial Studies 7 (1), 97–124.

Welch, I., 2000. Herding among security analysts. Journal of Financial Economics 58, 369–396.