Informational externalities, herding, and incentives

Lluís Bru

Universidad de Málaga

and*

Xavier Vives

INSEAD and Institut d'Anàlisi Econòmica, CSIC

* We are grateful to Meg Meyer for very helpful conversations on the topic of this paper. Bru acknowledges the support of DGICYT project PB98-1402 and Vives of PB98-0696.

Abstract

A version of the herding prediction model with a rational expectations flavor is reexamined at the light of incentive theory. The welfare loss at the market solution with respect to the incentive efficient solution can be decomposed into an information externality term minus an incentive cost term. It is found that the inefficiency of herding at the market solution is low when the cost of providing incentives is high. When the cost of providing incentives is low (and this happens when prior information is diffuse) the incentive efficient solution approaches the team solution that fully internalizes the information externality. Then the herding problem at the market solution is at its worst.

Keywords: information aggregation, rational expectations, coordination, teams, mechanism design.

1. Introduction

This paper shows that the inefficiency associated to herding is moderated when incentive provision is taken into account. Herding inefficiencies are maximal when the cost of providing incentives is low. In contrast, when the cost of providing incentives is high "market" solutions involving herding are close to constrained optimal.

Herding or, more in general, the insufficient reliance of agents on their private information, has been put forward recently as an explanation for different phenomena like financial crises, fashion and technology adoption. The herding literature has put the finger on the welfare consequences of information externalities in a very stark statistical prediction model. The root of inefficiency in herding models is an informational externality not taken into account by agents when making decisions. Attention has been focused typically on the full (shared) information equilibrium as a benchmark for comparison. However, this may not be attainable. The inefficiency of herding should be compared with what is achievable once incentive considerations are taken into account. It is not enough to say that a herding outcome is inefficient if it is not compared with what could be achieved.

Herding models (Banerjee (1992), Bikhchandani et al (1992)) have typically a lot of structure and the results do depend on assumptions like a discrete action space or signals of bounded precision (see Lee(1993) and Smith and Sorensen (2000)).[1] However, the informational externality problem is general and it is not qualitatively different from the one present in rational expectations models (Vives (1993, 1996), Gale (1996)). A central issue is therefore to characterize the inefficiency associated to informational externalities. In this paper we follow the herding literature in considering a pure informational externality model. We replace the sequential prediction model used in the

[1] See Chamley and Gale (1994) and Gul and Lundholm (1995) for models with endogenous order of moves.

literature by a static prediction model with a (noisy) competitive rational expectations flavor. Each agent out of a continuum has to predict a random variable with the help of a private signal and the (noisy) average of the predictions of other agents. This allows us to study in the simplest possible way the informational externality arising out of the decisions of agents. The unique linear (RE) equilibrium is characterized. At the RE equilibrium agents make an efficient use of public information and a privately optimal use of private information.

The information externality arises because an agent does not take into account the effect of his decision on the informativeness of the average prediction. The coordination problem underlying the information externality is solved at the team solution where each agent acts to minimize the average prediction error using a decentralized strategy. At the team solution agents make an efficient use of public information but depart from the privately optimal use of private information. The distance between the team optimum and the market solution is large when the information externality is important and this happens for intermediate values of the noise in the public and private signals. This is consistent also with the analysis in Vives (1997) where a dynamic version of the model is explored.

The team solution is only a theoretical benchmark because it does not take into account incentives. To account for incentives we take a mechanism design approach. In this way we can find the efficient solution in the class of incentive compatible ones. We restrict the mechanism to have a similar structure as the market solution: to be linear and to use the same communication constraints. That is, the mechanism should use the same information aggregator as the market: a noisy average of the predictions of the agents in the economy. We do not allow transfer payments. This defines the class of linear Bayesian incentive compatible mechanisms. The approach taken is similar to the welfare analysis of rational expectations equilibria in Laffont (1985).

We characterize the incentive-efficient solution and compare it to the market and team solutions. At the incentive-efficient solution agents make a privately optimal use of private information (to preserve incentive compatibility) but depart from the efficient use of public information. Indeed, to provide incentives is costly.

It is found that there is herding at the market solution. Agents put too little weight on their private information as compared to the team solution. The incentive-efficient solution lies in between the market and the team solutions. The reason is that to provide incentives is costly because public information is not used efficiently. The incentive and the team solutions are close together when to provide incentives is cheap. This happens when there is a lot of (prior) uncertainty. In the limit with a diffuse prior the team and the incentive-efficient solutions coincide. With a diffuse prior it is cheap to provide incentives because then agents are very responsive to private signals. In this case the herding problem at the market solution is at its worst. Conversely, when it is costly to provide incentives then the market solution is close to the incentive-efficient solution. We see therefore that the cost of incentive provision tends to make herding not so inefficient.

It is worth noting that at the incentive-efficient solution the loss may be increasing in the precision of the prior. The reason is that a more diffuse prior diminishes the cost of providing incentives (relaxes the incentive compatibility constraint). This means that there are circumstances in which a planner would like to add noise to the prior.

Simulations show that the welfare loss at the market solution, relative to incentive-efficiency, is decreasing in the precision of the prior, increasing in the precision of public information and hump-shaped shaped in the precision of private signals. Relative welfare losses can be substantial for a low precision of the prior, high precision of public information and intermediate precision of private signals. However, for many parameter combinations the relative loss remains very moderate. The relative cost of providing incentives is increasing in the precision of the prior and decreasing in the precision of public information and in the precision of private signals.

The plan of the paper is as follows. Section 2 presents a prediction model with a rational expectations flavor. Section 3 deals with the team optimal solution and Section 4 with the incentive-efficient solution. Section 5 comments briefly the possibility of introducing transfers with a tax-subsidy scheme. Concluding remarks follow.

2. A model with a rational expectations flavor

Consider a continuum of agents indexed in the unit interval [0,1] (endowed with the Lebesgue measure) trying to predict a random variable θ, normally distributed with mean $\bar{\theta}$ and finite variance σ_θ^2 ($\theta \sim N(\bar{\theta}, \sigma_\theta^2)$) . The payoff to an agent when choosing an action q_i is $-(\theta - q_i)^2$.

Agent i receives at the start a *private signal* $s_i = \theta + \varepsilon_i$, where the errors terms are iid with $\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$ and cov $(\theta, \varepsilon_i) = 0$. The convention is made that errors on average cancel out: $\int_0^1 \varepsilon_i \, di = 0$ almost surely (a.s.).[2] In consequence, the average signal reveals θ a.s. Apart from their private information agents observe also a *public information* signal p, where p is a noisy version of the average action of the agents: $p = \int_0^1 q_i \, di + u$ with $u \sim N(0, \sigma_u^2)$ and cov $(u, \varepsilon_i) =$ cov $(u, \theta) = 0$.

Agent i has available the information vector $I_i = \{s_i, p\}$ and solves the problem:

---

2 See Feldman and Gilles (1985) for the measure-theoretic issues involved. Suppose that $(q_i)_{i \in [0,1]}$ is a process of independent random variables with $Eq_i = 0$ for all i and that variances (Var q_i) are uniformly bounded. We define $\int_0^1 q_i \, di = 0$ almost surely (a.s.). This convention will be used taking as given the usual linearity properties of integrals. For example, if signals have uniformly bounded variances, we will write $s = \int_0^1 s_i \, di = \int_0^1 (\theta + \varepsilon_i) \, di = \theta + \int_0^1 \varepsilon_i \, di = \theta$ (a.s.), using the linearity of the integral (which is being defined) and the convention (which implies that $\int_0^1 \varepsilon_i \, di = 0$).

$\text{Min}_q\, E\{(\theta - q)^2 \mid I_i\}$.

As is well known the solution to this problem is $q_i = E(\theta|I_i)$. This information structure corresponds to a rational expectations solution. Indeed, think of agent i submitting a schedule contingent on the realizations of p, $q_i (s_i, \cdot)$, where p solves the equation $p = \int_0^1 q_i (s_i, p)\, di + u$. The forecast of agent i is contingent on his private signal and the (noisy) average forecast.

The minimization of the square loss function may arise from agents having quadratic utility functions. Suppose agent i has a utility function given by $U_i = (\theta + \eta_i)\, q_i - \frac{1}{2}\, q_i^2$ where η_i is an idiosyncratic random term (uncorrelated with everything else). Agent i may be a retailer selling in market and facing a random price $\theta + \eta_i$. The quantity sold q_i is supplied by a producer, the retail cost is $\frac{1}{2}\, q_i^2$. The retailer submits a request schedule for supplies $q_i (s_i, p)$, indexed by his private information s_i, contingent on the aggregate requests p. Some retailers just send noisy requests, u in the aggregate. We have then that the expected welfare loss with respect to the full-information first best (where θ is known and $q_i = \theta$) is easily seen to be $E(\theta - q_i)^2 /2$.

Other examples include the following. The agent can be a firm making a decision about capacity, with $\theta + \eta_i$ indexing its marginal value and quadratic costs A firm decides about capacity based on its private information and the aggregate capacity choices in the industry which includes some firms that invest for non-informational exogenous reasons. Alternatively, competitive firms decide about investment with macroeconomic uncertainty represented by the random variable θ which determines average profitability. Firms invest taking into account that the profits of investment will depend on the realization of θ. To predict θ each firm has access to a private signal as well as to public information, which is formed by aggregate investment figures compiled by a government agency. Data on aggregate investment incorporates measurement error.[3] Still the agent could be a consumer facing a good of random quality and $\theta + \eta_i$ is his willingness to pay. Assume that firms produce at zero cost and that prices are fixed at marginal cost. The consumer makes his quantity decision based on his private information and contingent on the realized value of aggregate sales $p = \int_0^1 q_i \, di + u$, where u are purchases by made by "noise" consumers.

We concentrate attention on linear REE. Given the structure of the model REE will necessarily be symmetric. Let a be the coefficient of s_i in the candidate linear equilibrium strategy of agent i. Then from the normality assumption and $p = \int_0^1 q_i \, di + u$ it follows that p will be a linear transformation of $z \equiv a\theta + u$ and that $E\{\theta | p\} = E\{\theta | z\}$.[4]

Let $\theta^* = E\{\theta | z\}$. We can write the equilibrium strategy as $q(s_i, z)$ (= $E\{\theta | s_i, z\}$) = $a s_i + (1-a)\, \theta^*$ where $a = \tau_\varepsilon / (\tau_\varepsilon + \tau)$ and $\tau = \tau_\theta + \tau_u a^2$.[5] The precision $\tau \equiv (Var(\theta | p))^{-1}$ is the *informativeness of public information* in the estimation of θ. The posterior mean of θ is a weighted average of the signals of the agent with weights according to their precisions (the private signal with precision τ_ε and the public with precision τ).

Proposition 1 characterizes the equilibrium. Denote by a^m the weight to private information in the REE or market solution. Let $v^m \equiv E(\theta - E\{\theta | s_i, z\})^2$.

---

3 For example, quarterly data on national accounts are subject to measurement error.

4 Indeed, let $q_i = a_i s_i + h_i p + b_i$. Then $p = a\theta + \int_0^1 a_i \varepsilon_i \, di + h p + b + u$. If $h \neq 1$ then $p = (a\theta + u + b)/(1-h)$, where $a = \int_0^1 a_i \, di$. This follows from the fact that $\int_0^1 s_i \, di = a\theta + \int_0^1 a_i \varepsilon_i \, di = a\theta$ because, according to our convention, $\int_0^1 a_i \varepsilon_i = 0$ provided the a_i's are uniformly bounded.

5 It is immediate from the properties of normal distributions that $E\{\theta | s_i, z\} = (\tau_\varepsilon s_i + \tau_u a z + \tau_\theta \bar{\theta}) / (\tau_\varepsilon + \tau)$ and $\tau = \tau_\theta + \tau_u a^2$. Equivalently, $E\{\theta | s_i, z\} = a s_i + (1-a) E\{\theta | z\}$ where $a = \tau_\varepsilon / (\tau_\varepsilon + \tau)$ because $E\{\theta | z\} = (\tau_u a z + \tau_\theta \bar{\theta}) / \tau$.

Proposition 1. The equilibrium strategy at the LREE is given by $q_i = a^m s_i + (1-a^m)\,\theta^*$ where a^m is the unique positive real solution to the cubic equation $a = \tau_\varepsilon / (\tau_\varepsilon + \tau_\theta + \tau_u a^2)$. We have that $0 < a^m < \alpha \equiv \tau_\varepsilon / (\tau_\varepsilon + \tau_\theta)$ and $v^m = (\tau_\varepsilon + \tau^m)^{-1} = a^m/\tau_\varepsilon$. Furthermore, a^m and v^m decrease with τ_θ and τ_u; a^m increases and v^m decreases with τ_ε.

Remark: The REE makes (privately) efficient use of private and public information. That is, at the REE, $Cov((\theta - q_i), s_i) = Cov((\theta - q_i), \theta^*) = 0$ for all s_i and θ^* (equivalently, $E((\theta - q_i), s_i)) = E((\theta - q_i), \theta^*) = 0$ for all s_i and θ^*). (See Figure 1.) The result follows from the equilibrium condition $E((\theta - q_i)|s_i, \theta^*) = 0$ for all s_i and θ^*, using the projection theorem for normal random variables (according to which $Cov\{(\theta - q_i) - E((\theta - q_i)|s_i, \theta^*), E((\theta - q_i)|s_i, \theta^*)\} = 0$).

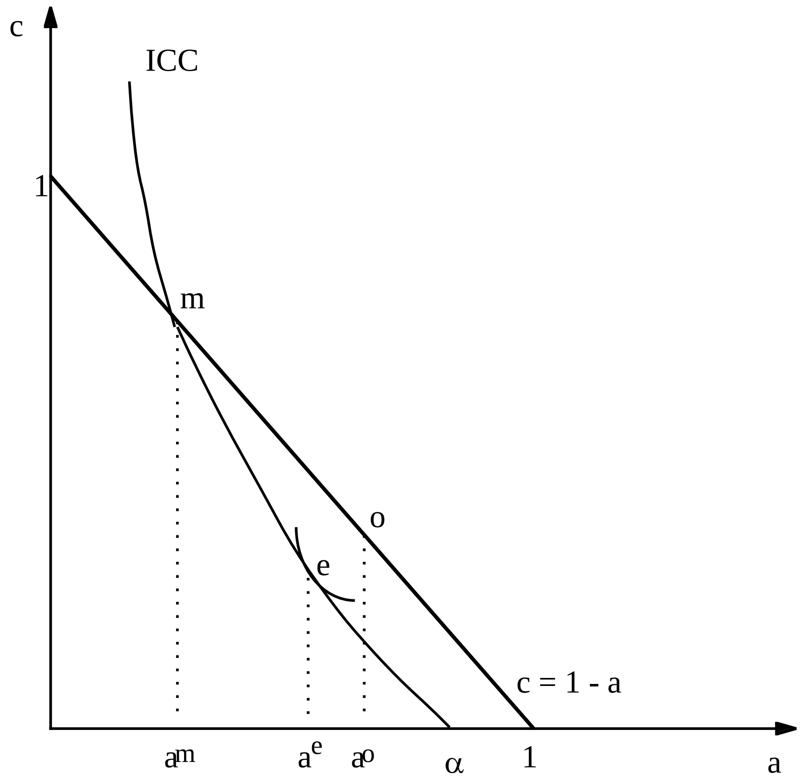

Figure 1

3. The team optimal solution

We will take as decentralized welfare benchmark the *team solution* where the planner is able to impose decision rules to agents but cannot manipulate the information flows. That is, agent i will continue to have available information $I_i = \{s_i, p\}$. In the retailer example it is as if the producer could impose the schedules of supplies requested by the retailers.

The planner will internalize the information externality derived from the decisions of agents who do not take into account the benefits they generate to other agents when they respond to their private information. It is as if all agents had a common objective but attempted to optimize it using their private information.

Furthermore, the planner will be restricted to use linear rules, the same family of simple rules that agents in the market use. Given decision rules $q_i(I_i)$ the average expected loss is then $\int_0^1 E(\theta - q_i(I_i))^2 \, di$. Let, without loss of generality, $q_i(I_i) = a_i s_i + c_i \theta^*$. Then, contingent on a_i, it is immediate that to minimize $E(\theta - q_i(I_i))^2$ with a linear function $q_i(\cdot)$ one has to set $c_i = 1 - a_i$. Indeed, otherwise public information would not be exploited efficiently. Put it another way, at the team optimal solution it has to hold that public information is used efficiently, $E((\theta - q_i), \theta^*) = 0$, and this is equivalent to $c_i = 1 - a_i$.

This means that the loss associated to agent i is given by

$L(a_i, \tau) = E(\theta - q_i(I_i))^2 = \frac{(1-a_i)^2}{\tau} + \frac{a_i^2}{\tau_\varepsilon}$ where $\tau = \tau_\theta + \tau_u \; a^2$ and $a = \int_0^1 a_i \, di$.[6] Given that L is convex in a_i it is immediate that $\int_0^1 L(a_i, \tau) \, di \;\; L(\int_0^1 a_i \, di, \tau)$ and therefore there can not be any gain from using asymmetric rules.

---

[6] Similarly as before p can be seen to be a linear transformation of $z = a\theta + u$ where $a = \int_0^1 a_i \, di$. This follows from the fact that $\int_0^1 s_i \, di = a\,\theta + \int_0^1 a_i \varepsilon_i \, di = a\,\theta$ because, according to our convention, $\int_0^1 a_i \varepsilon_i = 0$ provided the a_i's are uniformly bounded. Therefore, $\tau = \tau_\theta + \tau_u a^2$.

Let us restrict attention therefore to symmetric rules: $q(I_i) = a s_i + (1-a)\theta^*$. The associated loss is then $L(a) = \frac{(1-a)^2}{\tau_\theta + \tau_u a^2} + \frac{a^2}{\tau_\varepsilon}$. This function is strictly quasiconvex in a. Furthermore, $L' = -2\tau^{-1}(1-a(\tau_\varepsilon + \tau)(\tau_\varepsilon)^{-1} + (1-a)^2 \tau_u \tau^{-1}) = 2(a(\tau_\varepsilon)^{-1} - (1-a)(\tau_\theta + \tau_u a)\tau^{-2}$, where $\tau = \tau_\theta + \tau_u a^2$. It follows that at the market solution (for which $a^m = \tau_\varepsilon / (\tau_\varepsilon + \tau^m)$) $L'(a^m) < 0$ and $L'(1) > 0$. Denote by a^o the (unique) team solution and let $v^o = L(a^o)$. The following proposition characterizes the team solution.

<u>Proposition 2</u>. At the unique linear team solution $1 > a^o > a^m$ and a^o is increasing in τ_ε and decreasing in τ_θ. The equilibrium loss v^o is decreasing in τ_ε, τ_θ and τ_u.[7]

It follows that the weight to private (public) information is too low (high) at the market solution. The reason is that agents at the market solution do not internalize the positive effect on others, via the increase in the informativeness of the public statistic p, of their response to private information.

4. The cost of providing incentives. The incentive efficient solution.

We have characterized the team optimal solution disregarding incentive issues. That is, assuming that the agents will follow the recommended strategy of the center (or alternatively that agents internalize the common team objective). Is this incentive compatible? Can the market solution be improved upon when the recommended rules must be incentive compatible?

Let us consider direct revelation mechanisms which are of the same class as the market solution: linear of the form $q_i = a_i s_i + c_i \theta^* + b_i$. In this class of mechanisms the same communications constraints as in the market hold. That is, the action of an agent depends on his reported signal and a public statistic p which is formed in the same way as in the

7 Comparative statics follow from $^2L(a^o)/a\tau_\theta > 0$ and $^2L/a\tau_\varepsilon < 0$. It is immediate also that $L/\tau_u < 0$, $L/\tau_\varepsilon < 0$, and $L/\tau_\theta < 0$. However, sign$\{^2L/a\tau_u\}$ is ambiguous.

market solution. This defines the class of Linear Bayesian Incentive Compatible Mechanisms (LBICM). We are comparing the market solution with what is best in the class LBICM. We restrict therefore, as in Laffont (1985) for example, the search of efficient allocations to the class of mechanisms which are close to the properties of the market solution (linearity and reliance on noisy public information).

This is how the mechanism works. Agents are asked to reveal their signal and the center recommends privately action $q_i = a_i \hat{s}_i + c_i \hat{\theta}^* + b_i$ to agent i when he has reported $\hat{s}_i$, where $\hat{\theta}^* = E\{\theta \mid \hat{z}\}$ with $\hat{z} = \int_0^1 a_i \hat{s}_i \, di + u$. (If agents tell the truth $\hat{s}_i = s_i$ and $z = a\theta + u$ where $a = \int_0^1 a_i \, di$.)

In the retailer example the producer would ask each retailer for a report on his private information and then supply retailer i with $q_i = a_i \hat{s}_i + c_i \hat{\theta}^* + b_i$. It must be explained then why the producer can not aggregate the signals to obtain θ and instead use the aggregator θ^*. We may think, reasonably, that the information s_i can not be communicated and in fact an indirect mechanism must be used. Each retailer is asked to make a request for supplies. Retailer i submits request $\hat{q}_i$ and the center assigns the retailer a supply of q_i which is a linear function of $\hat{q}_i$ and the aggregate requests $\hat{p} = \int_0^1 \hat{q}_i \, di + u$.

Suppose other agents tell the truth. The problem of agent i is now to choose $\hat{s}_i$ to solve

Min $E\{(\theta - q_i)^2 \mid s_i\}$ where $q_i = a_i \hat{s}_i + c_i \theta^* + b_i$.

Provided that a_i is not zero this is a strictly convex problem with first order condition

$$-2a_i E\{(\theta - q_i) \mid s_i\} = 0$$

and therefore $E\{(\theta - q_i) \mid s_i\} = 0$ for all s_i. Truthful revelation requires then $\hat{s}_i = s_i$ to be optimal when others tell the truth. This means that $E\{(\theta - (a_i s_i + c_i \theta^* + b_i)) \mid s_i\} = 0$ has to hold for all s_i. From this equation we derive the incentive compatibility constraints (ICC): $E\{(\theta - q_i)\} = 0$ (by taking expectations) and $E\{(\theta - q_i), s_i\} = 0$ (from normality).

The constraint $E\{(\theta - q_i)\} = 0$ implies that $b_i = (1-a_i-c_i)\bar{\theta}$.[8] The constraint $E\{(\theta - q_i), s_i\} = 0$ yields $\alpha_i (1-c_i \tau_u a^2 \tau^{-1}) = a_i$.

In a symmetric mechanism, $c = (\tau_u a^2)^{-1} \tau (1-a \alpha^{-1})$. This is a decreasing function in a with c = 0 for a = α.

At the market solution we have that c =1-a. Therefore, the REE is determined by the intersection of the ICC, $E\{(\theta - q_i), s_i\} = 0$ or $c = (\tau_u a^2)^{-1} \tau (1-a \alpha^{-1})$, and the constraint determining the efficient exploitation of public information, $E((\theta - q_i), \theta^*)) = 0$ or line c =1-a. (See Figure 1.)

In order to provide truth-telling incentives if it is desired that agents put a weight to private information larger than the market a^m, then the decision rule must give a lower weight to public information (lower c) and this weight must be less than 1-a (otherwise we would go back to the REE). This means that a distortion is introduced because public information is not used efficiently, c < 1- a, and a positive weight, 1-a -c > 0, must be given to the prior mean $\bar{\theta}$. It is worth noting that the distortion diminishes as τ_θ decreases. Indeed, as τ_θ tends to 0 we have that the ICC, $E\{(\theta - q_i), s_i\} = 0$ or $c = (\tau_u a^2)^{-1} \tau (1-a \alpha^{-1})$, and the efficient use of public information line, c =1-a, coincide.

<u>Remark</u>: Indeed, the REE fulfils the incentive compatibility constraints. In fact the REE can be obtained as the solution to the following problem:

Min_{a_i} $E(\theta - q_i)^2$ subject to $q_i = a_i s_i + c_i \theta^* + b_i$, $\theta^* = E\{\theta | z\}$ with $z \equiv a\theta + u$

and the ICC:

$E\{(\theta - q_i)\} = 0$ or $b_i = (1-a_i-c_i)\bar{\theta}$, and

$E\{(\theta - q_i), s_i\} = 0$ or $c_i = (\tau_u a^2)^{-1} \tau (1-a_i \alpha_i^{-1})$.

---

8 Indeed, $E\{E\{(\theta - (a_i s_i + c_i \theta^* + b_i)) | s_i\}\} = (1-(a_i+c_i))\bar{\theta} - b_i = 0$.

The FOC of the problem yields that $E((\theta - q_i), \theta^*)) = 0$. The REE makes efficient use of public information $(E((\theta - q_i), \theta^*)) = 0)$ and *privately* efficient use of private information $(E\{(\theta - q_i), s_i\} = 0)$. Note that in the minimization problem a is taken as given by agent i.

To determine an incentive efficient allocation we maximize the expected payoff (minimize the expected loss) subject to ICC and taking into account the information externality. This is in the symmetric case (and without loss of generality):

$\text{Min}_a\ E(\theta - q_i)^2$ subject to $q_i = a s_i + c\,\theta^* + b$, $\theta^* = E\{\theta \mid z\}$ with $z \equiv a\theta + u$

and the ICC:

$E\{(\theta - q_i)\} = 0$ or $b = (1-a-c)\bar{\theta}$ and

$E\{(\theta - q_i), s_i\} = 0$ or $c \equiv c(a) = (\tau_u a^2)^{-1}\, \tau\, (1 - a\,\alpha^{-1})$.

Using the constraint $b = (1-a-c)\,\bar{\theta}$ we can write the objective as $E(\theta - q_i)^2 =$

$L(a) + \phi(a, c)$, where $\phi(a, c) = (1-a-c)^2\, \tau_u a^2\, \tau^{-1}\, (\tau_\theta)^{-1}$.

Incorporating the ICC we obtain $g(a) \equiv \phi(a, c(a)) = \tau_\theta\, (\tau_u \tau)^{-1} \left(\frac{\tau_\varepsilon + \tau}{\tau_\varepsilon} - \frac{1}{a}\right)^2$.

An incentive efficient allocation solves $\text{Min}_a\ L(a) + g(a)$. We can write $L(a) + g(a) = a^2\, \alpha^{-1}\, (\tau_\varepsilon)^{-1} + (\tau_u)^{-1}(\alpha^{-1} - a^{-1})^2$. This is a strictly convex problem on $(0, \alpha)$ and it is easy to see that the solution never lies on $[\alpha, \infty)$. Denote by a^e the (unique and interior) solution.

With this formulation it is transparent that the team solution (not bound by incentive constraints) sets $c = 1-a$ and obtains $\phi(a, c) = 0$. The term g represents the cost of providing incentives which boils down to the departure from the efficient use of public information (which implies that $E((\theta - q_i), \theta^*)) = 0$ or $c = 1-a$). We have also that $g(a^m)$

= 0. Indeed at the market solution we have that $\frac{\tau_\varepsilon + \tau}{\tau_\varepsilon} - \frac{1}{a}$ = 0 and the REE makes efficient use of public information.

At the unique and interior solution we have that L'(a) +g'(a) = 0. It is possible to show that g'(a) > 0 provided a < 1 (which is the relevant range) and therefore we have that 1 > $a^o > a^e$ (at the team solution $L'(a^o) + g'(a^o) > 0$ because $L'(a^o) = 0$). Furthermore we have that $a^e > a^m$. Indeed, it is easy to see that $L'(a^m) < 0$ and $g'(a^m) = 0$.

At the REE the marginal impact on welfare of an increase in the weight put on private information is the same at the incentive efficient and at the team problems. Indeed, we have that $g'(a^m) = 0$ and therefore $L'(a^m) + g'(a^m) = L'(a^m) < 0$. This is so because at the REE public information is exploited efficiently.

The team solution calls for a larger weight on private information than the incentive efficient solution. The reason is that the more we depart from the REE solution (see Figure 1) by making agents put more weight on their private signals, the more costly is to provide incentives because we have to exploit public information less and less efficiently. The incentive efficient solution resolves the trade off at an intermediate point between the market and the team optimal solutions.

The cost to provide incentives is increasing with the prior precision τ_θ. With more (prior) uncertainty it is cheaper to provide incentives and a^o and a^e are closer together. Indeed as τ_θ tends to 0, both 1-c-a (note that α tends to 1) and g(a) tend to 0. Let us consider the limit case in which the prior is diffuse: $\tau_\theta = 0$. Then the constraints yielding the efficient use of public information $(E((\theta - q_i), \theta^*)) = 0)$ and the *privately* efficient use of private information $(E\{(\theta - q_i), s_i\} = 0)$ (ICC) collapse to c = 1-a and the team solution is incentive efficient. When $\tau_\theta = 0$ agents do not put any weight on the prior mean $\overline{\theta}$ in any circumstance and therefore necessarily the weights to private and public information have to add up to one.

Let $v^e = L(a^e) + g(a^e)$. In summary:

<u>Proposition 4</u>. At the incentive efficient decision rule the weight to private information a^e fulfils $1 > a^o > a^e > a^m$ and is increasing in τ_ε and decreasing in τ_θ and τ_u. The loss v^e is decreasing in τ_ε and τ_u.[9] When $\tau_\theta = 0$ we have that the team solution is incentive efficient, $a^o = a^e$.

The loss at the incentive efficient solution v^e can be increasing or decreasing in τ_θ (note that $L/\tau_\theta < 0$ but $g/\tau_\theta > 0$). We have that $v^e / \tau_\theta = a^e \alpha^{-1} (a^e/\tau_\varepsilon)^2 ((a^e)^{-1} - 2\alpha^{-1})$.
This means that in some circumstances the planner would like to add noise to the prior information. The benefit of making prior information more noisy is relaxing the ICC. For example, when $\tau_\theta = 0$, $\tau_u = 8$ and $\tau_\varepsilon = 1$ we have that $a^e = 1/2$ and $v^e / \tau_\theta = 0$. It follows that for $\tau_u > 8$ or $\tau_\varepsilon < 1$, $a^e < 1/2$ and v^e is increasing in τ_θ.

For extreme values of the parameters τ_ε and τ_u the information externality disappears. Indeed, as τ_ε tends to 0 (uninformative signals), α and a^o, a^e, and a^m tend to 0 and as τ_ε tends to infinity (perfectly informative signals), α and a^o, a^e, and a^m tend to 1. As τ_u tends to 0 (no public information), the ICC c(a) tends to 0 and a^o, a^e, and a^m tend to α and as τ_u tends to infinity (perfect public information), c(a) tends to $c = 1 - \alpha/a$ and a^o, a^e, and a^m tend to 0.

With respect to the team solution the weight to private (public) information is too low (high) at the market solution.

---

9 We have that $^2(L+g)/a\tau_u > 0$, $^2(L+g)/a\tau_\theta > 0$ and $^2(L+g)/a\tau_\varepsilon < 0$. It is immediate also that $(L+g)/\tau_u < 0$ and that $(L(a^e)+g(a^e))/\tau_\varepsilon < 0$.

This is also the case with respect to the incentive efficient solution although the latter introduces a (second best) inefficiency putting some weight to the prior mean on top of θ^*.

Remark: Neither at the team nor at the incentive efficient solution it pays to add noise to the public statistic (increase τ_u).

We can decompose the welfare loss at the market solution with respect to the incentive efficient solution $v^m - v^e$ as an information externality term $(v^m - v^o)$ minus the incentive cost $(v^e - v^o)$:

$$v^m - v^e = (v^m - v^o) - (v^e - v^o).$$

We can think of the ratio $(v^e - v^o)/(v^m - v^o)$ as the cost of providing incentives relative to the benefit of coordination (internalizing the information externality).

When the incentive cost is large then the market is close to the incentive efficient solution; when it is small then the team solution is close to the incentive efficient solution. For a high (low) cost of providing incentives the incentive efficient solution looks like the market (team) solution.

It is instructive to simulate the relative welfare losses at the market solution $(v^m - v^e)/v^e$ and the relative cost of providing incentives $(v^e - v^o)/(v^m - v^o)$ as a function of the deep parameters of the model. The simulations (see Table 1) show that $(v^m - v^e)/v^e$ is decreasing in τ_θ, increasing in τ_u and hump-shaped shaped in τ_ε. The same pattern holds for $(v^m - v^o)/v^o$ and $(v^e - v^o)/v^o$ with the exception that $(v^e - v^o)/v^o$ is hump-shaped shaped in τ_θ. Relative welfare losses can be substantial for τ_θ low, τ_u high and τ_ε intermediate. For example, for $\tau_\theta = .1$, $\tau_u = 5$ and $\tau_\varepsilon = .5$, $(v^m - v^e)/v^e$ is more than 9%. However, for many parameter combinations the relative loss remains very moderate. The relative cost of providing incentives $(v^e - v^o)/(v^m - v^o)$ is increasing in τ_θ and decreasing in τ_u and τ_ε. The market is close to incentive efficient when this relative cost is high.

Table 1: Comparative statics of the relative welfare loss (τ_u and τ_ε)

$\tau_\theta = 1$, $\tau_\varepsilon = 0.5$

τ_u	$\frac{v^m - v^e}{v^e}$	$\frac{v^m - v^o}{v^o}$	$\frac{v^e - v^o}{v^o}$	$\frac{v^e - v^o}{v^m - v^o}$
0.1	$3.72 * 10^{-5}$	0.01	0.01	99.7
0.5	$3.4 * 10^{-3}$	0.21	0.21	98.4
1	0.0196	0.71	0.69	97.3
2	0.092	2.01	1.91	95.3
5	0.49	5.99	5.47	91.3
20	2.93	18.31	14.93	81.5
100	10.80	42.05	28.19	67.1

$\tau_\theta = 1$, $\tau_\varepsilon = 2$

τ_u	$\frac{v^m - v^e}{v^e}$	$\frac{v^m - v^o}{v^o}$	$\frac{v^e - v^o}{v^o}$	$\frac{v^e - v^o}{v^m - v^o}$
0.1	$2.74 * 10^{-4}$	0.01	$9.72 * 10^{-3}$	97.3
0.5	0.0196	0.18	0.16	89.3
1	0.092	0.54	0.45	82.9
2	0.34	1.38	1.03	75
5	1.33	3.76	2.40	63.8
20	5.54	11.34	5.49	48.4
100	16.29	27.36	9.52	34.8

$\tau_\theta = 1$, $\tau_u = 0.5$

τ_ε	$\frac{v^m - v^e}{v^e}$	$\frac{v^m - v^o}{v^o}$	$\frac{v^e - v^o}{v^o}$	$\frac{v^e - v^o}{v^m - v^o}$
0.05	$1.24 * 10^{-7}$	$2.35 * 10^{-3}$	$2.35 * 10^{-3}$	100
0.1	$5.07 * 10^{-6}$	$1.41 * 10^{-2}$	$1.41 * 10^{-2}$	100
0.5	$3.4 * 10^{-3}$	0.218	0.215	98.5
1	0.01298	0.267	0.254	95.1
2	0.0196	0.183	0.163	89.3
5	0.0107	0.054	0.043	80.2
20	$7.9 * 10^{-4}$	$2.73 * 10^{-3}$	$1.95 * 10^{-3}$	71.2

$\tau_\theta = 1$, $\tau_u = 5$

τ_ε	$\frac{v^m - v^e}{v^e}$	$\frac{v^m - v^o}{v^o}$	$\frac{v^e - v^o}{v^o}$	$\frac{v^e - v^o}{v^m - v^o}$
0.05	$1.11 * 10^{-4}$	0.262	0.262	100
0.1	0.003539	1.38	1.37	99.8
0.5	0.49	5.99	5.47	91.3
1	1.06	5.35	4.24	79.3
2	1.33	3.76	2.40	63.8
5	0.96	1.74	0.78	44.7
20	0.194	0.26	0.07	26.7

Table 1: Comparative statics of the relative welfare loss (τ_θ)

$\tau_\varepsilon = 0.5$, $\tau_u = 0.5$,

τ_θ	$\frac{v^m - v^e}{v^e}$	$\frac{v^m - v^o}{v^o}$	$\frac{v^e - v^o}{v^o}$	$\frac{v^e - v^o}{v^m - v^o}$
0	1.91	1.91	0	0
0.01	1.78	1.90	0.11	5.99
0.1	0.95	1.74	0.78	44.8
0.5	0.064	0.79	0.72	91.8
1	$3.4 * 10^{-3}$	0.218	0.215	98.5
2	$4.66 * 10^{-5}$	0.024	0.024	99.8
5	$4.2 * 10^{-8}$	$5.64 * 10^{-4}$	$5.64 * 10^{-4}$	100

$\tau_\varepsilon = 0.5$, $\tau_u = 5$,

τ_θ	$\frac{v^m - v^e}{v^e}$	$\frac{v^m - v^o}{v^o}$	$\frac{v^e - v^o}{v^o}$	$\frac{v^e - v^o}{v^m - v^o}$
0	13.01	13.01	0	0
0.01	12.55	12.95	0.35	2.7
0.1	9.17	12.37	2.93	24
0.5	2.45	9.52	6.89	72.4
1	0.49	5.99	5.47	91.7
2	0.023	1.71	1.69	98.6
5	$3.88 * 10^{-5}$	0.056	0.056	100

$\tau_\varepsilon = 2$, $\tau_u = 0.5$,

τ_θ	$\frac{v^m - v^e}{v^e}$	$\frac{v^m - v^o}{v^o}$	$\frac{v^e - v^o}{v^o}$	$\frac{v^e - v^o}{v^m - v^o}$
0	0.25	0.25	0	0
0.01	0.24	0.25	$8.2 * 10^{-3}$	3.22
0.1	0.19	0.26	0.07	26.7

0.5	0.06	0.24	0.18	73
1	0.019	0.18	0.16	89.3
2	$2.16 * 10^{-3}$	0.079	0.077	97.4
5	$1.85 * 10^{-5}$	$8.2 * 10^{-3}$	$8.1 * 10^{-3}$	99.8
20	$6.61 * 10^{-10}$	$3.52 * 10^{-5}$	$3.52 * 10^{-5}$	100

$\tau_\varepsilon = 2$, $\tau_u = 5$,

τ_θ	$\frac{v^m - v^e}{v^e}$	$\frac{v^m - v^o}{v^o}$	$\frac{v^e - v^o}{v^o}$	$\frac{v^e - v^o}{v^m - v^o}$
0	4.8	4.8	0	0
0.01	4.74	4.79	0.05	1
0.1	4.21	4.71	0.477	10.1
0.5	2.52	4.32	1.75	40.8
1	1.33	3.76	2.4	64.1
2	0.37	2.61	2.23	85.4
5	$6.47 * 10^{-7}$	$3.53 * 10^{-3}$	$3.53 * 10^{-3}$	100

5. The role of tax-subsidy schemes

We have seen how the degree of inefficiency of herding at the market solution depends on the cost of incentive provision. We have derived our results without allowing transfers in the mechanism. With transfers agents could be incentivated directly to put a larger weight on their private signals. This is how it could be done. Suppose that a tax-subsidy scheme is instituted according to which a tax $t\,(q_i - \bar{\theta})^2$, with t a real number, is levied for any departure of the prediction q_i form the prior mean $\bar{\theta}$. Note that $\bar{\theta}$ is known, as well as q_i ex post. The problem of agent i is now

$$\text{Min}_q\ [E\{(\theta - q)^2 \mid I_i\} + t\, E\{(q - \bar{\theta})^2 \mid I_i\}],$$

where $I_i = \{s_i, p\}$ and p solves the equation $p = \int_0^1 q_i\,(s_i, p)\, di + u$.

It is very easy to see that the solution to this problem is $q_i = E\{\tilde{\theta} \mid I_i\}$ where $\tilde{\theta} = (\theta + t\bar{\theta})/(1+t)$. It follows then that the weight to private information a is decreasing in t and that the range of possible a's is $[a^m$, ¥) as t ranges from 0 to -1. Indeed, to induce agents to put a larger weight on their signal a subsidy must be given to reward departures from the prior mean.

The optimal subsidy scheme will then minimize the average prediction loss plus subsidy (where an extra cost of raising public funds could be included). The result will be a weight to private information larger than the market weight a^m but smaller than the team optimal one a^o. Increasing the cost of public funds will diminish the optimal subsidy and approach the solution to the subsidy-free market solution a^m.

6. Concluding remark

The paper has analyzed a very stylized model inspired in the herding prediction models. Several extensions could be examined. First, the analysis could be extended to non-linear environments to check the robustness of results. Second, we have only considered information externalities but not direct payoff externalities among agents. Obviously, most economic situations involve payoff externalities. Further work should study the interaction between informational and payoff externalities. In Messner and Vives (2001) we study informational and economic efficiency in markets in which firms compete in supply functions.

References

Banerjee, A. (1992), "A Simple Model of Herd Behavior", *Quarterly Journal of Economics*, 107 (3), 797-817.

Bikhchandani, S, D. Hirshleifer and I. Welch (1992), "A Theory of Fads, Fashion, Custom and Cultural Change as Informational Cascades", *Journal of Political Economy*, 100 (51), 992-1026.

Chamley, C. and D. Gale (1994), "Information Revelation and Strategic Delay in a Model of Investment", *Econometrica*, 62 (5), 1065-85

Feldman, M and C. Gilles (1985), "An Expository Note on Individual Risk without Aggregate Uncertainty", *Journal of Economic Theory,* 35 (1), 26-32.

Gale, D. (1996), "What Have We Learned from Social Learning", *European Economic Review* , 40 (3-5), 617-628.

Gul, F. and R. Lundholm (1995), "Endogenous Timing and the Clustering of Agents' Decisions", J*ournal of Political Economy*, 103(5), 1039-66.

Laffont, J. J. (1985), "On the Welfare Analysis of Rational Expectations Equilibria with Asymmetric Information," *Econometrica*, 53, 1-29.

Lee, I. (1993), "On the Convergence of Informational Cascades", *Journal of Economic Theory*, 61, 395-411.

Messner, S. and X. Vives (2001), "Informational and Economic Efficiency in REE with Asymmetric Information", CEPR DP 2678, London.

Smith, L. and P. Sorensen (2000), "Pathological Outcomes of Observational Learning", *Econometrica*, 68 (2), 371-398.

Vives, X. (1993), "How fast do rational agents learn ?", *Review of Economic Studies,* 60, 329-347.

Vives, X. (1996), "Social Learning and Rational Expectations", *European Economic Review*, 40, 589-601.

Vives, X. (1997), "Learning from Others: A Welfare Analyis," *Games and Economic Behavior*, 20, 177-200.