Available online at www.sciencedirect.com

SciVerse ScienceDirect

Journal of Economic Theory 147 (2012) 910–940

JOURNAL OF Economic Theory

www.elsevier.com/locate/jet

Learning from private and public observations of others' actions ☆

Manuel Amador [a,b,1], Pierre-Olivier Weill [c,b,d,*]

[a] *Landau Economics Building, 579 Serra Mall, Stanford, CA 94305-6072, United States*
[b] *NBER, United States*
[c] *Bunche 8283, University of California, Los Angeles, Los Angeles, Box 951477, CA 90095, United States*
[d] *CEPR, United Kingdom*

Received 7 September 2008; final version received 20 June 2011; accepted 15 August 2011

Available online 24 February 2012

Abstract

We study the diffusion of dispersed private information in a large economy, where agents learn from the actions of others through two channels: a public channel, such as equilibrium market prices, and a private channel, for example local interactions. We show that, when agents learn *only* from the public channel, an initial release of public information increases agents' total knowledge at all times and increases welfare. When a private learning channel is present, this result is reversed: more initial public information reduces agents asymptotic knowledge by an amount in order of $\log(t)$ units of precision. When agents are sufficiently patient, this reduces welfare.

☆ For fruitful discussions and suggestions, we'd like to thank George-Marios Angeletos, Bruno Biais, Ricardo Caballero, Amil Dasgupta, Chris Edmond, Raquel Fernandez, Doireann Fitzgerald, Alessandra Fogli, Boyan Jovanovic, Douglas Gale, Christian Hellwig, Thomas Mariotti, Andy Skrzypacz, Stijn Van Nieuwerburgh, Dimitri Vayanos, Laura Veldkamp, Xavier Vives, Ivan Werning, and seminar participants at ECB, UCSD, Princeton, NYU, Harvard Junior Lunch, Stanford Junior Lunch, University of Montreal, LSE, Toulouse, Columbia, MIT, Penn State, USC, Society of Economics Dynamics Meeting (2006), Stanford Institute for Theoretical Economics (2006), Southwestern Economic Theory Conference (2007) at USC, Minnesota Workshop in Macroeconomic Theory (2006), Federal Reserve Bank of Cleveland Summer Workshop in Money, Banking, and Payment (2006), IESE's Conference on Complementarities and Information (2007), for their comments. We are grateful to the Associate Editor and two anonymous referees for comments that greatly improved the paper. We also thank Romans Pancs for excellent research assistance.
* Corresponding author at: Bunche 8283, University of California, Los Angeles, Los Angeles, Box 951477, CA 90095, United States. Fax: +1 (310) 825 9528.

E-mail addresses: amador@stanford.edu (M. Amador), poweill@econ.ucla.edu (P.-O. Weill).

[1] Fax: +1 (650) 725 5702.

0022-0531/$ – see front matter
doi:10.1016/j.jet.2012.02.001

JEL classification: D80; D83; E58

Keywords: Information aggregation; Learning; Value of information

1. Introduction

Households and firms have knowledge from their local markets about aggregate economic conditions. This dispersed private knowledge diffuses in the population over time as households and firms learn from each other valuable bits of information that are not yet known to everyone. Some channels of information diffusion are private: for example private information comes from observing the actions of others in local markets. Other channels are public, such as the information aggregated in prices, or in the macroeconomic figures published by agencies. In this paper, we study how private information dispersed in a large economy diffuses through such private and public channels. The main result of the paper is to show that these two channels of information diffusion generate different diffusion dynamics and more importantly, generate opposite social values of public information. When agents learn only from public channels, a release of public information would always increase welfare. By contrast, when agents also learn from private channels, a release of public information can reduce welfare.

Our baseline model is set in continuous time and builds on the discrete-time environments of [38,39]. We consider a continuum of agents who, at time zero, receive both public and private signals about the state of the world. This is the only exogenous source of information in the model. After time zero, each agent takes an action at every moment until some random time when the state of the world is revealed and her payoff is realized. We assume that an agent's payoff decreases with the distance between the sequence of actions and the revealed state, and it is independent of the actions of any other agent. After receiving the initial information, but before the state of the world is revealed, each agent learns continuously by observing two noisy signals about the actions of others. The first signal is private, only observed by the agent: this constitutes the private learning channel and represents information gathered through private communication and local interactions. The second signal is public, shared with everyone: this constitutes the public learning channel and represents information gathered from observing an endogenous aggregate variable, such as a price or some macroeconomic aggregate.

We solve for an equilibrium in which agents eventually learn the truth. Agents in the equilibrium take actions that are the convex combinations of the endogenous signals generated by the private and the public information observations of others' actions. As is standard, the convex weight agents put on their private information determines the informativeness of the endogenously generated signals. We show that the existence of the private social learning channel guarantees that the weight on private information and thus the informativeness of all the endogenous signals converges to a strictly positive constant in the long run. Therefore, asymptotically, the precision of agents' beliefs increases linearly towards infinity.

As in the herding literature, a public information release creates a negative externality as it reduces the informativeness of endogenously generated signals. To see this, suppose now that a benevolent agency holds some relevant, but partial, information and ponders whether to release it publicly at time zero. Clearly, the release has the direct beneficial effect of making agents' current decisions better informed. The negative externality is that agents respond by reducing, temporarily, the weight that they put on private information, which slows down learning. We show that when agents only learn through the public channel, this negative externality on learning is sufficiently weak that a release of public information increases the amount known at all times and is always socially beneficial. By contrast, when agents also learn through the private channel, the negative externality is amplified: since each agent accumulates less information from the private channel, her actions become less sensitive to private information, slowing down information diffusion even more, and a negative feedback loop ensues. This so slows information diffusion that agents end up less informed in the long run: namely, we show that, asymptotically, increasing the precision of the initial public signal by one unit reduces knowledge at time t by an amount in order of $\log(t)$ units of precision. That is, public information releases generate unbounded negative effects on agents' knowledge. Further, and in line with this intuition, we show that the negative effect of public information is larger when the private learning channel is more effective.

The social benefit of releasing public information depends then on the trade-off between increasing the amount known currently by all agents, and reducing the amount known in the future. We show that if agents are sufficiently patient then a given marginal increase in the precision of the initial public signal is always socially costly.

In the final section of the paper we explore the robustness of our results by analyzing the socially optimal diffusion of information: we study the problem of a planner that can choose the sensitivity of the agents' actions to their private and public forecasts, as in [39]. We show that, after the planner corrects the information externality, public information always increases *ex-ante* social welfare. Surprisingly, welfare is reduced *ex-post*: the planner finds it optimal to make agents less informed in the long run, by the same $\log(t)$ amount as in the decentralized equilibrium.

1.1. Related literature

Our work is related to the recent literature on the social value of public information. As initially shown by [29], public information can reduce welfare in the presence of a payoff externality. [4] provided a complete characterization of the effect in general linear–quadratic models, and [24] studied the implications for a monetary economy. However, the models used so far have been essentially static, and abstracted from learning. Our contribution is to analyze an alternative mechanism based on a dynamic information externality: in our baseline model there are no payoff externalities, but public information slows down the diffusion of private information in the population.[2]

The social learning literature, started by [10] and [13], has been concerned with information externalities. Its central result is the possibility of informational cascades and herds: agents may choose to disregard their private information, acting solely on the basis of the public information, and take the "wrong" action. Thus public information, by facilitating the emergence of herds, can reduce welfare. However, the standard herding models are sequential move games, and the appearance of cascades and herds requires bounded beliefs and a discrete set of actions. Our model is closest to [38] and [39] instead, where beliefs are unbounded and actions lie in a continuous space. The maintained assumption among these social-learning papers is that private information diffuses through public channels. Our paper allows information to diffuse through both public

[2] [30] sets up a model in which a central bank ends up learning less from the actions of private agents after disclosing public information. But since private agents do not accumulate information over time, they do not suffer from our dynamic learning externality. In fact, absent payoff externalities, public information improves social welfare in their model.

and private channels, and we show that this has different implications for dynamics and welfare. Another closely related paper that also emphasizes negative effects of public information, when coupled with an information externality is [3]. The focus of that paper is however essentially static, while the focus of the present paper is on the dynamics of information diffusion and the associated long-run welfare effects.

[16] presents a model where agents exert effort to collect private information, and show that public information reduces the incentives of agents to gather new information and can reduce welfare. In our model, agents do not collect new information: we focus instead on the question of how information, that is already in the hands of agents, aggregates and diffuses.[3] In many situations it is reasonable to expect that agents can learn at no cost from each other, for example through their ongoing market interactions or through the observation of public prices.

Strategic experimentation models study the interplay between information acquisition and information diffusion. Recent examples include [15,12,17,26,33]. The key inefficiency is a free-riding problem: given that other agents' actions generate observable new information, an individual agent has less incentive to learn on his own by taking the costly action. No such free riding is at a play in our model where information is generated at no cost, and where an individual agent's decision problem is not affected by other agents' contemporaneous and future decisions, but only by their past decisions.

Some recent work on social learning has focused on learning in networks: [8,21,35] study deterministic networks with a finite number of agents, [9] provides a continuum-of-agents setup, and [18] proposes a network of boundedly rational agents. The private learning channel of the present paper is arguably a reduced form model of local interactions through networks. However, our model is tractable enough to address questions that would be more difficult in an explicit network model.

Another body of research uses search-and-matching models to study how agents learn from local interactions with others (see for example [41] seminal work and the recent work of [27]). The issue of convergence to the truth has also been addressed in [23,14,22]. The independent work of [19] characterizes learning dynamics in a search and matching model where agents exchange information in multilateral meetings.[4] Their subsequent work, [20], provides results on learning speed related to ours. They focus on the positive impact of private learning on information dissemination, but they abstract from its interaction with learning externalities and the associated negative welfare effect of public information. [40,25,6,7] address learning about a money supply shock in a random-matching model similar to [37]. Our setup is related to this literature because the private signals about aggregate actions can be interpreted as the result of random local interactions. The benefit of this simplification is that we can characterize transitional dynamics of beliefs and compute the social value of public information.

The rest of this paper is organized as follows. Section 2 introduces the setup. Section 3 solves for an equilibrium. Section 4 studies the impact on dynamics and welfare of changes in the quality of public information. Section 5 analyzes optimal information diffusion. Section 6 concludes. Appendices A–C collect all the proofs not in the main text. An addendum to this paper [2], presents additional results and extensions of our model.

[3] Our results regarding the social value of public information complement theirs, and would apply in a costly information acquisition setup after the agents have stopped gathering information but could still observe signals about the aggregate market behavior.

[4] A previous version of our current paper, [1], also studied the learning dynamics within a matching framework.

2. Set up

Time is continuous and runs forever. We fix a probability space $\{\Omega, \mathcal{G}, Q\}$ together with an information filtration $\{\mathcal{G}_t, t \geqslant 0\}$ satisfying the usual conditions [32]. The economy is populated by a $[0, 1]$-continuum of agents. Each agent's payoff depends on some unknown state of the world $x \in \mathbb{R}$. At each time before some exponentially distributed "day of reckoning" $\tau > 0$, with parameter λ, each agent takes an action $a_{it} \in \mathbb{R}$. At time τ, the state of the world x is revealed and each agent receives the payoff

$$-\int_0^\tau (a_{it} - x)^2 \, dt.$$

Agents are endowed with a diffused common prior that x is normally distributed with mean zero and zero precision.[5] Over time, they observe a public signal Z_t and a private signal z_{it}. The initial realizations of these signals are

$$Z_0 = x + \frac{W_0}{\sqrt{P_0}} \quad \text{and} \quad z_{i0} = x + \frac{\omega_{i0}}{\sqrt{p_0}} \tag{1}$$

where W_0 and $(\omega_{i0})_{i \in [0,1]}$ are normally distributed with mean zero and variance one, pairwise independent, and independent of everything else. In Eq. (1), P_0 and p_0 represent the respective precisions of the public and the private signal.

The initial signal, Z_0, represents information released by a public agency. By continuously varying its precision, P_0, we will obtain the impact on diffusion and welfare of varying the size of an information release. The continuum of initial private signals, $(z_{i0})_{i \in [0,1]}$, makes agents asymmetrically informed about x, and represents dispersed information about aggregate economic conditions.

At all times after time zero the public and the private signals evolve according to the stochastic differential equations:

$$dZ_t = A_t \, dt + \frac{dW_t}{\sqrt{P_\varepsilon}} \quad \text{and} \quad dz_{it} = A_t \, dt + \frac{d\omega_{it}}{\sqrt{p_\varepsilon}}, \tag{2}$$

where $A_t \equiv \int_0^1 a_{it} \, di$ is the cross-sectional average action at time t, and where W and $(\omega_i)_{i \in [0,1]}$ are pairwise independent Wiener processes with initial conditions W_0, ω_{i0}, and are independent from everything else.

There is a key difference between the initial signal realization (1) and the subsequent signal realizations (2): the former is centered around the true state of the world, x, while the later is centered around the average action, A_t. This means that, after time zero, all the information about x that agents learn comes from others, and there is no additional arrival of "new" information. Indeed, the realizations dZ_t and dz_{it} are signals about the information that others accumulated in the past so, ultimately, they are signals about the initial realizations Z_0 and z_{i0}.

The public signal, dZ_t, represents the information conveyed by some endogenous aggregate variable (such as prices). The private signal, dz_{it}, on the other hand, captures the decentralized

[5] The assumption of a diffused prior is without loss of generality. In an online Addendum, available from the authors' website, we consider the general case of a prior with a strictly positive precision, and show that our results for the dynamics of beliefs and welfare do not change.

gathering of information. One could think, for instance, of local interaction and of private communication, such as gossips.[6,7]

Given a process A for the average action, we let $\mathcal{G}_{it}$ be the filtration generated by $\{(Z_s, z_{is}),\ 0 \leqslant 0 \leqslant t\}$, representing all the information available to agent $i \in [0,1]$ at any time $t > 0$. Because agents are infinitesimal, their actions do not affect the average action process A, and hence do not affect the information they receive. So, the agents' inter-temporal problems are essentially static, and together with their quadratic payoffs, this implies that optimal actions are the expectation

$$a_{it} = \mathbb{E}[x \mid \mathcal{G}_{it}] \tag{3}$$

of the random variable x, conditional on their information filtration $\{\mathcal{G}_{it},\ t \geqslant 0\}$. Finally, in an equilibrium, these individual actions have to generate the average:

$$A_t = \int_0^1 a_{it}\, di. \tag{4}$$

We summarize all the above in the following:

Definition 1. An *equilibrium* is a collection of processes a_i and A solving (2), (3), and (4).

3. An equilibrium

We now show that there exists an equilibrium in which agent i's action at any time is the convex combination of two forecasts of the state of the world: a forecast containing *only* the information shared with everyone in the economy, which we denote by $\hat{X}_t$, and a forecast containing *only* the information observed by agent i and no one else, which we denote by $\hat{x}_{it}$. In what follows, we will abuse language and call $\hat{X}_t$ the "public forecast," and $\hat{x}_{it}$ the "private forecast."

Let us guess for now that these forecasts are normally distributed, independent given x, and that their precisions are common knowledge. Denoting by P_t and p_t the precision of the public and the private forecast, Bayesian updating implies that the action taken by agent i at time t is, then,

$$a_{it} = \mathbb{E}[x \mid \mathcal{G}_{it}] = \frac{P_t}{P_t + p_t}\hat{X}_t + \frac{p_t}{P_t + p_t}\hat{x}_{it}, \tag{5}$$

a "precision weighted" convex combination of the public and private forecasts. The public forecast is, of course, the same for every agent. The private forecasts, on the other hand, are unbiased and based on independent private information: thus, their cross-sectional average must equal x.

[6] In the previous version of this paper, [1], we suggest that specification (2) may arise when each agent continuously observes, with idiosyncratic noise, the action of other randomly chosen agents. Intuitively, observing the action of a randomly chosen agent amounts to sampling from a distribution centered around the average action, A_t. When the time between periods and the precision of the noise go to zero at the same rate, we informally arrive at specification (2).

[7] In [3], we provide a different interpretation of the endogenous private signal, dz_{it}: they are generated because agents receive exogenous private signals about the noise in public endogenous aggregates. We show how such signals about the noise naturally arise in a monetary economy.

These observations mean that the average action is

$$A_t = \int_0^1 a_{it}\,di = \frac{P_t}{P_t + p_t}\hat{X}_t + \frac{p_t}{P_t + p_t}x.$$

Recall that the public and private signals dZ_t and dz_{it} of Eq. (2) are centered around the average action A_t. But since the public forecast, $\hat{X}_t$, and the precisions, P_t and p_t, are all common knowledge, we obtain an informationally equivalent set of signals

$$x\,dt + \frac{dW_t}{\sqrt{P_\varepsilon(\frac{p_t}{P_t+p_t})^2}} \quad \text{and} \quad x\,dt + \frac{d\omega_{it}}{\sqrt{p_\varepsilon(\frac{p_t}{P_t+p_t})^2}} \tag{6}$$

after first subtracting off $P_t/(P_t + p_t)\hat{X}_t\,dt$ from dZ_t and dz_{it}, and then dividing by $p_t/(P_t + p_t)$. Eq. (6) defines two equivalent "transformed" public and private signals which have the convenient properties of being centered around x and independent given x.

It is then straightforward to verify our guess. We let the public forecast $\hat{X}_t$ be the expectation of x conditional on the common prior and the history of the "transformed" public signal defined by the left side of Eqs. (1) and (6). Similarly, the private forecast $\hat{x}_{it}$ is the expectation of x conditional on the history of the "transformed" private signal, defined by the right side of Eqs. (1) and (6). The precisions of the public and private forecasts, P_t and p_t, are readily characterized by a system of Ordinary Differential Equations (ODEs):

$$dP_t = P_\varepsilon\left(\frac{p_t}{P_t + p_t}\right)^2 dt, \tag{7}$$

$$dp_t = p_\varepsilon\left(\frac{p_t}{P_t + p_t}\right)^2 dt, \tag{8}$$

with initial conditions P_0 and p_0, respectively. The intuition follows standard Bayesian updating formulas with independent signals: for instance, the change in the public precision, dP_t, is equal to the precision of the transformed public signal $P_\varepsilon(\frac{p_t}{p_t+P_t})^2$. The term inside the brackets is the weight that agents put on their private information when taking their actions. This weight controls the informativeness of the signals and as a result, the speed at which information diffuses.

The ODEs show that the informativeness of the public and the private signals at time t is a function of the precisions P_t and p_t of the public and private forecasts. This informativeness decreases with P_t and increases with p_t. That is, the more the agents know privately, the more informative the new signals become, and the faster agents learn. Improvements in public knowledge have the opposite effect: they slow down subsequent learning. This suggests that public and private learning affect the diffusion of information in the economy differently.

Closed form solution and learning asymptotic. Note that ODE (7) is equal to ODE (8) multiplied by $P_\varepsilon/p_\varepsilon$. So, as long as $p_\varepsilon > 0$, this implies that $P_t - (P_\varepsilon/p_\varepsilon)p_t$ stays constant over time, and $(P_t - P_0) = (p_t - p_0)P_\varepsilon/p_\varepsilon$. Plugging this into (8), we obtain

$$\dot{p}_t = p_\varepsilon\left(\frac{p_t}{\alpha + \beta p_t}\right)^2 \tag{9}$$

where $\alpha \equiv P_0 - (P_\varepsilon/p_\varepsilon)p_0$ and $\beta \equiv 1 + P_\varepsilon/p_\varepsilon$. Hence given an initial condition for the precision p_0 of the private forecast, and using (9), it is possible to solve analytically for p_t.

Theorem 1 *(A closed form solution). There exists an equilibrium in which an agent's belief can be decomposed into two independent public and private forecasts with respective precisions P_t and p_t solving*:

- *for* $p_\varepsilon = 0$:

$$p_t = p_0 \quad \text{and} \quad P_t + p_t = \left(3p_0^2 P_\varepsilon t + (p_0 + P_0)^3\right)^{1/3}, \tag{10}$$

- *for* $p_\varepsilon > 0$:

$$H(p_t) = H(p_0) + p_\varepsilon t \quad \text{and} \quad P_t + p_t = \alpha + \beta p_t, \tag{11}$$

where $H(p) = 2\alpha\beta \log p + \beta^2 p - \alpha^2/p$, $\alpha = P_0 - (P_\varepsilon/p_\varepsilon)p_0$, *and* $\beta = 1 + P_\varepsilon/p_\varepsilon$.

This result allows us to characterize the asymptotic dynamics of the entire system:

Corollary 1 *(Precisions asymptotic).* (i) *The precision of the private forecast monotonically converges to infinity as long as* $p_\varepsilon > 0$, (ii) *the weight on private information,* $p_t/(P_t + p_t)$ *decreases towards* $p_\varepsilon/(p_\varepsilon + P_\varepsilon)$ *if* $p_0/(P_0 + p_0) > p_\varepsilon/(p_\varepsilon + P_\varepsilon)$, *and increases towards* $p_\varepsilon/(p_\varepsilon + P_\varepsilon)$ *otherwise, and* (iii) *as* $t \to \infty$ *the total precision,* $P_t + p_t$, *is such that*

$$P_t + p_t = \begin{cases} (3p_0^2 P_\varepsilon t)^{1/3} + Q_t, & \text{for } p_\varepsilon = 0, \\ (\frac{p_\varepsilon}{P_\varepsilon + p_\varepsilon})^2 (p_\varepsilon + P_\varepsilon)t + 2(\frac{P_\varepsilon}{p_\varepsilon} p_0 - P_0)\log(t) + R_t, & \text{for } p_\varepsilon > 0, \end{cases} \tag{12}$$

where Q_t *converges to zero and* R_t *has a finite limit,* R_∞, *which satisfies*:

$$\frac{\partial^2 R_\infty}{\partial P_0 \partial p_\varepsilon} < 0. \tag{13}$$

To start with, let us contrast the leading terms of these two asymptotic expansions (in Section 4.1.2 we provide a detailed intuition for the higher-order, $\log(t)$, term). Consider first the case $p_\varepsilon > 0$. Note that the weight on private information, $p_t/(p_t + P_t)$, is equal to $p_t/(\alpha + \beta p_t)$ which converges to $1/\beta = p_\varepsilon/(p_\varepsilon + P_\varepsilon)$ as p_t goes to infinity. And thus, the precisions of the signals generated by the average action converge to a strictly positive number as long as $p_\varepsilon > 0$. By Eqs. (7) and (8), the sum of the precisions of the endogenously generated private and public signals converges in the limit to:

$$\lim_{t\to\infty} \left(\frac{p_t}{P_t + p_t}\right)^2 (P_\varepsilon + p_\varepsilon) = \left(\frac{p_\varepsilon}{P_\varepsilon + p_\varepsilon}\right)^2 (P_\varepsilon + p_\varepsilon), \tag{14}$$

and this is the coefficient on t in the asymptotic expansion of the precision of total beliefs. As long as $p_\varepsilon > 0$, when time goes to infinity, the social learning process converges to a situation that is as if agents repeatedly observe independent signals centered around x with precision given by (14).

At $p_\varepsilon = 0$, the path of precision is discontinuous, and so its asymptotic behavior is quite different. In case $p_\varepsilon = 0$, from part (ii) of the corollary, the precision of the signals generated by the average action converges to zero. As a result, the speed of convergence of the learning process is greatly reduced asymptotically. In this case, the corollary is the continuous-time counterpart of the well-known result of [38]: when social learning is constrained to public observations of the average action, the precision of total beliefs goes to infinity at rate $t^{1/3}$. This rate is one order of magnitude slower than the rate at which beliefs would converge if agents were to observe, for example, noisy exogenous signals of x in every period with i.i.d. noise – in that case the precision would go to infinity at a linear rate t. However, the precision of beliefs is going to infinity at a linear rate when $p_\varepsilon > 0$, even though no new information is being exogenously provided to agents: that is, the social learning generated from endogenous private signals, no matter how noisy, is sufficient to restore the speed of convergence to its usual linear rate. When $p_\varepsilon > 0$, the accumulation of knowledge is taking place through both the public and the private social learning channels, and this is the key element that maintains the informativeness of the endogenous signals bounded away from zero, thus avoiding that the rate of convergence be dramatically reduced. The existence of a private social learning channel is important, and as will be shown below, it will also be crucial for the social value of public information.

4. The impact of public information

4.1. Comparative dynamics

We fist study the impact of public information on precision dynamics.

4.1.1. When there is no endogenous private learning channel

We first show that when agents do not learn privately from others' action, $p_\varepsilon = 0$, then more public information at the beginning increases agents' knowledge at all times. We also argue that this result is robust: it only requires the social learning process to be smooth.

We consider a generalized version of our model where, in addition to observing a public signal of others' actions, we let agents accumulate exogenous information over time. Think for instance, of additional public and private exogenous signals centered around x. However, social learning is restricted to be *public* and "*smooth*": agents continuously observe public signals about the average action in population. This generalization leads us to formulate the following ODE:

$$dP_t = P_\varepsilon \left(\frac{\pi_t}{P_t + \pi_t} \right)^2 dt + d\Pi_t, \tag{15}$$

where Π_t and π_t denote the cumulative precision of exogenous public and private information. Note that we allow for Π_t and π_t to have jumps: that is, the exogenous information could arrive in a lumpy fashion.

Proposition 1. *Suppose exogenous public and private information increase over time according to piecewise continuously differentiable cumulative precisions Π_t and π_t. Consider two initial levels $P_0' > P_0$ of public information and their associated paths P_t' and P_t. Then $P_t' > P_t$ for all t.*

The proposition follows from (i) the standard mathematical result that two different solutions of a (possibly time dependent) ODE can never cross,[8] and (ii) from our assumption that lumpy arrivals of public information can only arise from exogenous signals, i.e. $d\Pi_t$ does not depend on the current level of public information.

[8] Our ODE is slightly non-standard since we allow for positive jumps. However, it is straightforward to extend the standard existence and uniqueness result to this case. See the online Addendum.

Namely, the differential equation implies that P_t' and P_t are continuous except when there is a lumpy arrival of public information, in which case they increase by the exact same amount. Thus, given $P_0' > P_0$, if there were some time such that $P_t' \leqslant P_t$ then there must be a time $s < t$ such that $P_s = P_s'$. But since both precisions solve the same differential equation the paths of P_t and P_t' would be the same for all times, which is a contradiction.

In the Addendum, we show that the result of Proposition 1 also holds in [38] original *discrete time* setup, under either one of the following three conditions: if the initial public precision is large, if the release in public information occurs after the first period, or if the observational noise per period is small. A large observational noise per period results in a small amount of endogenous learning per period, which is the natural discrete-time counterpart of our continuous time Brownian setup where information flows smoothly from the endogenous learning channel. In fact, the result of the proposition holds for any "smooth" endogenous public learning channel: formally, the same proof would go through if the first term on the right-hand side of the ODE is replaced by some bounded locally Lipschitz function $F(t, P_t)$ of time and current public precision.

4.1.2. When there is an endogenous private learning channel

We now show that, because the path of precision is discontinuous at $p_\varepsilon = 0$, the effect of public information releases change dramatically when $p_\varepsilon > 0$ and the accumulation of private information is endogenous.

From the asymptotic expansion (12) one sees that a unit increase in the initial public precision, P_0, decreases total precision by approximately $2\log(t)$ at time t. Thus, small differences in time-zero public information result in asymptotically unbounded differences in total information: after an initial release of public information, agents eventually know strictly less. This is in sharp contrast with the result obtained when there was no private social learning.

To understand how this $\log(t)$-impact comes about, let us go back to the ODE (9) for $\dot{p}_t$:

$$\dot{p}_t = p_\varepsilon \left(\frac{p_t}{\alpha + \beta p_t} \right)^2 = p_\varepsilon \left(\frac{\alpha}{p_t} + \beta \right)^{-2},$$

where the term inside the first bracket is the weight agents put on their private information at time t.

Note that, since $\alpha = P_0 - P_\varepsilon / p_\varepsilon p_0$, it follows that an increase in P_0 reduces the weight that agents put on their private information, and therefore reduces $\dot{p}_t$. But we know from Corollary 1 that this negative effect on the weight disappears in the limit, as the weight on private information converges to $p_\varepsilon/(p_\varepsilon + P_\varepsilon)$, a constant that is independent on initial conditions. However, as we will show below, the negative effect washes out at the slow speed of $1/t$, where t is given by the asymptotic learning speed. Thus, an increase in initial public precision reduces the rate of precision accumulation, $\dot{p}_t$, by an amount in order $1/t$. These vanishing reductions in $\dot{p}_t$ add up, however, to unbounded differences, as $\int dt/t = \log(t)$.

To see this, first note that $\dot{p}_t$ can be approximated by

$$\dot{p}_t = \frac{p_\varepsilon}{\beta^2} - \frac{2 p_\varepsilon}{\beta^3} \frac{\alpha}{p_t} + o\left(\frac{1}{p_t} \right), \tag{16}$$

as $p_t \to \infty$. The first term of the Taylor expansion, $p_\varepsilon (0 + \beta)^{-2} = p_\varepsilon/\beta^2$, is the asymptotic rate of private precision accumulation. It reflects the fact that the weight on private information goes to a strictly positive constant, $1/\beta$, and implies that private precision goes to infinity at a linear speed:

$$p_t = \frac{p_\varepsilon}{\beta^2} t + o(t). \tag{17}$$

Next, we turn to the second term of the Taylor expansion. The coefficient $-2p_\varepsilon/\beta^3$ measures the effect of a change in α/p_t on $\dot{p}_t$, evaluated at the limit $\alpha/p_t = 0$. This coefficient is strictly negative because, even in the limit, the equilibrium weight that agents put on private precision remains sensitive to changes in α/p_t.[9] After plugging the approximation (17) in (16), the second term of the Taylor expansion becomes

$$-\frac{2p_\varepsilon}{\beta^3}\frac{\alpha}{p_t} = -\frac{2}{\beta}\frac{\alpha}{t}.$$

This term describes how fast the precision of the endogenous private signal is converging to its long-run value of p_ε/β^2. Since $\alpha = P_0 - p_\varepsilon/P_\varepsilon p_0$, it also determines the rate at which the negative impact of increasing P_0 washes out in the long run. Namely, a unit increase in P_0 reduces the time t accumulation of private precision by $(2/\beta)/t$ units, so the negative impact of public information is indeed vanishing with time. But it is doing so at a slow speed: these small $(2/\beta)/t$ differences in the *change* of private precision accumulation add up through time to unbounded differences in the *level* of private precision, which will be in order $\int (2/\beta)(dt/t) = (2/\beta)\log(t)$. The corresponding differences in total precision, $\alpha + \beta p_t$, are thus obtained after multiplying by β, and are in order $2\log(t)$. We summarize this discussion with the following corollary[10]:

Corollary 2. *Let $p_\varepsilon > 0$. Consider two initial levels $P_0' > P_0$ of public information and their associated paths of private and public precisions (p_t', P_t') and (p_t, P_t). Then, for all $M > 0$ there exists $\bar{t} < \infty$ such that $p_t' + P_t' + M < p_t + P_t$ for all $t > \bar{t}$.*

Intuitively, a release of public information has a self-reinforcing negative feedback on the accumulation of private information. Better initial public information causes agents to put more weight on the public information and less weight on their accumulated private information. This reduces subsequent information accumulation from all channels, in particular from the private social learning channel. But this implies that agents accumulate less private information, so they will put even more weight on their public information, and less on their private information, reducing further private information accumulation, and a negative feedback ensues. The corollary makes explicit that the negative feedback is always strong enough to eventually overtake the initial gains in information.

[9] This is indeed the result of the equilibrium behavior. Contrast this with the following: suppose that agents were to follow an exogenous rule that forces them take an action with a weight on private information equal to some given smooth and decreasing function of P_t/p_t. Since this weight governs the accumulation of both public and private information, this dynamic system still has the property that $P_t = \alpha + (\beta - 1)p_t$ so we can write the weight as some function $f(\alpha/p_t) \in [0, 1]$. Let us also impose the reasonable assumption that $f(0) > 0$: in words, as the agents accumulate lots of private information, the rule put some positive weight on it. The evolution of private information is then given by $\dot{p}_t = p_\varepsilon(f(\alpha/p_t))^2$, and p_t will grow linearly in the limit. Whether the second term of a Taylor expansion in α/p_t of the ODE is strictly negative will depend on whether $f'(0)$ is strictly negative: that is, whether a change in α/p_t has a first-order negative effect on the limiting weight.

[10] We omit the proof as it follows directly from the asymptotic expansion (12).

Corollary 2 showed that, with the introduction of a private learning channel, public information starts having a negative effect on asymptotic precision. The next corollary, that arises directly from part (iii) of Corollary 1, shows that this negative effect gets further amplified as the private channel becomes more informative, i.e., as p_ε increases:

Corollary 3. *Consider two initial levels $P'_0 > P_0$ of public information, and let $\Delta_t[p_\varepsilon] \equiv (p'_t + P'_t)[p_\varepsilon] - (p_t + p_t)[p_\varepsilon]$ denote the associated difference in total precision, as a function of p_ε. Then, for any $\hat{p}_\varepsilon > \tilde{p}_\varepsilon$, the double difference, $\Delta_t[\hat{p}_\varepsilon] - \Delta_t[\tilde{p}_\varepsilon]$, has a negative limit.*

The corollary states that given two different levels of initial public information, $P'_0 > P_0$, the *long-run* difference between the total precision generated by initial condition P'_0 and the total precision generated by initial condition P_0 becomes more negative when the private channel is more important, as measured by a higher p_ε. In this sense, the private channel amplifies the long-run negative effects of public information.

4.2. The social value of public information

Whether an initial release of public information is socially beneficial depends on a trade-off between a short-term gain and a long-term loss. The short-term gain is that public information initially improves the precision of agents, which can start making better decisions right away. The long-term loss, is that public information eventually reduces the amount known by everyone as long as the private social learning channel is active. In the proposition that follows, we provide conditions that ensure that the long-term loss dominates: we show that if the state is revealed in a sufficiently long time, on average, then a marginal increase in public information always reduces utilitarian welfare. Hence, unlike [29], we conclude that even in the absence of a payoff externality, more public information can be welfare reducing.

Let the welfare criterion be the equally weighted sum of agents' expected utility. By the law of large numbers, this criterion coincides with the *ex-ante* utility of a representative agent,

$$W \equiv -\lambda \mathbb{E}\left[\int_0^\tau (a_{it} - x)^2 \, dt\right] = -\lambda \mathbb{E}\left[\int_0^\infty e^{-\lambda t}(a_{it} - x)^2 \, dt\right] = -\int_0^\infty \frac{\lambda e^{-\lambda t}}{\alpha + \beta p_t} \, dt,$$

where we have normalized welfare by the intensity λ. The first equality follows from the random end time τ being geometrically distributed: $e^{-\lambda t}$ is the probability density that the economy ends at time $\tau \geqslant t$. The second equality follows from $a_{it} = \mathbb{E}[x \mid \mathcal{G}_{it}]$ so $\mathbb{E}[(a_{it} - x)^2 \mid \mathcal{G}_{it}] = 1/(P_t + p_t) = 1/(\alpha + \beta p_t)$, and an application of Fubini's Theorem.

Public information increases the total precision $\alpha + \beta p_t$ of agents' beliefs in the short run. But, as shown by Corollary 2, in the long run it results in unbounded losses of total precision, $\alpha + \beta p_t$. Because the welfare function is the present value of $-1/(\alpha + \beta p_t)$, it is natural to conjecture that as long as λ is close enough to 0 (i.e. the state is revealed in a long time, on average), public information reduces welfare.

Although intuitive, this result does not directly follow because, even when λ goes to zero, the trade-off between the short-term gain and the long-term loss remains non-trivial. Indeed, since $1/(\alpha + \beta p_t)$ converges to zero, the flow losses of increasing public information are vanishingly small. However, the next theorem shows that it is always possible to find a small enough λ to make a marginal release of information welfare-reducing:

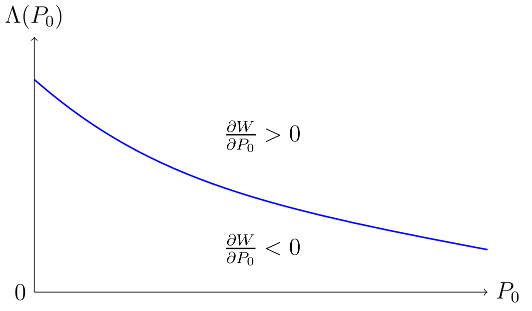

Fig. 1. This figure illustrates the boundary in the (P_0, λ) plane below which a marginal increase in information reduces welfare.

Theorem 2 *(Social value of public information). The welfare function, $W(P_0)$, is continuous, negative, and converges to zero as P_0 goes to infinity. Moreover, if we fix any $p_\varepsilon > 0$, then there exists a function $\Lambda(P_0)$, positive, decreasing and converging to zero as P_0 goes to infinity, such that $\partial W/\partial P_0 < 0$ for $0 < \lambda < \Lambda(P_0)$, and $\partial W/\partial P_0 > 0$ for $\lambda > \Lambda(P_0)$.*

The first part of the theorem has the intuitive implication that a sufficiently large public release of information is always welfare improving. Indeed, an infinite increase in precision would reveal the state of the world and would clearly improve welfare. By continuity, a sufficiently large release of public information must also improve welfare.

The second part of the theorem shows that a public release can have a negative welfare effect, as illustrated in Fig. 1. Namely, for any given P_0, a marginal increase in public information reduces welfare as long as λ lies below $\Lambda(P_0)$. In other words, for the negative welfare results, agents need to be patient enough (which in the model means that the state is revealed in a sufficiently long time). Alternatively, if λ lies above $\Lambda(P_0)$, then a marginal increase in public information increases welfare. This is intuitive: if agents are impatient enough, then they mostly care about the short-term gains of public information.

The theorem also implies that, if we fix some $\lambda < \Lambda(0)$, then welfare is a U-shaped function of P_0. This has important implications for optimal communication: if information can only be revealed at $t = 0$, then it is always optimal to reveal all or none of the information. By the same token, this also suggests that a *dynamic* communication is likely to involve a delay: indeed, public information becomes welfare improving when P_0 is large enough, i.e., after agents have learned sufficient information about the state of the world.[11]

A key assumption for the theorem is that $p_\varepsilon > 0$, that is, the private social learning channel is active. Indeed, when $p_\varepsilon = 0$ and only the public social learning channel is active, an increase in public information always increases welfare, given that it increases the precision of agents' information at all times. Thus, the channel of information diffusion is crucial for the social value of public information.

5. Optimal information diffusion

In the previous sections we have shown that the equilibrium allocation is independent of the impatience rate λ, and that, when a private learning channel is active, a release of public

[11] For some recent work on the optimal timing of public announcements see [28].

information always leads, in the long run, to a reduction of agents knowledge. In addition, if agents are sufficiently patient and if the public release is not too large, this long-run reduction in knowledge is welfare reducing. As we explained, these results arise because of a learning externality. In this section we study how the results differ when the externality is corrected.

Our notion of constrained efficiency is a natural one: it forces the social planner to respect the informational restrictions present in the economy, while also imposing that agents' actions remain linear functions of their observed signals. As we argue below, this notion of constrained efficiency has the benefit of tractability and, moreover, corresponds to the situation where ex-post linear Pigouvian taxes are the only policy tool available.

We show that, in the constrained efficient allocation, public releases are always welfare enhancing and the efficient allocation is affected by the impatience rate λ. Both results are in sharp contrast with their decentralized equilibrium counterparts. Surprisingly, however, public releases still reduce agents' knowledge in the long run.

5.1. Constrained efficiency

We let a planner choose an adapted action process a_i in order to maximize the *ex-ante* utility of a randomly chosen agent, subject to the learning technology. In setting up our planning problem, we follow [39] and restrict attention to actions that are convex combinations of the public and private forecasts defined in Section 3:

$$a_{it} = (1 - \gamma_t)\hat{X}_t + \gamma_t \hat{x}_{it}. \tag{18}$$

That is, instead of using the individually optimal weight $p_t/(P_t + p_t)$ on private information, the planner prescribes agents to use some other deterministic weight $\gamma_t \in [0, 1]$.

The set of controls specified in (18) is consistent with constrained efficiency: the planner is subject to the constraint that agents only learn from public and private signals of others' average actions. It is important to note that the specification (18) is restrictive: in principle, one could let the planner use actions which are arbitrary functions of the public and private endogenous signals received by the agents.[12] We make the restriction (18) for several reasons. First, as in [39] and subsequent work, it has the advantage of tractability: agents endogenous signals remain normally distributed and Kalman filtering methods continue to apply. Second, as it is well known, linear decision rules based on public and private forecasts arise in the equilibrium of a variety of rational expectations models with dispersed information. In many of these environments, simple *ex-post* Pigouvian taxes can be chosen to arbitrarily change the coefficients of the agents' linear decision rule. The associated optimal taxation problem becomes, then, equivalent to the planning problem considered in this section (see [1] and [5]).

Given decision rules given by (18), we can then solve for the learning dynamics exactly as we did before in the equilibrium analysis. Namely, given a weight γ_t, the precisions P_t and p_t of the public and private forecasts solve:

$$dP_t = P_\varepsilon \gamma_t^2 \, dt \quad \text{and} \quad dp_t = p_\varepsilon \gamma_t^2 \, dt. \tag{19}$$

As before, the speed of learning is controlled by the weight, γ_t, that agents put on their private information. Also, these laws of motion imply that P_t is an affine function of p_t: $P_t = \alpha + (\beta - 1) p_t$

[12] In the online Addendum, we allow for actions that are general affine functions of the public and the private forecast. However, we show that it is always optimal for the planner to use a convex combination.

where α and β take the same values as before, $\alpha = P_0 - p_0 P_\varepsilon / p_\varepsilon$ and $\beta = 1 + P_\varepsilon / p_\varepsilon$. This affine transformation simplifies the analysis because it allows to formulate the planner's problem in a one-dimensional state space. One should bear in mind, however, that the transformation imposes a restriction between the parameter α and the initial condition p_0: $p_0 = (\alpha - P_0)/(\beta - 1)$.

The utility flow of a representative agent is

$$-\mathbb{E}\big[(a_{it} - x)^2\big] = -\mathbb{E}\big[\big\{\gamma_t(\hat{x}_{it} - x) + (1 - \gamma_t)(\hat{X}_t - x)\big\}^2\big]$$
$$= -\left(\frac{\gamma_t^2}{p_t} + \frac{(1 - \gamma_t)^2}{\alpha + (\beta - 1)p_t}\right),$$

where the first line follows from plugging in (18), and the second line from noting that the private and public forecast errors are independent with respective variance $1/p_t$ and $1/(\alpha + (\beta - 1)p_t)$. Let us, then, define the following welfare function for $p \geqslant p_0$:

$$v(p, \gamma \mid \alpha) \equiv -\lambda \int_0^\infty e^{-\lambda t} \left(\frac{\gamma_t^2}{p_t} + \frac{(1 - \gamma_t)^2}{\alpha + (\beta - 1)p_t}\right) dt,$$

subject to $\dot{p}_t = p_\varepsilon \gamma_t^2$ with initial condition p. Let us denote by $V(p \mid \alpha)$ the supremum of $v(p, \gamma \mid \alpha)$ in the set of admissible controls γ. This function gives the planner's value along any socially optimal precision path starting at p_0.

Our first result states that, once the planner has corrected for the information externality, the welfare effects of public information are always positive:

Proposition 2. *The value function $V(p_0 \mid \alpha)$ is increasing in $\alpha = P_0 - P_\varepsilon / p_\varepsilon p_0$.*

This is intuitive: starting with a higher α, the planner can always choose the same sequence of weights on private information that he would have followed with a lower one. These same weights would then imply the same time path of private precision, p_t. But a higher α increases total precision, $\alpha + \beta p_t$, and thus welfare. The planner's ability to control the weight guarantees that the social value of public information is always positive.

We now proceed to analyze the impact of a public information release on precision dynamics, which requires a tighter characterization of the planner's optimal weight.

5.2. The optimal weight and comparative dynamics

In Appendix C, we show that the value function $V(p \mid \alpha)$ is differentiable with respect to p and solves the Hamilton–Jacobi–Bellman (HJB) equation:

$$\lambda V(p \mid \alpha) = \max_{\gamma \in [0,1]} \left\{ -\lambda \left(\frac{\gamma^2}{p} + \frac{(1 - \gamma)^2}{\alpha + (\beta - 1)p}\right) + p_\varepsilon \gamma^2 V'(p \mid \alpha) \right\}. \tag{20}$$

Taking the derivative with respect to the weight, γ, yields:

$$-2\lambda \left(\frac{\gamma}{p} + \frac{\gamma - 1}{\alpha + (\beta - 1)p}\right) + 2 p_\varepsilon \gamma V'(p \mid \alpha). \tag{21}$$

The second term of (21), $2p_\varepsilon\gamma V'(p \mid \alpha)$, corresponds to the marginal welfare gain of speeding up information dissemination.[13] The first term, on the other hand, corresponds to the marginal change in agents' forecast error. This term is equal to zero if the planner chooses the individually optimal weight $p/(\alpha + \beta p)$ of the decentralized equilibrium we studied before, since it minimizes the forecast error. Clearly, the planner has incentive to increase its weight above the individual optimum $p/(\alpha + \beta p)$: at the margin, it has no impact on agents' forecast error, while the marginal welfare gain of speeding up information dissemination is strictly positive, given that $V'(p \mid \alpha) > 0$. Finally, after plugging $\gamma = 1$ in the right-hand side of (20), it follows that $0 > \lambda V(p \mid \alpha) \geqslant -\lambda/p + p_\varepsilon V'(p \mid \alpha)$, which implies that (21) is strictly negative when evaluated at $\gamma = 1$, and thus the planner's optimal weight, $\gamma^*(p)$, is less than 1. It follows then that the planner's optimal weight satisfies

$$\frac{p}{\alpha + \beta p} < \gamma^\star(p) < 1. \tag{22}$$

Note that this implies that the precision increases faster in the social optimum than in the decentralized equilibrium, and hence, that it also converges to infinity: that is, full revelation is socially optimal.

Next, we analyze the asymptotic behavior of the planner's solution, which can be heuristically characterized from the HJB equation as follows (the formal proofs are in Appendix C). First, we note that, as $p \to \infty$, agents are almost fully informed about x, so the planner's marginal value of speeding up information becomes very small. Therefore, the planner chooses to reduce the forecast error instead, that is, the planner's control becomes very close to $1/\beta$, the limit of the individual optimum. Plugging this into the value function (20), we obtain that, as $p \to \infty$,

$$\lambda V(p \mid \alpha) \simeq -\frac{\lambda}{\beta p} + \frac{p_\varepsilon}{\beta^2} V'(p \mid \alpha), \tag{23}$$

where we used the approximation that $\alpha + (\beta - 1)p \simeq (\beta - 1)p$. But this means that, in (23), the second term is negligible relative to the first term – indeed, if $V(p \mid \alpha)$ is of order $1/p$, then its derivative is, heuristically, of order $1/p^2$. It follows then that

$$V(p \mid \alpha) \simeq -\frac{1}{\beta p} \quad \text{and} \quad V'(p \mid \alpha) \simeq \frac{1}{\beta p^2}.$$

Importantly, the initial conditions have, to a first-order approximation, no impact on the asymptotic level and derivative of the value function. Plugging the approximation $V'(p \mid \alpha) \simeq 1/(\beta p^2)$ into the first-order condition (21) delivers:

Theorem 3. *Suppose that $p_\varepsilon > 0$. In the planner's solution, as $p \to \infty$, the socially optimal weight admits the following asymptotic expansion*:

$$\gamma^\star(p) = \underbrace{\frac{1}{\beta} - \frac{\alpha}{\beta^2}\frac{1}{p}}_{\textit{individual optimum}} + \underbrace{\frac{1}{\lambda}\frac{(\beta - 1)p_\varepsilon}{\beta^2}\frac{1}{p}}_{\textit{planner's correction}} + O\left(\frac{1}{p^2}\right). \tag{24}$$

[13] Note that the value function is increasing: indeed if the planner applies the same optimal control but with a higher initial condition, his flow utility is higher at each time. This implies that $V'(p \mid \alpha) \geqslant 0$. In Appendix C we use the envelope condition to show that the derivative is, in fact, strictly positive.

The first two terms of the asymptotic expansion are the same as in the decentralized equilibrium. The correction that the planner makes shows up in the third term. As expected, this third term is positive: the planner would like agents to put more weight on their private information. Also, because the initial conditions have no impact on the first-order approximation of the value function, the planner's correction to the individual optimum is not affected by α.

The asymptotic expansion of $\gamma^\star(p)$ also reveals another feature of the solution: impatience reduces the planner's asymptotic correction to the weight. This is intuitive, as a more impatient planner discounts more heavily the future benefits of correcting the information externality.

To derive the asymptotic behavior of the total precision, recall that the socially optimal p_t solves the ODE $\dot{p}_t = p_\varepsilon^2 \gamma^*(p_t)^2$. Therefore, as in the decentralized equilibrium, the expansion of $\gamma^*(p)$ determines the terms of order t and $\log(t)$ in the asymptotic expansion of the total precision. The crucial observation is that the planner's correction term is independent of α, and hence, differences in initial public information will continue to have the same negative and unbounded $\log(t)$-impact on total precision. So, surprisingly, our main equilibrium comparative statics still holds: an increase in initial public information eventually leads to less total knowledge. Even though public information always increases *ex-ante* welfare, we find that *ex-post* welfare will eventually decrease.

Corollary 4. *Suppose that $p_\varepsilon > 0$. In the planner's solution, the total precision of agents beliefs at time t equals*:

$$P_t + p_t = (P_\varepsilon + p_\varepsilon)\left(\frac{p_\varepsilon}{P_\varepsilon + p_\varepsilon}\right)^2 t + 2\left(\frac{P_\varepsilon}{p_\varepsilon} p_0 - P_0 + \underbrace{\frac{1}{\lambda}\left(\frac{P_\varepsilon p_\varepsilon}{p_\varepsilon + P_\varepsilon}\right)}_{\textit{planner's correction}}\right)\log(t) + S_t$$

where S_t is bounded.

Lastly, we observe that Corollary 4 is a result of optimality, not feasibility: although, after a release of public information, the planner could choose a weight that increases knowledge at all times (as noted after Proposition 2), he chooses not to do so, and knowledge is reduced in the long run.

6. Conclusion

This paper analyzed how private information diffuses among a continuum of agents who learn from both public and private observations of each others' actions. We showed that when agents learn from a private channel, a release of public information at the beginning always slows the diffusion of private information in the economy, eventually reduces the amount known by everyone, and sometimes reduces welfare. We studied the optimal diffusion of information and showed that, relative to the private optimum, the planner corrects the learning externality by recommending agents to put more weight on their private information. We showed that the social value of public information after the planner corrects the externality is positive. However, the optimal response to a release of public information eventually leads to agents knowing strictly less, just as in the equilibrium.

Appendix A. Proof of Theorem 1

We prove the theorem using a guess and verify approach. Namely, we start by guessing that agents' signals, in Eq. (2) are observationally equivalent to the transformed signals of Eq. (6). We then derive the dynamics of agents' beliefs and of the average action. We then verify our guess, given this stochastic process for the average action.

A.1. Beliefs dynamics

Assume, then, that agents observe the pair of public and private signals given by the Stochastic Differential Equations (SDE) (6), with initial conditions (1). Let the public forecast, $\hat{X}_t$ be the expectation of x conditional on the public signal on the left-hand side of (6), given the common prior. Similarly, let public forecast, $\hat{x}_{it}$ be the expectation of x conditional on the private signal on the left-hand side of (6), given a fully diffuse prior. Denote by P_t and p_t the associated public and private precision. Then, we have:

Lemma 1 *(Dynamics of private and public forecasts). The private and public forecasts* $(\hat{x}_{it}, \hat{X}_t)$ *and the precisions* (p_t, P_t) *solve the system of SDE*

$$d\hat{x}_{it} = \frac{p_\varepsilon}{p_t} \frac{p_t^2}{(P_t + p_t)^2} \left[(x - \hat{x}_{it})\, dt + \frac{d\omega_{it}}{\sqrt{p_\varepsilon \frac{p_t}{P_t + p_t}}} \right], \tag{25}$$

$$d\hat{X}_t = \frac{P_\varepsilon}{P_t} \frac{p_t^2}{(P_t + p_t)^2} \left[(x - \hat{X}_t)\, dt + \frac{dW_t}{\sqrt{P_\varepsilon \frac{p_t}{P_t + p_t}}} \right], \tag{26}$$

$$dp_t = p_\varepsilon \left(\frac{p_t}{P_t + p_t} \right)^2 dt, \tag{27}$$

$$dP_t = P_\varepsilon \left(\frac{p_t}{P_t + p_t} \right)^2 dt, \tag{28}$$

with the initial conditions p_0, *and* P_0. *In addition, the above system can be integrated into*

$$\hat{x}_{it} = x + \frac{1}{p_t} \left[\sqrt{p_0}\, \omega_{i0} + \int_0^t \sqrt{p_\varepsilon} \frac{p_u}{P_u + p_u}\, d\omega_{iu} \right], \tag{29}$$

$$\hat{X}_t = x + \frac{1}{P_t} \left[\sqrt{P_0}\, W_0 + \int_0^t \sqrt{P_\varepsilon} \frac{p_u}{P_u + p_u}\, dW_u \right]. \tag{30}$$

Eqs. (25)–(28) follow from a direct application of one-dimensional continuous-time Kalman filtering formula (see, for instance, [31, pp. 85–105]). In order to derive Eq. (30), we multiply both sides of (26) by P_t. We find

$$P_t\, d\hat{X}_t = P_\varepsilon \frac{p_t^2}{(P_t + p_t)^2} \left[(x - \hat{X}_t)\, dt + \frac{dW_t}{\sqrt{P_\varepsilon \frac{p_t}{P_t + p_t}}} \right]$$
$$\Rightarrow \quad P_t\, d\hat{X}_t + dP_t(\hat{X}_t - x) = \sqrt{P_\varepsilon} \frac{p_t}{P_t + p_t}\, dW_t$$

$$\Rightarrow \quad d\big[P_t(\hat{X}_t - x)\big] = \sqrt{P_\varepsilon}\frac{p_t}{P_t + p_t}\,dW_t$$

$$\Rightarrow \quad P_t(\hat{X}_t - x) - P_0(\hat{X}_0 - x) = \int_0^t \sqrt{P_\varepsilon}\frac{p_u}{P_u + p_u}\,dW_u$$

$$\Rightarrow \quad \hat{X}_t = \frac{P_0}{P_t}\hat{X}_0 + \left(1 - \frac{P_0}{P_t}\right)x + \frac{1}{P_t}\int_0^t \sqrt{P_\varepsilon}\frac{p_u}{P_u + p_u}\,dW_u \tag{31}$$

where the second line follows from the fact that $dP_t = P_\varepsilon p_t^2/(P_t + p_t)^2\,dt$. Because P_t is a deterministic function of time it follows that $d[(\hat{X}_t - x)P_t] = d\hat{X}_t\,P_t + (\hat{X}_t - x)\,dP_t$, which implies the third line. The fourth line follows from integrating the third line from $u = 0$ to $u = t$, and the fifth line follows from rearranging. Now note that $\hat{X}_0$ and P_0 are the posterior mean and precision at time zero, after observing the public signal $Z_0 = x + W_0/\sqrt{P_0}$ and starting from the fully diffused common prior. Therefore $\hat{X}_0 = Z_0$. Eq. (30) then follows from plugging $\hat{X}_0 = Z_0 = x + W_0/\sqrt{P_0}$ back into (31). Eq. (29) follows from exactly the same algebraic manipulations.

Under our guess, the signals that generate the public and the private forecast, $\hat{X}_t$ and $\hat{x}_{it}$, are independent conditional on x. Hence, an agent's forecast conditional on all his information will be a linear combination of the public and his private forecasts, with weights given by their respective precisions:

$$\frac{P_t}{P_t + p_t}\hat{X}_t + \frac{p_t}{P_t + p_t}\hat{x}_{it}. \tag{32}$$

The precision of the agents belief conditional on all his information is $P_t + p_t$.

A.2. Verifying the guess

Let $\tilde{Z}_t$ and $\tilde{z}_{it}$ be the "transformed" signals of Eq. (6), with initial condition (1). We need to verify that the filtration generated by the transformed signals $(\tilde{Z}_t, \tilde{z}_{it})$ is the same as that generated by the signals of others' action, (Z_t, z_{it}). First, after plugging $A_t = P_t/(P_t + p_t)\hat{X}_t + p_t/(P_t + p_t)x$ into Eq. (2) we find that

$$dZ_t = \frac{P_t}{P_t + p_t}\hat{X}_t\,dt + \frac{p_t}{p_t + P_t}\,d\tilde{Z}_t,$$

$$dz_{it} = \frac{P_t}{P_t + p_t}\hat{X}_t\,dt + \frac{p_t}{p_t + P_t}\,d\tilde{z}_{it}.$$

Keeping in mind that $\hat{X}_t$, by construction, is adapted to the filtration generated by $(\tilde{Z}_t, \tilde{z}_{it})$, the above equation shows that the filtration generated by (Z_t, z_{it}) is included in the filtration generated by $(\tilde{Z}_t, \tilde{z}_{it})$. To show the reverse inclusion, first rearrange the above equation into

$$d\tilde{Z}_t = \frac{P_t + p_t}{p_t}\left(dZ_t - \frac{P_t}{P_t + p_t}\hat{X}_t\,dt\right), \tag{33}$$

$$d\tilde{z}_{it} = \frac{P_t + p_t}{p_t}\left(dz_{it} - \frac{P_t}{P_t + p_t}\hat{X}_t\,dt\right). \tag{34}$$

Now, we also know from Lemma 1 that

$$d\hat{X}_t = \frac{P_\varepsilon}{P_t}\left(\frac{p_t}{P_t+p_t}\right)^2\left[(x-\hat{X}_t)\,dt + \frac{dW_t}{\sqrt{P_\varepsilon \frac{p_t}{P_t+p_t}}}\right] = \frac{P_\varepsilon}{P_t}\left(\frac{p_t}{P_t+p_t}\right)^2(-\hat{X}_t\,dt + d\tilde{Z}_t).$$

After plugging Eq. (33) in the equation above and rearranging, we find:

$$d\hat{X}_t = \frac{P_\varepsilon}{P_t}\left(\frac{p_t}{P_t+p_t}\right)^2\left\{-\left(1+\frac{P_t}{p_t}\right)\hat{X}_t\,dt + \frac{P_t+p_t}{p_t}\,dZ_t\right\}.$$

Therefore, $\hat{X}_t$ is adapted to the filtration generated by Z. Together with (33), this means that $\tilde{Z}_t$ is adapted to the filtration generated by Z_t. Together with (33) and (34) this implies that $(\tilde{Z}_t, \tilde{z}_{it})$ is adapted to the filtration generated by (Z_t, z_{it}).

A.3. Closed form solution

This can be verified directly.

A.4. Proof of Corollary 1

Part (i): When $p_\varepsilon = 0$, this follows directly from the solution (10). When $p_\varepsilon > 0$, then from ODE (9) it is clear that p_t is strictly increasing, so it has a limit as $t \to \infty$. The limit must be infinite otherwise, as $t \to \infty$, the left-hand side of Eq. (11) would remain bounded, while the right-hand side would go to infinity.

Part (ii): When $p_\varepsilon = 0$ the ratio is $p_0/(p_0 + P_t)$ and is clearly decreasing towards zero. When $p_\varepsilon > 0$, we have

$$\frac{p_t}{P_t+p_t} = \frac{p_t}{\alpha+\beta p_t} = \frac{1}{\frac{\alpha}{p_t}+\beta}. \tag{35}$$

This implies that the ratio is converging towards $1/\beta = p_\varepsilon/(p_\varepsilon + P_\varepsilon)$. Since p_t is increasing, the ratio is strictly increasing if $\alpha > 0$, strictly decreasing if $\alpha < 0$, and constant if $\alpha = 0$.

Part (iii): When $p_\varepsilon = 0$, the result follows directly from the closed form solution (10). When $p_\varepsilon > 0$, we note that

$$\begin{aligned}\dot{p}_t &= \left(\frac{p_t}{\alpha+\beta p_t}\right)^2 p_\varepsilon = \left(\frac{\alpha}{p_t}+\beta\right)^{-2} p_\varepsilon \\ &= \frac{p_\varepsilon}{\beta^2} - \frac{2\alpha p_\varepsilon}{\beta^3}\frac{1}{p_t} + O\left(\frac{1}{p_t^2}\right),\end{aligned}$$

where $O(1/p^2)$ is the standard Landau notation for a function $f(p)$ such that $p^2 f(p)$ is bounded. An application of Lemma 2 below shows that

$$p_t = \frac{p_\varepsilon}{\beta^2}t - \frac{2\alpha}{\beta}\log(t) + C_t, \tag{36}$$

for some bounded function C_t. The expansion follows after noting that $P_t + p_t = \alpha + \beta p_t$, and defining $R_t \equiv \alpha + \beta C_t$. For the last part of the corollary, we need to show that R_t has a limit,

R_∞, and characterize the cross partial of R_∞ with respect to (P_0, p_ε). To that end we plug the expansion (36) back into Eq. (11):

$$H(p_t) = H(p_0) + p_\varepsilon t \quad \Leftrightarrow \quad C_t = p_0 + \frac{\alpha^2}{\beta^2}\left(\frac{1}{p_t} - \frac{1}{p_0}\right) - \frac{2\alpha}{\beta}\log\left(\frac{p_t}{p_0}\right).$$

Now, as $t \to \infty$, we have that $p_t/t \to p_\varepsilon/\beta^2$, and $\alpha^2/p_t \to 0$. Therefore, both C_t and R_t have limits, and the limit of R_t equals:

$$R_\infty \equiv \alpha + \beta \lim C_t = \alpha + 2\alpha\log(p_0) + \beta p_0 - \frac{\alpha^2}{\beta p_0} - 2\alpha\log\left(\frac{p_\varepsilon}{\beta^2}\right).$$

The cross partial of R_∞ can be shown to be:

$$\frac{\partial^2 R_\infty}{\partial P_0 \partial p_\varepsilon} = -\frac{2}{(p_\varepsilon + P_\varepsilon)^2}\left(p_\varepsilon + 5P_\varepsilon + \frac{P_0 P_\varepsilon}{p_0} + 3\frac{P_\varepsilon^2}{p_\varepsilon}\right) < 0.$$

Lemma 2 *(ODE asymptotics). Suppose*

$$\dot{x}_t = A + \frac{B}{x_t} + O\left(\frac{1}{x_t^2}\right) \tag{37}$$

and suppose that $x_t \to \infty$. Then

$$x_t = At + \frac{B}{A}\log(t) + C_t,$$

for some bounded function C_t.

Given that $x_t \to \infty$, there exists a finite T such that $A - \varepsilon \leqslant \dot{x}_t \leqslant A + \varepsilon$ for all $t > T$, where $0 < \varepsilon < A$. Then $(A-\varepsilon)(t-T) \leqslant x_t - x_T \leqslant (A+\varepsilon)(t-T)$, and thus $1/x_t = O(1/t)$. Plugging this back into the ODE (37) we have that $\dot{x}_t = A + O(1/t)$, which after integrating delivers that $x_t = At + O(\log(t))$. Taking the inverse of this last equation gives:

$$\frac{1}{x_t} = \frac{1}{At}\left[1 + O\left(\frac{\log(t)}{t}\right)\right]^{-1} = \frac{1}{At}\left[1 + O\left(\frac{\log(t)}{t}\right)\right] = \frac{1}{At} + O\left(\frac{\log(t)}{t^2}\right).$$

Plugging back into the ODE (37):

$$\dot{x}_t = A + \frac{B}{At} + O\left(\frac{\log(t)}{t^2}\right).$$

Now, the function $\log(t)/t^2$ is absolutely integrable, so the final result follows by integrating both sides of this equation.

A.5. Proof of Corollary 3

The first two terms of the asymptotic expansion (12) have a zero cross derivative with respect to (P_0, p_ε). Therefore, the limit of the double difference, $\Delta_t[\hat{p}_\varepsilon] - \Delta_t[\tilde{p}_\varepsilon]$, is the limit of the double difference of R_t, the third term of the expansion. The limit of the double difference of R_t is the double difference of R_∞, which has the same sign as its cross derivative with respect to (P_0, p_ε), and is thus negative.

Appendix B. Proof of Theorem 2

In the first part of the theorem, continuity follows because solutions of ODE are continuous with respect to their initial conditions (see, e.g., Theorem 2.10 in [36]). Negativity follows because knowledge is finite at all times. Lastly, since knowledge is always rising over time, we have that $p_t + P_t \geqslant p_0 + P_0 = \alpha + \beta P_0$, and so $0 > W \geqslant -1/(\alpha + \beta P_0)$, implying in turns that welfare goes to zero as P_0 goes to infinity. Now, let us turn to the second part of the theorem. Using (9) the welfare function can be rewritten as

$$W = -\int_0^\infty \left(\frac{\alpha + \beta p_t}{p_t^2}\right)\frac{\lambda}{p_\varepsilon} e^{-\lambda t}\,(\dot{p}_t\,dt) = -\int_{p_0}^\infty \left(\frac{\alpha + \beta p}{p^2}\right)\frac{\lambda}{p_\varepsilon} e^{-\frac{\lambda}{p_\varepsilon}[H(p)-H(p_0)]}\,dp, \tag{38}$$

where the last equality follows from $p_\varepsilon t = H(p_t) - H(p_0)$ for $H(p) \equiv 2\alpha\beta \log p + \beta^2 p - \alpha^2/p$ and, given that p_t monotonically approaches infinity through time, a change in the integrating variable from t to p. From the above equation, one sees that the welfare function W depends on the initial precision P_0 of public information only through $\alpha = P_0 - (P_\varepsilon/p_\varepsilon)p_0$. Clearly, this means that welfare is decreasing in P_0 if and only if it is decreasing in α. We thus calculate $\partial W/\partial\alpha$ and show:

Lemma 3 *(Preliminary). The partial derivative of the welfare function (38) with respect to α is*:

$$\frac{\partial W}{\partial \alpha} = \frac{\lambda}{p_\varepsilon}\int_{p_0}^\infty \Phi(p)\frac{\lambda}{p_\varepsilon}\frac{\partial H}{\partial p}(p) e^{-\frac{\lambda}{p_\varepsilon}[H(p)-H(p_0)]}\,dp, \quad \textit{where} \tag{39}$$

$$\Phi(p) \equiv -\left(\frac{1}{p_0} - \frac{1}{p}\right) + \frac{2}{\alpha + \beta p}\left[\beta \log\left(\frac{p}{p_0}\right) + \alpha\left(\frac{1}{p_0} - \frac{1}{p}\right)\right]. \tag{40}$$

In the above lemma and in all what follows, we simplify notations by suppressing the explicit dependence of functions (H, Φ, etc.) on α. To derive the formula of the lemma, note:

$$\begin{aligned}\frac{\partial W}{\partial \alpha} &= -\int_{p_0}^\infty \left\{\frac{1}{p^2} - \frac{\alpha + \beta p}{p^2}\frac{\lambda}{p_\varepsilon}\frac{\partial}{\partial \alpha}\big[H(p) - H(p_0)\big]\right\}\frac{\lambda}{p_\varepsilon} e^{-\frac{\lambda}{p_\varepsilon}[H(p)-H(p_0)]}\,dp \\ &= \int_{p_0}^\infty \left\{-\frac{1}{p^2} + 2\frac{\alpha + \beta p}{p^2}\frac{\lambda}{p_\varepsilon}\left[\beta \log\left(\frac{p}{p_0}\right)\right.\right. \\ &\qquad \left.\left. + \alpha\left(\frac{1}{p_0} - \frac{1}{p}\right)\right]\right\}\frac{\lambda}{p_\varepsilon} e^{-\frac{\lambda}{p_\varepsilon}[H(p)-H(p_0)]}\,dp.\end{aligned} \tag{41}$$

Next, integrate by parts the first term of the integral (41), noting that $1/p^2 = d/dp(1/p_0 - 1/p)$:

$$-\int_{p_0}^\infty \frac{1}{p^2}\frac{\lambda}{p_\varepsilon} e^{-\frac{\lambda}{p_\varepsilon}[H(p)-H(p_0)]}\,dp = -\frac{\lambda}{p_\varepsilon}\int_{p_0}^\infty \left(\frac{1}{p_0} - \frac{1}{p}\right)\frac{\partial H}{\partial p}(p)\frac{\lambda}{p_\varepsilon} e^{-\frac{\lambda}{p_\varepsilon}[H(p)-H(p_0)]}\,dp,$$

because $H(p) \to \infty$ as $p \to \infty$. Finally, manipulate the second term of the integral (41) as follows:

$$2\int_{p_0}^{\infty}\frac{\alpha+\beta p}{p^2}\frac{\lambda}{p_\varepsilon}\left[\alpha\left(\frac{1}{p_0}-\frac{1}{p}\right)+\beta\log\left(\frac{p}{p_0}\right)\right]\frac{\lambda}{p_\varepsilon}e^{-\frac{\lambda}{p_\varepsilon}[H(p)-H(p_0)]}\,dp$$
$$=2\int_{p_0}^{\infty}\frac{\frac{\partial H}{\partial p}(p)}{\frac{\partial H}{\partial p}(p)}\frac{\alpha+\beta p}{p^2}\frac{\lambda}{p_\varepsilon}\left[\alpha\left(\frac{1}{p_0}-\frac{1}{p}\right)+\beta\log\left(\frac{p}{p_0}\right)\right]\frac{\lambda}{p_\varepsilon}e^{-\frac{\lambda}{p_\varepsilon}[H(p)-H(p_0)]}\,dp$$
$$=2\frac{\lambda}{p_\varepsilon}\int_{p_0}^{\infty}\frac{1}{\alpha+\beta p}\left[\alpha\left(\frac{1}{p_0}-\frac{1}{p}\right)+\beta\log\left(\frac{p}{p_0}\right)\right]\frac{\lambda}{p_\varepsilon}\frac{\partial H}{\partial p}(p)e^{-\frac{\lambda}{p_\varepsilon}[H(p)-H(p_0)]}\,dp,$$

where the third line follows from the fact that $\partial H/\partial p = [(\alpha+\beta p)/p]^2$. Collecting terms leads to the formula of the lemma.

Lemma 4 *(Behavior for small and large λ). For any $P_0 \geqslant 0$, $\partial W/\partial\alpha < 0$ for λ small enough, and $\partial W/\partial\alpha > 0$ for λ large enough.*

For the first part of the lemma, note that since $\Phi(p) \to -1/p_0$ as $p \to \infty$, there exists some p^* such that $\Phi(p) < -1/(2p_0)$ for all $p > p^*$. Letting $M^* = \sup_{p\in[p_0,p^*]}\Phi(p)$, Eq. (39) implies that

$$\frac{\partial W}{\partial\alpha}=\frac{\lambda}{p_\varepsilon}\int_{p_0}^{p^*}\Phi(p)\frac{\lambda}{p_\varepsilon}\frac{\partial H}{\partial p}(p)e^{-\frac{\lambda}{p_\varepsilon}[H(p)-H(p_0)]}\,dp$$
$$+\frac{\lambda}{p_\varepsilon}\int_{p^*}^{\infty}\Phi(p,\alpha)\frac{\lambda}{p_\varepsilon}\frac{\partial H}{\partial p}(p)e^{-\frac{\lambda}{p_\varepsilon}[H(p)-H(p_0)]}\,dp$$
$$\leqslant\frac{\lambda}{p_\varepsilon}\left\{M^*\int_{p_0}^{p^*}\frac{\lambda}{p_\varepsilon}\frac{\partial H}{\partial p}(p)e^{-\frac{\lambda}{p_\varepsilon}[H(p)-H(p_0)]}\,dp\right.$$
$$\left.-\frac{1}{2p_0}\int_{p^*}^{\infty}\frac{\lambda}{p_\varepsilon}\frac{\partial H}{\partial p}(p)e^{-\frac{\lambda}{p_\varepsilon}[H(p)-H(p_0)]}\,dp\right\}$$
$$=\frac{\lambda}{p_\varepsilon}\left\{M^*\left(1-e^{-\frac{\lambda}{p_\varepsilon}[H(p^*)-H(p_0)]}\right)-\frac{1}{2p_0}e^{-\frac{\lambda}{p_\varepsilon}[H(p^*)-H(p_0)]}\right\}.$$

The term inside the curly brackets is negative as long as λ is small enough.

For the second part of the lemma, we proceed with another integration by parts in (39), keeping in mind that $\Phi(p_0)=0$, $\Phi(p)$ is bounded, and $H(p)\to 0$ as $p\to\infty$.

$$\frac{\partial W}{\partial\alpha}=\frac{\lambda}{p_\varepsilon}\int_{p_0}^{\infty}\frac{\partial\Phi}{\partial p}(p)e^{-\frac{\lambda}{p_\varepsilon}[H(p)-H(p_0)]}\,dp=\int_{p_0}^{\infty}\frac{\frac{\partial H}{\partial p}(p)}{\frac{\partial H}{\partial p}(p)}\frac{\partial\Phi}{\partial p}(p)\frac{\lambda}{p_\varepsilon}e^{-\frac{\lambda}{p_\varepsilon}[H(p)-H(p_0)]}\,dp$$
$$=\int_{p_0}^{\infty}\Psi(p)\frac{\lambda}{p_\varepsilon}\frac{\partial H}{\partial p}(p)e^{-\frac{\lambda}{p_\varepsilon}[H(p)-H(p_0)]}\,dp$$

where

$$\Psi(p) \equiv \frac{1}{(\alpha+\beta p)^2} - \frac{2\beta p^2}{(\alpha+\beta p)^4}\left\{\beta \log\left(\frac{p}{p_0}\right) + \alpha\left(\frac{1}{p_0} - \frac{1}{p}\right)\right\}.$$

Note that $p \mapsto \Psi(p)$ is continuous, $\Psi(p_0) = 1/(\alpha+\beta p_0)^2 > 0$, and that $p \mapsto 1 - e^{-\frac{\lambda}{p_\varepsilon}[H(p)-H(p_0)]}$ converges to the Dirac measure at p_0 as $\lambda \to \infty$. It thus follows that

$$\lim_{\lambda\to\infty} \frac{\partial W}{\partial \alpha} = \Psi(p_0) > 0$$

which establishes the second part of the lemma.

Lemma 5 *(Properties of $\Phi(p)$). The function $\Phi(p)$ has the following properties*: (i) *it is bounded,* (ii) *it tends to $-1/p_0$ as p goes to infinity,* (iii) *it is positive for all $p \in (p_0, \hat{p})$ and negative for all $p \in (\hat{p}, \infty)$, for some $\hat{p} > p_0$,* (iv) *it is increasing in α.*

The first two properties, (i) and (ii), follow from the definition of $\Phi(p)$. For the third one, note that

$$\frac{\partial \Phi}{\partial p}(p) = \frac{N(p)}{p^2(\alpha+\beta p)^2} \quad \text{where}$$

$$N(p) = (\alpha+\beta p)^2 - 2\beta p^2\left\{\beta \log\left(\frac{p}{p_0}\right) + \alpha\left(\frac{1}{p_0} - \frac{1}{p}\right)\right\}.$$

Clearly, $N(p_0) = (\alpha+\beta p_0)^2 > 0$, and $\lim_{p\to\infty} N(p) = -\infty$. Moreover, for $p > p_0$:

$$\frac{\partial N}{\partial p}(p) = -4\beta p\left\{\beta \log\left(\frac{p}{p_0}\right) + \alpha\left(\frac{1}{p_0} - \frac{1}{p}\right)\right\} < 0,$$

where the last inequality follows because the term in curly bracket is zero at $p = p_0$ and is increasing, as its derivative is equal to $(\alpha+\beta p)/p^2 > 0$. So $N(p)$ changes sign only once. Taken together, the analysis of $N(p)$ implies that $\Phi(p)$ is a hump-shaped function of p with a peak at a value strictly higher than p_0. Since $\Phi(p)$ starts at 0 when evaluated at p_0 and has a negative limit as $p \to \infty$, part (iii) of the lemma follows.

Lastly, for the fourth result of the lemma note that

$$\frac{\partial \Phi}{\partial \alpha}(p) = \frac{2\beta}{(\alpha+\beta p)^2}\left\{\left(\frac{p}{p_0} - 1\right) - \log\left(\frac{p}{p_0}\right)\right\} \geqslant 0,$$

with an equality for $p = p_0$ only. This is because $\log(x)$ is concave and so $\log(x) \leqslant x - 1$.

Finally, we show:

Lemma 6. *Given P_0, there exists a unique $\lambda = \Lambda(P_0)$ solving $\partial W/\partial\alpha = 0$. Moreover, the function $\Lambda(P_0)$ decreasing and converges to zero as P_0 goes to infinity.*

Given that all functions under consideration are continuous in all their arguments, Lemma 4 implies that, given $P_0 \geqslant 0$, there exists some $\lambda = \Lambda(P_0)$ solving $\partial W/\partial\alpha = 0$. To show uniqueness, note that solving $\partial W/\partial\alpha = 0$ for $\lambda > 0$ is equivalent to solving:

$$\Delta(\lambda) = 0, \quad \text{where } \Delta(\lambda) = \int_{p_0}^{\infty} \Phi(p) \frac{\partial H}{\partial p}(p) e^{-\frac{\lambda}{p_\varepsilon}[H(p)-H(p_0)]}\, dp.$$

Now differentiate $\Delta(\lambda)$ and evaluate at $\lambda = \Lambda(P_0)$:

$$\Delta'\big(\Lambda(P_0)\big) = -\frac{1}{p_\varepsilon}\int_{p_0}^{\infty}\big[H(p) - H(p_0)\big]\Phi(p)\frac{\partial H}{\partial p}(p)e^{-\frac{\Lambda(P_0)}{p_\varepsilon}[H(p)-H(p_0)]}\,dp > 0,$$

where the inequality follows from an application of Lemma 7 below, with

$$x = p, \quad k(p) = -\big[H(p) - H(p_0)\big] \quad \text{and} \quad f(p) = \Phi(p)\frac{\partial H}{\partial p}(p)e^{-\frac{\Lambda(P_0)}{p_\varepsilon}[H(p)-H(p_0)]}.$$

Note that $\Delta'(\Lambda(P_0)) > 0$ also implies that, at $\lambda = \Lambda(P_0)$, $\partial^2 W/\partial\alpha\partial\lambda > 0$. Therefore, to show that $\Lambda(P_0)$ is decreasing, the Implicit Function Theorem implies that it is enough to show that $\partial^2 W/\partial\alpha^2 > 0$ at $\lambda = \Lambda(P_0)$. To establish this, note that

$$\frac{\partial^2 W}{\partial\alpha^2} = \frac{\lambda}{p_\varepsilon}\int_{p_0}^{\infty}\frac{\partial\Phi}{\partial\alpha}(p)g(p)\,dp + \frac{\lambda}{p_\varepsilon}\int_{p_0}^{\infty}\frac{\frac{\partial g}{\partial\alpha}(p)}{g(p)}\Phi(p)g(p)\,dp, \tag{42}$$

where

$$g(p) \equiv \frac{\lambda}{p_\varepsilon}\frac{\partial H}{\partial p}(p)e^{-\frac{\lambda}{p_\varepsilon}[H(p)-H(p_0)]}.$$

The first term of $\partial^2 W/\partial\alpha^2$ is positive since, by Lemma 5, $\partial\Phi/\partial\alpha > 0$ for $p > p_0$. For the second term note that

$$d\left[\frac{\frac{\partial g}{\partial\alpha}(p)}{g(p)}\right]\Big/dp = -2\frac{\beta}{(\alpha+\beta p)^2} - 2\frac{(\alpha+\beta p)\lambda}{p^2 p_\varepsilon} < 0.$$

Therefore, $[\partial g/\partial\alpha](p)/g(p)$ is a decreasing function of p and we can apply Lemma 7 to (42) with:

$$k(p) = \frac{\frac{\partial g}{\partial\alpha}(p)}{g(p)} \quad \text{and} \quad f(p) = \Phi(p)g(p), \quad \text{evaluated at } \lambda = \Lambda(P_0).$$

The last result to establish is that $\Lambda(P_0) \to 0$ as $P_0 \to \infty$. To see why, note that $\Lambda(P_0)$ must have a nonnegative limit Λ^*. If $\Lambda^* > 0$, then consider some $\lambda \in (0, \Lambda^*]$. Then, the above results show that, given such λ, $\partial W/\partial P_0 < 0$ for all P_0, which is impossible given that $W \to 0$ as $P_0 \to \infty$.

Finally, we now state and prove the following result which was used above:

Lemma 7. *Consider a function $f : [\underline{x}, \overline{x}] \mapsto \mathbb{R}$, where $0 \leqslant \underline{x} < \overline{x} \leqslant \infty$, and suppose that the following holds: there exists an $\hat{x}$ such that $f(x) > 0$ for $x \in (\underline{x}, \hat{x})$; $f(x) < 0$ for $x \in (\hat{x}, \overline{x})$; $\int_{\underline{x}}^{\overline{x}} f(x)\,dx = 0$. Then for any decreasing function $k : [\underline{x}, \overline{x}] \mapsto \mathbb{R}$, $\int_{\underline{x}}^{\overline{x}} f(x)k(x)\,dx > 0$.*

To prove the above, note that

$$\int_{\underline{x}}^{\overline{x}} f(x)k(x)\,dx = \int_{\underline{x}}^{\overline{x}} f(x)k(x)\,dx - \underbrace{\int_{\underline{x}}^{\overline{x}} f(x)k(\hat{x})\,dx}_{=0}$$

$$= \int_{\underline{x}}^{\hat{x}} f(x)\big(k(x) - k(\hat{x})\big)\,dx + \int_{\hat{x}}^{\overline{x}} f(x)\big(k(x) - k(\hat{x})\big)\,dx.$$

The right-hand side is positive since, for $x \in (\underline{x}, \hat{x})$, $f(x) > 0$ and $k(x) > k(\hat{x})$, while for $x \in (\hat{x}, \overline{x})$ we have $f(x) < 0$ and $k(x) < k(\hat{x})$.

Appendix C. Proofs omitted in Section 5

In all what follows, we keep the parameters (α, β) the same, and we omit the dependence of functions on the parameters (α, β). In setting up and solving the planner's problem, we also impose that the control γ_t lies in a compact set, which we take to be the interval [0, 1]. However, as will become clear, this constraint is never binding.

C.1. Lipschitz continuity

We start by establishing that the value function is Lipschitz continuous:

Lemma 8. *For all $p' > p$:*

$$V(p') - V(p) \leqslant (p' - p)\left(\frac{1}{p^2} + \frac{\beta - 1}{(\alpha + (\beta - 1)p)^2}\right).$$

Indeed, consider two initial conditions $p' > p$. By definition of the value function, for any $\varepsilon > 0$ there is some control γ' such that $v(p', \gamma') \geqslant V(p') - \varepsilon$, and $V(p) \geqslant v(p, \gamma')$. Combining these two inequalities, we obtain:

$$\begin{aligned}
V(p') - V(p) &\leqslant v(p', \gamma') - v(p, \gamma') + \varepsilon \\
&= \int_0^\infty \lambda e^{-\lambda_t}\left[\gamma_t'^2\left(\frac{1}{p_t} - \frac{1}{p_t'}\right)\right. \\
&\quad \left. + (1 - \gamma_t')^2\left(\frac{1}{\alpha + (\beta - 1)p_t} - \frac{1}{\alpha + (\beta - 1)p_t'}\right)\right] dt + \varepsilon \\
&= \int_0^\infty \lambda e^{-\lambda_t}\left[\frac{\gamma_t'^2(p_t' - p_t)}{p_t' p_t} + \frac{(1 - \gamma_t')^2(\beta - 1)(p_t' - p_t)}{(\alpha + (\beta - 1)p_t)(\alpha + (\beta - 1)p_t')}\right] dt + \varepsilon \\
&\leqslant \int_0^\infty \lambda e^{-\lambda_t}\left[\frac{(p' - p)}{p' p} + \frac{(\beta - 1)(p' - p)}{(\alpha + (\beta - 1)p)(\alpha + (\beta - 1)p')}\right] dt \\
&\leqslant (p' - p)\left(\frac{1}{p^2} + \frac{\beta - 1}{(\alpha + (\beta - 1)p)^2}\right) + \varepsilon,
\end{aligned}$$

where the second to last line follows since $\gamma_t' \in [0, 1]$, and because $\dot{p}_t' = \dot{p}_t = p_\varepsilon(\gamma_t')^2$ implies that $p_t' - p_t = p' - p$ and $p_t' > p_t \geqslant p$ for all times. The result obtains by letting $\varepsilon \to 0$.

C.2. Dynamic programming

Let us define

$$V'(p^+) \equiv \limsup_{p' \to p^+} \frac{V(p') - V(p)}{p' - p}.$$

Since Lipschitz continuity implies absolute continuity, we know from [34, Theorem 7.20] that the value function is differentiable almost everywhere: therefore, $V'(p^+)$ coincides is the classical derivative almost everywhere and

$$V(p') = V(p) + \int_p^{p'} V'(x^+)\,dx, \tag{43}$$

for all $p' > p$. Equipped with this definition of the derivative, the results of Chapter III in [11] allow us to state (see the Addendum for a step-by-step explanation):

Lemma 9. *The value function solves the HJB equation (20) shown in the text for all p. Let $\gamma^*(p)$ achieve the maximum in (20)*

$$\gamma^*(p) \equiv \min\left\{1, \frac{p}{\alpha + \beta p - (\alpha + (\beta - 1)p)pp_\varepsilon V'(p^+)/\lambda}\right\}. \tag{44}$$

Then $\gamma_t^* = \gamma^*(p_t)$ where $p_t = p + \int_0^t p_\varepsilon (\gamma_t^*)^2\,dt$ is an optimal control for the planner's problem with initial condition $p > p_0$.

As explained in the text, one easily shows that $\gamma^*(p) < 1$. Together with the finding that $\gamma^*(p) \geqslant p/(\alpha + \beta p)$, this implies that the constraint $\gamma^*(p) \in [0, 1]$ is never binding. Next, plugging $\gamma^*(p)$ back into the HJB, one finds that

$$V(p) = -\frac{1 - \gamma^*(p)}{\alpha + (\beta - 1)p}. \tag{45}$$

But $V(p)$ is continuous: therefore, $\gamma^*(p)$ and, by implication, $V'(p^+)$, are also continuous. Then, it follows from (43) that $V(p)$ is, in fact, continuously differentiable.

Plugging the expression for $\gamma^*(p)$ as a function of $V'(p)$ into (45), one finds that $V(p)$ solves the Ordinary Differential Equation (ODE):

$$V'(p) = \frac{\lambda}{p_\varepsilon} \frac{V(p)(\alpha + \beta p) + 1}{p[V(p)(\alpha + (\beta - 1)p) + 1]}. \tag{46}$$

Clearly, since the right-hand side is continuous, so is the left-hand side, and so on, implying that $V(p)$ admits continuous derivatives at all orders. This allows us to use the envelope condition in (20), and obtain:

$$V'(p) = \lambda \frac{\gamma^*(p)^2}{p^2} + \lambda \frac{(\beta - 1)(1 - \gamma^*(p))^2}{(\alpha + (\beta - 1)p)^2} + p_\varepsilon \gamma^*(p)^2 V''(p).$$

Integrating this up along the socially optimal path of precision starting at p, and using that $\lim_{t\to\infty} e^{-\lambda t} V'(p_t) = 0$ (which follows from the Lipschitz bound), gives:

Lemma 10. *The value function is continuously differentiable and its derivative is*:

$$V'(p) = \int_0^\infty \left(\frac{(\gamma_t^*)^2}{(p_t^*)^2} + \frac{(\beta - 1)(1 - \gamma_t^*)^2}{(\alpha + (\beta - 1)p_t^*)^2}\right)\lambda e^{-\lambda t}\,dt, \tag{47}$$

where p_t^* and γ_t^* are, respectively, the socially optimal path of precision and the socially optimal weight starting at p. In particular, $V'(p) > 0$ for all p.

Plugging that $\gamma_t^* \geqslant p_t^*/(\alpha+\beta p_t^*) > 0$ in the right-hand side of (47) implies that the derivative is strictly positive.

C.3. Preliminary asymptotic results

Our first preliminary result is

Lemma 11. *As $p \to \infty$, $\gamma^*(p) = 1/\beta + O(1/p)$.*

To prove this result first note that, since $\gamma(p) \in [0,1]$, it follows from (47) that

$$V'(p) \leqslant \frac{1}{p^2} + \frac{\beta-1}{(\alpha+(\beta-1)p)^2}.$$

Now plugging this into the (44), this gives

$$\begin{aligned}
&\frac{p}{\alpha+\beta p} \leqslant \gamma^*(p) \leqslant \frac{p}{\alpha+\beta p - \frac{p_\varepsilon}{\lambda}\left(\frac{\alpha+(\beta-1)p}{p} + \frac{(\beta-1)p}{\alpha+(\beta-1)p}\right)} \\
\Leftrightarrow \quad & \left[\frac{\alpha}{p}+\beta\right]^{-1} \leqslant \gamma^*(p) \leqslant \left[\frac{\alpha}{p}+\beta - \frac{p_\varepsilon}{\lambda p}\left(\frac{\alpha+(\beta-1)p}{p} + \frac{(\beta-1)p}{\alpha+(\beta-1)p}\right)\right]^{-1} \\
\Leftrightarrow \quad & \left[\beta + O\left(\frac{1}{p}\right)\right]^{-1} \leqslant \gamma^*(p) \leqslant \left[\beta + O\left(\frac{1}{p}\right)\right]^{-1},
\end{aligned}$$

and the result follows.

Next, we prove:

Lemma 12. *As $p \to \infty$, $p^2 V'(p) = 1/\beta + O(1/p)$.*

To see this, multiply both sides of (47) by p^2 and obtain

$$p^2 V'(p) = \int_0^\infty \frac{p^2 \gamma(p_t^*)^2}{(p_t^*)^2} \lambda e^{-\lambda t}\, dt + \int_0^\infty \frac{p^2(\beta-1)(1-\gamma(p_t^*))^2}{(\alpha+(\beta-1)p_t^*)^2} \lambda e^{-\lambda t}\, dt. \tag{48}$$

We start by showing that the first integral is $1/\beta^2 + O(1/p)$. Indeed, since we know from Lemma 11 that $\gamma(p) = 1/\beta + O(1/p)$, it follows that

$$\gamma\left(p_t^*\right)^2 = \frac{1}{\beta^2} + O\left(\frac{1}{p_t^*}\right) = \frac{1}{\beta^2} + O\left(\frac{1}{p}\right)$$

where the second equality follows because $p \leqslant p_t$ for all t. Substituting this in Eq. (48) and subtracting $1/\beta^2$, we find:

$$\begin{aligned}
\left|\int_0^\infty \frac{p\gamma(p_t^*)^2}{(p_t^*)^2} \lambda e^{-\lambda t}\, dt - \frac{1}{\beta^2}\right| &= \left|\int_0^\infty \left(\frac{p^2(\frac{1}{\beta^2} + O(\frac{1}{p}))}{(p_t^*)^2} - \frac{1}{\beta^2}\right) \lambda e^{-\lambda t}\, dt\right| \\
&= \left|\int_0^\infty \frac{p^2 + O(p) - (p + (p_t^* - p))^2}{\beta^2 (p + (p_t^* - p))^2} \lambda e^{-\lambda t}\, dt\right|
\end{aligned}$$

$$= \left| \int_0^\infty \frac{O(p) - 2p(p_t^* - p) + (p_t^* - p)^2}{\beta^2 (p + (p_t^* - p))^2} \lambda e^{-\lambda t}\, dt \right|$$

$$\leqslant \int_0^\infty \frac{O(p) + 2p(p_t^* - p) + (p_t^* - p)^2}{\beta^2 p^2} \lambda e^{-\lambda t}\, dt$$

$$\leqslant \int_0^\infty \left(O\left(\frac{1}{p}\right) + \frac{2p_\varepsilon t}{\beta^2 p} + \frac{(p_\varepsilon t)^2}{\beta^2 p^2} \right) \lambda e^{-\lambda t}\, dt = O\left(\frac{1}{p}\right)$$

where the last inequality follows because $p_t^* - p = \int_0^t p_\varepsilon (\gamma_s^*)^2\, ds$ and $\gamma_s^* \in [0, 1]$. Following the same steps one shows that the second integral in (48) is $(\beta - 1)/\beta^2 + O(1/p)$, and the result follows.

C.4. Proof of Theorem 3

Combining Eq. (44) and Lemma 12, we obtain:

$$\begin{aligned}
\gamma(p) &= \frac{p}{(\alpha + \beta p - (\alpha + (\beta - 1)p) p \frac{p_\varepsilon}{\lambda} V'(p^+ \mid \alpha))} \\
&= \left[\frac{\alpha}{p} + \beta - \frac{1}{p} \left((\beta - 1) + \frac{\alpha}{p} \right) \frac{p_\varepsilon}{\lambda} p^2 V'(p \mid \alpha) \right]^{-1} \\
&= \left[\frac{\alpha}{p} + \beta - \frac{1}{p} \left((\beta - 1) + \frac{\alpha}{p} \right) \frac{p_\varepsilon}{\lambda} \left(\frac{1}{\beta} + O\left(\frac{1}{p}\right) \right) \right]^{-1} \\
&= \left[\beta + \frac{1}{p} \left(\alpha - \frac{p_\varepsilon}{\lambda} \frac{\beta - 1}{\beta} \right) + O\left(\frac{1}{p^2}\right) \right]^{-1} \\
&= \frac{1}{\beta} \left[1 - \frac{1}{p} \left(\frac{\alpha}{\beta} - \frac{1}{\lambda} \frac{(\beta - 1) p_\varepsilon}{\beta^2} \right) \right] + O\left(\frac{1}{p^2}\right),
\end{aligned}$$

as claimed.

C.5. Proof Corollary 4

Theorem 3 implies that

$$\gamma^*(p)^2 = \frac{1}{\beta^2} \left[1 - \frac{2}{p} \left(\frac{\alpha}{\beta} - \frac{1}{\lambda} \frac{(\beta - 1) p_\varepsilon}{\beta^2} \right) \right] + O\left(\frac{1}{p^2}\right).$$

Given that $\dot{p}_t^* = p_\varepsilon \gamma(p_t^*)^2$, this implies:

$$\dot{p}_t^* = \frac{p_\varepsilon}{\beta^2} - \frac{1}{p_t^*} \left(\frac{2 p_\varepsilon \alpha}{\beta^3} - \frac{2 p_\varepsilon}{\lambda} \frac{\beta - 1}{\beta^4} \right) + O\left(\frac{1}{(p_t^*)^2}\right),$$

and the result follows from Lemma 2.

Appendix D. Supplementary material

Supplementary material related to this article can be found online at doi:10.1016/j.jet.2012.02.001.

References

[1] Manuel Amador, Pierre-Olivier Weill, Learning by matching, 2006.
[2] Manuel Amador, Pierre-Olivier Weill, Addendum to "Learning from private and public observations of others' actions", Working Paper, Stanford and UCLA, 2007.
[3] Manuel Amador, Pierre-Olivier Weill, Learning from prices: Public communication and welfare, J. Polit. Economy 118 (2010) 866–907.
[4] George-Marios Angeletos, Alessandro Pavan, Efficient use of information and social value of information, Econometrica 75 (2007) 1103–1142.
[5] George-Marios Angeletos, Alessandro Pavan, Policy with dispersed information, J. Europ. Econ. Assoc. 7 (1) (2009) 60.
[6] Luis Araujo, Andrei Shevchenko, Price dispersion, information and learning, J. Monet. Econ. 53 (2006) 1197–1223.
[7] Luis Araujo, Braz Camargo, Information, learning, and the stability of fiat money, J. Monet. Econ. 53 (2006) 1571–1591.
[8] Venkatesh Bala, Sanjeev Goyal, Learning from neighbours, Rev. Econ. Stud. 65 (1998) 595–621.
[9] Abhijit Banerjee, Drew Fudenberg, Word-of-mouth learning, Games Econ. Behav. 46 (2004) 1–22.
[10] Abhijit V. Banerjee, A simple model of herd behavior, Quart. J. Econ. 107 (3) (August 1992) 797–817.
[11] Marco Bardi, Italo Capuzzo-Dolcetta, Optimal Control and Viscosity Solutions of Hamilton–Jacobi–Bellman Equations, Birkhäuser, Boston, Basel, Berlin, 1997.
[12] Dirk Bergemann, Jusso Välimäki, Experimentation in Markets, Rev. Econ. Stud. 67 (2000) 213–234.
[13] Sushil Bikhchandani, David Hirshleifer, Ivo Welch, A theory of fads, fashion, custom, and cultural change in informational cascades, J. Polit. Economy 100 (5) (October 1992) 992–1026.
[14] Max R. Blouin, Roberto Serrano, A decentralized market with common values uncertainty: Non-steady states, Rev. Econ. Stud. 68 (2001) 323–346.
[15] Partick Bolton, Christopher Harris, Strategic experimentation, Econometrica 67 (1999) 349–374.
[16] Roberto Burguet, Xavier Vives, Social learning and costly information acquisition, Econ. Theory 15 (1) (2000) 185–205.
[17] Jean-Paul Décamp, Thomas Mariotti, Investment timing and learning externality, J. Econ. Theory 18 (2004) 80–102.
[18] Peter DeMarzo, Dimitri Vayanos, Jeffrey H. Zwiebel, Persuasion bias, social influence, and uni-dimensional opinions, Quart. J. Econ. 118 (2003) 909–968.
[19] Darrell Duffie, Gustavo Manso, Information diffusion in large population, Amer. Econ. Rev., Papers and Proceedings 97 (2007).
[20] Darrell Duffie, Semyon Malamud, Gustavo Manso, The relative contributions of private information sharing and public information releases to information aggregation, J. Econ. Theory 145 (2010) 1574–1601.
[21] Douglas Gale, Shachar Kariv, Bayesian learning in social networks, Working Paper, New York University, 2003.
[22] Mikhail Golosov, Aleh Tsyvinski, Guido Lorenzoni, Decentralized trading with private information, Working Paper, MIT and Yale, 2007.
[23] Edward J. Green, Eliciting traders' knowledge in "frictionless" asset markets, Working Paper, 1991.
[24] Christian Hellwig, Heterogeneous information and the welfare effects of public information disclosures, 2005.
[25] Brett Katzman, John Kennan, Neil Wallace, Output and price level effects of monetary uncertainty in a matching model, J. Econ. Theory 108 (2003) 217–255.
[26] Godfrey Keller, Sven Rady, Martin Cripps, Strategic experimentation with exponential bandits, Econometrica 73 (2005) 39–68.
[27] Philipp Kircher, Andrew Postlewaite, Strategic firms and endogenous consumer emulation, Quart. J. Econ. 123 (2008) 621–661.
[28] Marcello Miccoli, Optimal dynamic public communication, Working Paper, 2011.
[29] Stephen Morris, Hyun Song Shin, The social value of public information, Amer. Econ. Rev. 92 (2002) 1521–1534.
[30] Stephen Morris, Hyun Song Shin, Central bank transparency and the signal value of prices, Brookings Pap. Econ. Act. 2 (2005) 1–66.
[31] Bernt Oksendal, Stochastic Differential Equations, Springer, Berlin, 1985.
[32] Philip Protter, Stochastic Integration and Differential Equations, Springer-Verlag, New York, 1990.
[33] Dinah Rosenberg, Eilon Solan, Nicolas Vieille, Social learning in one-arm bandit problems, Econometrica 75 (2007) 1591–1611.
[34] Walter Rudin, Real and Complex Analysis, McGraw–Hill, New York, 1974.
[35] Lones Smith, Peter Norman Sørensen, Rational social learning by random sampling, Working Paper, University of Michigan, 2005.

[36] Andrei N. Tikhonov, Adelaida B. Vasil'eva, Aleksei G. Sveshnikov, Differential Equations, Springer-Verlag, Berlin, 1985.
[37] Alberto Trejos, Randall Wright, Search, bargaining, money, and prices, J. Polit. Economy 103 (1) (1995) 118–141.
[38] Xavier Vives, How fast do rational agents learn?, Rev. Econ. Stud. 60 (1993) 329–347.
[39] Xavier Vives, Learning from others: A welfare analysis, Games Econ. Behav. 20 (1997) 177–200.
[40] Neil Wallace, Short-run and long-run effects of changes in money in a random-matching model, J. Polit. Economy 105 (1997) 1293–1307.
[41] Asher Wolinsky, Information revelation in a market with pairwise meetings, Econometrica 58 (1990) 1–23.