Learning from Private and Public Observations of Others' Actions*

Manuel Amador[†] and Pierre-Olivier Weill[‡]

September 25, 2007

Abstract

We study the diffusion of dispersed private information in a large economy. We assume that agents learn from the actions of others through both a private channel, which represents learning from local interactions, and a public channel, which represents learning from prices. We show that the private and public channels of information diffusion generate different learning dynamics and welfare implications. In particular, while public information always increases welfare when agents learn only from a public channel, it can reduce it when a private learning channel is present.

*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, 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 also thank Romans Pancs for excellent research assistance.

†Department of Economics, Stanford University, Stanford, CA 94305-6072, e-mail: `amador@stanford.edu`

‡Department of Economics, University of California Los Angeles, Bunche 8283, Los Angeles, Box 951477, CA 90095, e-mail: `poweill@econ.ucla.edu`

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 strikingly 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 Vives (1993, 1997). 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 where the state of the world is revealed and a 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 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 is the private learning channel and represents information gathered through private communication and local interactions. The second signal is public, shared with everyone: it is the public learning channel and represents an endogenous aggregate variable, such as a price or some macroeconomic aggregate.

We solve for an equilibrium in which agents eventually learn the truth. Like in the herding literature, there is a public information externality: better public information reduces the informational content of both the public and the private channels and slows down learning. Indeed, with an increase in the precision of public information, each agent's action becomes more sensitive to the public information that is already known by everyone, and hence useless for the purpose of learning. At the same time, each agent's action becomes less sensitive to private information. This makes it harder to glean private information from the noisy observation of agents' actions, and thus generates a negative learning externality: information will diffuse more slowly through all channels.

Now imagine that a benevolent agency holds some relevant, but partial, information and ponders whether to release it publicly. A release would have a direct beneficial effect as it would make agents' current decisions better informed. However, it has the negative effect of slowing the diffusion of private information. We show that when agents only learn through the public channel, the negative externality 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 so slows information diffusion that all agents end up less informed in the long run. 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 have enough time to learn from others actions before the state of the world is revealed, then a given marginal increase in the precision of the initial public signal is always socially costly.

We characterize the dynamic behavior of the distribution of agents' beliefs, in closed form. The average belief converges along an S-shaped curve, as long as the initial information is sufficiently dispersed and the private channel is active.[1] Agents can become more and more confident about the information they gathered privately, taking actions that are increasingly sensitive to this private information. This makes

[1]The S-shaped pattern has been documented in a number of empirical studies of social learning. See Chamley (2004, chap. 9) and also Jovanovic and Nyarko (1995)). See also the subsequent learning models of Fernández (2007) and Fogli and Veldkamp (2007) to explain the S shape in women labor force participation.

the private and public learning channels more informative, generating more precise signals, faster learning and in some cases, a convex learning curve at the beginning. On the other hand, convergence to the truth implies that, in the limit, the average belief is concave. Although the average belief follows an S-shaped curve, we show that the cross-sectional dispersion of agents' beliefs will converge to zero along a hump-shaped curve. Indeed, when agents learn privately, they learn independently, and their learning histories are increasingly heterogeneous: hence, the dispersion of beliefs might increase early on. However, this dispersion eventually must converge to zero as agents learn the truth. Importantly, we show that when agents learn only from a public channel and the private channel is shut down, then the S and the hump disappear: the average belief converges to the truth along a concave curve, and the dispersion of beliefs converges to zero along a decreasing curve.

The final section of the paper applies and generalizes our results. In a first extension, we relax our baseline-model assumption that agents do not learn from their own experience, in that they cannot observe their payoffs until the state is revealed. We show that relaxing this assumption does not significantly alter our results. Next, in a second extension, we provide an example of public signals being generated by the observation of market prices. We assume that firms face quadratic adjustment costs in investment and buy capital goods in a competitive market. After some random time to build a capital stock, these goods become inputs in a production technology with a common but initially unknown productivity. As anticipated by our baseline model, the price of capital generates a public signal centered around the average action, which in this case is the average investment. Different from our baseline model, though, this example features a payoff externality: the price of the capital affects firms' profits, and hence depends on the aggregate investment. We show that even in the presence of this externality, the learning dynamics and the welfare results remain the same as in the baseline model. The extensions demonstrate the tractability of our continuous-time set up, and suggest that the results derived are robust to more explicit economic environments.

Our work is closely related to the recent literature on the social value of public information. As initially shown by Morris and Shin (2002), public information can reduce welfare in the presence of a payoff externality. They set up a beauty contest game where the private payoffs from coordination distort the actions of agents, but wash out in the aggregate. The release of a public signal may then end up reducing welfare by facilitating wasteful coordination among agents. Hellwig (2005), and Angeletos and Pavan (2005)) have characterized more generally the conditions under which the social value of public information is negative. However, the models used so far have been essentially static. Our contribution is analyzing an alternative mechanism based on a dynamic information externality: in our baseline model there are no payoff externalities[2], but public information crowds out the diffusion of private information in the population. Morris and Shin (2005) set up an overlapping-generations beauty-contest game in which a similar learning externality arises: they show that when it discloses its information, a central bank ends up learning less from the actions of private agents. Morris and Shin's model is designed for studying the learning process of the central bank, while our work focuses on the learning process and welfare of private agents.[3]

The social learning literature, started by Banerjee (1992) and Bikhchandani et al. (1992), has been concerned with information externalities. The surprising result is the possibility of informational cascades and herding: agents may choose to disregard their private information, acting solely on the basis of their public information, and take the "wrong" action. Thus public information, by facilitating the emergence of herds, can reduce welfare. However, the appearance of cascades and herds requires bounded beliefs and a discrete set of actions. Our model is closest to Vives (1993) and Vives (1997) 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 and private channels, and we argue that this has strikingly different implications for dynamics and welfare, as described in Section 5.2.

[2]See Section 6 for a market application with a payoff externality.

[3]Because private agents in Morris and Shin's model do not accumulate information over time, either publicly or privately, they do not suffer from the dynamic learning externality emphasized in our paper. Hence, in the absence of a payoff externality, a release of public information always improves their welfare (although it reduces the precision of the information gathered by the central bank). This result stands in contrast to the welfare results of our paper.

Burguet and Vives (2000) present a model where more public information can reduce welfare because it reduces the incentives to acquire private information which is costly. In contrast, in our model, private information (and public) is continuously acquired by agents at no cost. Differently from all of the above, our continuous-time set up has the advantage of being highly tractable: we solve the model in closed-form and in Section 6 we easily extend it in various directions. Some recent work on social learning has focused on learning in networks: Bala and Goyal (1998), Gale and Kariv (2003), and Smith and Sorensen (2005) study deterministic networks with a finite number of agents, and Banerjee and Fudenberg (2004) provide a continuum-of-agents setup (see also DeMarzo et al. (2003) for a network of boundedly rational agents). But because they lack tractability, these models end up focusing almost exclusively on the question of convergence to the truth. Our continuous-time model, with a continuum of agents, allows us to take the learning-in-network literature a step further, analyzing transitional dynamics, welfare, and the impact of public information.

Another body of research uses search-and-matching models to study how agents learn from local interactions with others. In Wolinsky (1990) seminal work, information diffuses at the individual level but stays constant at the aggregate level: indeed, agents leave the economy after trading and uninformed agents continuously enter the economy (see also the recent work of Kircher and Postlewaite (2006)). The issue of convergence when information diffuses in the aggregate also has been addressed in Green (1991), Blouin and Serrano (2001), and in Duffie and Manso (2006). Wallace (1997), Katzman et al. (2003), and Araujo and Shevchenko (2006) address learning about a money supply shock in Trejos and Wright's [1995] random-matching model. For tractability, they assume that the money supply becomes public after one or two periods. Araujo and Camargo (2006) relax this assumption in a Kiyotaki and Wright (1989) model, and study the government incentives to expand the money supply. Our setup is somewhat simpler than these models, because agents do not learn from trading but rather from observing the action of others. The benefit of this simplification is that we can explicitly characterize the transitional dynamics of beliefs and study the welfare impact 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 provides the dynamics of information diffusion. Section 5 studies the impact on dynamics and welfare of changes in the quality of public information. Section 6 provides applications, and section 7 concludes. The appendix collects all the proofs not in the main text.

2 Set up

Time is continuous and runs forever. We fix a probability space $\{\Omega, \mathcal{G}, Q\}$ together with an information filtration $\{\mathcal{G}_t, t \geq 0\}$ satisfying the usual conditions (Protter (1990)). The economy is populated by a continuum of agents whose payoff depends on some random state of the world, $x \in \mathbb{R}$. At time zero, agents receive a public signal about x, which represents information released by some public agency. At the same time, they receive independent private signals, which represent the information that is dispersed in the population.

Given the continuum assumption, any particular agent who could gather the signals of all others would know the realized value of x. In our model, such pooling won't be feasible: instead, the initial private information will diffuse slowly in the population through local interactions and the noisy observation of aggregate variables.

Preferences and Information.

At each time until some random "day of reckoning" $\tau > 0$, each agent $i \in [0,1]$ takes an action a_{it}. At time τ, the state of world is revealed and the agent receives the payoff

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

We take the random time τ to be Poisson distributed with intensity $\lambda > 0$, and independent of everything else.

Agents share a common prior that x is normally distributed with mean zero and precision $\bar{P}$, but their information evolves over time in the following way. First at time zero, agents become asymmetrically informed about x: each of them receives both a private signal z_{i0} and a public signal Z_0

$$
\begin{aligned}
z_{i0} &= x + \tfrac{\omega_{i0}}{\sqrt{\pi_0}}, && (1)\\
Z_0 &= x + \tfrac{W_0}{\sqrt{\Pi_0}} && (2)
\end{aligned}
$$

where $(\omega_{i0})_{i\in[0,1]}$ and W_0 are normally distributed with mean zero and variance 1, pairwise independent, and independent of everything else. The precision of the private and public signals are measured by π_0 and Π_0, respectively.

The signal Z_0 represents information released by a public agency. By continuously varying its precision, Π_0, we can obtain the impact on diffusion and welfare of varying the size of an information release. The continuum of private signals, z_{i0}, represents dispersed information about aggregate economic conditions. In particular, a smaller π_0 means that information is more dispersed: indeed, the private information of an individual agent becomes less precise, but aggregate private information remains constant, as the average of the private signals (1) reveals the underlying state x. Hence, if agents were to pool their time-zero private signals, they would know x. Another standard way of obtaining such pooling of information in economic models is to assume that agents observe an endogenous aggregate variable. For instance, as will become clear later, if agents were able to observe the cross-sectional average action

$$
A_t \equiv \int_0^1 a_{it}\, di, \tag{3}
$$

in the first period, $t = 0$, then they would be able to infer the exact value of x. We prevent agents from instantly pooling all their private information by introducing noise into the observations of the average action in the following fashion.

Each agent i observes a private signal z_{it} of the average action satisfying

$$
dz_{it} = A_t dt + \frac{d\omega_{it}}{\sqrt{p_\varepsilon}}, \tag{4}
$$

with initial condition given by equation (1). Such a signal captures the decentralized gathering of information. One could think, for instance, of local interaction and of private communication, such as gossips.[4]

Agents also observe a noisy public signal of the average action, Z_t, solving the stochastic differential equation

$$dZ_t = A_t dt + \frac{dW_t}{\sqrt{P_\varepsilon}}, \tag{5}$$

with initial condition given by equation (2). This public signal represents the information conveyed by some endogenous aggregate variable. We let $(\omega_i)_{i\in[0,1]}$ and W, in equations (4) and (5), be pairwise independent Wiener processes with initial conditions ω_{i0}, W_0, independent from x and from τ.

Note that one crucial assumption in specification (4)-(5) is that agents' signals are centered around the cross-sectional average action, A_t. This assumption, which is standard in the social learning literature (see Chamley (2004) and references therein), means that agents do not learn by directly observing the information of others, but rather by observing noisy averages of others' actions.[5] In particular, Section 6.2 proposes a market setting that is equivalent to our current set up, and where observing the market price is the same as observing Z_t.

Equilibrium.

Given a process A for the average action, we can define the information sets generated in the economy as

Definition 1. *Let A be a process for the average action. Then public information is represented by the filtration $\{\mathcal{F}_t,\ t \geq 0\}$ generated by $\{Z_s,\ 0 \leq s \leq t\}$. Similarly, an agent's private information is represented by the filtration $\{\mathcal{F}_{it},\ t \geq 0\}$ generated by $\{z_{is},\ 0 \leq s \leq t\}$. The filtration $\mathcal{G}_{it}$ generated by $\mathcal{F}_t \cup \mathcal{F}_{it}$ represents all the information available to agent $i \in [0,1]$ at any time $t > 0$.*

[4]The work of Amador and Weill (2006) suggests that specification (4) 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, Amador and Weill (2006) informally arrive at specification (4).

[5]Regarding the endogenous private signals, Amador and Weill (2007) provide a different interpretation: they are generated because agents receive exogenous private signals about the noise in public endogenous aggregates. Their paper shows how such signals about the noise naturally arise in a monetary economy.

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' intertemporal problems are essentially static, and together with their quadratic payoffs, this implies that their optimal actions are the expectation

$$a_{it} = E\left[x \mid \mathcal{G}_{it}\right] \tag{6}$$

of the random variable x, conditional on their information filtration $\{\mathcal{G}_{it}\,,\, t \geq 0\}$. We can now state our notion of equilibrium.

Definition 2. *An* equilibrium *is a collection of processes a_i and A, where a_i is a $\mathcal{G}_{it}$-adapted individual action process and A is a $\mathcal{G}_t$-adapted average action process, such that:*

(i) at each time, a_{it} solves equation (6),

(ii) and A_t solves equation (3).

The notion of equilibrium is standard. The average action process A determines the information sets of each agent, both the private and public filtrations. Given their information, agents maximize their expected utility. Finally, the cross-sectional average action of the population, A, is consistent with the agents' optimal actions.

3 An Equilibrium

In this section we explicitly construct an equilibrium. We start by showing that an agent's action at any time is the convex combination of two forecasts of the state of the world, x: a public forecast shared with everyone in the economy, and a private forecast containing all the information observed by that agent and no one else.

Hypothesis (H). There exist two continuous precision functions, π_t and Π_t, such that observing Z_t is equivalent to observing $\tilde{Z}_t$, and observing (Z_t, z_{it}) is equivalent to observing $(\tilde{Z}_t, \tilde{z}_{it})$, where

$$d\tilde{z}_{it} = xdt + \frac{d\omega_{it}}{\sqrt{\pi_t}}. \tag{7}$$

$$d\tilde{Z}_t = x\,dt + \frac{dW_t}{\sqrt{\Pi_t}} \tag{8}$$

Formally, let $\tilde{\mathcal{F}}_{it}$ be the filtration generated by $\{z_{i0}, \tilde{z}_{iu},\ 0 < u \leq t\}$, and let $\tilde{\mathcal{F}}_t$ be the filtration generated by $\{Z_0, \tilde{Z}_u,\ 0 < u \leq t\}$. Then $\mathcal{F}_t = \tilde{\mathcal{F}}_t$ and $\mathcal{G}_{it} = \mathcal{F}_t \cup \mathcal{F}_{it} = \tilde{\mathcal{F}}_t \cup \tilde{\mathcal{F}}_{it}$.

Hypothesis (H) states that observing the original public and private signals, z_{it} and Z_t, which are centered around the average action in the population, is equivalent to observing two transformed signals centered instead around the true value of the parameter x. These transformed signals are generated with the same noises, (ω_i) and W, but with time-varying precisions, π_t and Π_t, which are determined in equilibrium.

We begin by introducing some notations. Conditional only on the history of the transformed private signal $\tilde{z}$, and the initial private signal z_0, together with a totally diffuse initial prior, we define an agent's best private forecast of x as $\hat{x}_{it} = E\big[x \mid \tilde{\mathcal{F}}_{it}\big]$. We refer to $\hat{x}_{it}$ as the **private forecast** of agent i at time t, and we denote the precision of this forecast by

$$p_t = \frac{1}{E\big[(x - \hat{x}_{it})^2 \,|\, \tilde{\mathcal{F}}_{it}\big]}\ , \tag{9}$$

where we anticipate that the precision p_t of the private forecast is the same for all agents at any given time t.

Similarly, conditional on the history of the transformed public signal $\tilde{Z}$, the initial public signal Z_0, together with the initial common prior that x is normally distributed with mean zero and precision $\bar{P}$, we denote an agent's best forecast of the state by $\hat{X}_t \equiv E\big[x \,|\, \tilde{\mathcal{F}}_t\big] = E\left[x \,|\, \mathcal{F}_t\right]$. We refer to $\hat{X}_t$ as the **public forecast** and denote its precision by P_t.

Our objective is to derive the dynamics of the action of any given agent, i, as a combination of both her private and the public forecasts, $\hat{x}_{it}$ and $\hat{X}_t$. Bearing this in mind, we first determine the dynamics of the private forecast, $\hat{x}_{it}$, and of the public forecast, $\hat{X}_t$.

In order to see how we obtain the law of motion of $\hat{x}_{it}$, it is convenient to consider the following discrete time approximation of the filtering problem (the formal proof in continuous time can be found in the appendix). Suppose that at the beginning of time t, the private forecast is that x is normally distributed with mean $\hat{x}_{it}$ and precision p_t. During the small time interval $[t, t+\Delta]$, equation (7) means that an agent receives a private signal, $\Delta\tilde{z}_{it}$, which is approximately equal to

$$\begin{aligned} & \Delta\tilde{z}_{it} = x\Delta + \sqrt{\Delta}\frac{\varepsilon_{it}}{\sqrt{\pi_t}} \\ \Leftrightarrow \quad & \frac{\Delta\tilde{z}_{it}}{\Delta} = x + \frac{\varepsilon_{it}}{\sqrt{\pi_t\Delta}}, \end{aligned} \tag{10}$$

for some standard normal random variable ε_{it}.[6] Equation (10) means that agents receive a signal centered around x, with precision $\pi_t\Delta$. Given that all distributions are normal, the private forecast of x after having observed $\Delta\tilde{z}_{it}$ is then a weighted average of the prior forecast and the newly received signal,

$$\hat{x}_{it+\Delta} = \left(\frac{p_t}{p_t + \pi_t\Delta}\right)\hat{x}_{it} + \left(\frac{\pi_t\Delta}{p_t + \pi_t\Delta}\right)\frac{\Delta\tilde{z}_{it}}{\Delta}, \tag{11}$$

where the weights on the prior and the signal reflect their relative precision. Subtracting $\hat{x}_{it}$ on both sides, and plugging equation (10), the change in the private forecast is

$$\hat{x}_{it+\Delta} - \hat{x}_{it} = \frac{\pi_t}{p_t + \pi_t\Delta}\left[\Delta\left(x - \hat{x}_{it}\right) + \sqrt{\Delta}\frac{\varepsilon_{it}}{\sqrt{\pi_t}}\right], \tag{12}$$

suggesting that, in the continuous-time limit as Δ goes to zero, the private forecast $\hat{x}_{it}$ solves the stochastic differential equation,

$$d\hat{x}_{it} = \frac{\pi_t}{p_t}\left[\left(x - \hat{x}_{it}\right)dt + \frac{d\omega_{it}}{\sqrt{\pi_t}}\right]. \tag{13}$$

To complete the characterization of $\hat{x}_{it}$ it is also necessary to obtain the law of motion

[6] The variance of the increment of a Wiener process is proportional to the time interval, which implies the presence of $\sqrt{\Delta}$ in the discrete time approximation.

for p_t. Note that at each time interval $[t, t+\Delta]$, the private signal has precision $\pi_t \Delta$. Since, the precision of the posterior is the sum of the precision of the prior, and of the precision of the signal, it follows that $p_{t+\Delta} = p_t + \pi_t \Delta$, and hence letting Δ go to zero,

$$dp_t = \pi_t \, dt,$$

which determines the law of motion for p_t. Given that p_0 and π_t are the same for all agents, it follows that p_t is also the same for all agents.

A similar analysis determines the law of motion of the public forecast $\hat{X}_t$ and its precision P_t under hypothesis (H). The following proposition formalizes these results.

Proposition 1 (Dynamics of Private and Public Forecasts)**.** *Suppose that hypothesis (H) holds. The private and public forecasts $(\hat{x}_{it}, \hat{X}_t)$ and the precisions (p_t, P_t) solve the system of stochastic differential equations*

$$d\hat{x}_{it} = \frac{\pi_t}{p_t}\left[(x - \hat{x}_{it})\, dt + \frac{d\omega_{it}}{\sqrt{\pi_t}}\right] \tag{14}$$

$$d\hat{X}_t = \frac{\Pi_t}{P_t}\left[\left(x - \hat{X}_t\right)\, dt + \frac{dW_t}{\sqrt{\Pi_t}}\right] \tag{15}$$

$$dp_t = \pi_t \, dt \tag{16}$$

$$dP_t = \Pi_t \, dt \ , \tag{17}$$

with the initial conditions $p_0 = \pi_0$, and $P_0 = \bar{P} + \Pi_0$. In addition, the above system can be integrated into

$$\hat{x}_{it} = x + \frac{1}{p_t}\left[\sqrt{\pi_0}\omega_{i0} + \int_0^t \sqrt{\pi_u} d\omega_{iu}\right] \tag{18}$$

$$\hat{X}_t = \left(1 - \frac{\bar{P}}{P_t}\right) x + \frac{1}{P_t}\left[\sqrt{\Pi_0} W_0 + \int_0^t \sqrt{\Pi_u} dW_u\right] \tag{19}$$

Proof. In the appendix. □

We next show that an agent's posterior belief is a weighted average of his private forecast and the public forecast, where the weights are given by the relative precision of those forecasts.

Corollary 1. *Suppose that hypothesis (H) holds. Then, at time $t > 0$, agent i's posterior belief about x is normally distributed with mean*

$$a_{it} = E\left[x \,|\, \mathcal{G}_{it}\right] = \frac{p_t}{p_t + P_t}\,\hat{x}_{it} + \frac{P_t}{p_t + P_t}\,\hat{X}_t, \tag{20}$$

and precision,

$$\left\{E\left[(x - a_{it})^2 \;\; |\, \mathcal{G}_{it}\right]\right\}^{-1} = p_t + P_t. \tag{21}$$

Proof. In the appendix. □

We can now aggregate the actions in the population, and obtain the average action,

$$A_t = \frac{p_t}{p_t + P_t}\int_0^1 \hat{x}_{it}\,di + \frac{P_t}{p_t + P_t}\,\hat{X}_t = \frac{p_t}{p_t + P_t}\,x + \frac{P_t}{p_t + P_t}\,\hat{X}_t\ , \tag{22}$$

where we use that the cross-sectional average private forecast is x. This follows from noting that the idiosyncratic errors of the private forecasts in equation (18) wash out.

So, under hypothesis (H), we have characterized the dynamics of the beliefs of an agent, after observing signals according to equation (7) and equation (8), and computed the implied aggregate behavior. To complete the characterization of equilibrium, we now provide an intuitive verification of our Hypothesis (H), and compute the law of motion for the precisions of the private and the public forecasts, p_t and P_t (the formal proof is left to the appendix.)

Consider the public signal, $dZ_t = A_t\,dt + dW_t/\sqrt{P_\varepsilon}$ and suppose that A_t evolves according to equation (22). Note that $\hat{X}_t$ is public under hypothesis (H) and known at time t. Hence, an agent can subtract from dZ_t the part $P_t/(p_t + P_t)\hat{X}_t$ of A_t that she already knows. This implies that the public signal is observationally equivalent to $p_t/(p_t + P_t)x\,dt + dW_t/\sqrt{P_\varepsilon}$. Dividing this through by $p_t/(p_t + P_t)$, we find that observing the public signal is equivalent to

$$d\tilde{Z}_t = x\,dt + \frac{dW_t}{\sqrt{P_\varepsilon}\frac{p_t}{p_t+P_t}},$$

as conjectured, implying that $\Pi_t = P_\varepsilon \left[p_t/(p_t + P_t)\right]^2$.

Similarly, together with the public signal that determines A_t, the private signal is observationally equivalent to

$$d\tilde{z}_{it} = x\,dt + \frac{d\omega_{it}}{\sqrt{p_\varepsilon}\frac{p_t}{p_t+P_t}},$$

implying that $\pi_t = p_\varepsilon \left[p_t/(p_t + P_t)\right]^2$. This confirms that hypothesis (H) holds. Hence we can state the following Theorem:

Theorem 1 (Equilibrium)**.** *Let*

$$\pi_t = p_\varepsilon \left(\frac{p_t}{p_t + P_t}\right)^2 \qquad \textit{and} \qquad \Pi_t = P_\varepsilon \left(\frac{p_t}{p_t + P_t}\right)^2 , \tag{23}$$

and where p_t and P_t evolve according to equations (16) and (17). The collection of processes a_i and A, where a_i solves equation (20) and A solves equation (22) is an equilibrium.

Proof. In the appendix. □

Note that π_t and Π_t represent the precisions of the newly received private and public signals. The informativeness of these newly received signals at any time is a function of the precisions p_t and P_t of the private and public forecasts: it increases with p_t and decreases with P_t. Hence, the more agents know privately, the more informative their new signals become, and the faster they learn. Improvements in the public forecast 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. The next section sheds some light on these different effects by studying the dynamics of the system.

4 Information Dynamics

This section analyzes the dynamics of the equilibrium just described. We first study the dynamics of the precisions of public and private forecasts. We show that, in the limit as time goes to infinity, each agent learns the realization of the state of the world x, and we characterize the speed of the learning process. We also show that, if the initial private information is sufficiently dispersed, then the mean belief in the population converges to the truth along a S-shaped curve, and the dispersion of beliefs in the population converges to zero along a hump-shaped curve.

4.1 Precisions

From (16) and (17) and using the equilibrium values of π_t and Π_t, the precisions p_t and P_t of the private and the public forecasts evolve according to the Ordinary Differential Equations (ODE)

$$\dot{p}_t = p_\varepsilon \left(\frac{p_t}{P_t + p_t} \right)^2 \tag{24}$$

$$\dot{P}_t = P_\varepsilon \left(\frac{p_t}{P_t + p_t} \right)^2 , \tag{25}$$

Note also that, by (21), the precision of agent i's beliefs is the sum of the precisions of his private and the public forecasts: $p_t + P_t$.

Note that ODE (25) is equal to ODE (24) 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 $P_t - P_0 = P_\varepsilon/p_\varepsilon(p_t - p_0)$. Plugging this into equation (24), we obtain

$$\dot{p}_t = \left(\frac{p_t}{p_t + \alpha/\beta} \right)^2 \frac{p_\varepsilon}{\beta^2}, \tag{26}$$

where $\alpha = P_0 - P_\varepsilon/p_\varepsilon p_0$ and $\beta = 1 + P_\varepsilon/p_\varepsilon$. Hence given an initial condition for the precision p_0 of the private forecast, and using equation (26), it is possible to characterize the dynamics of p_t.

Lemma 1 (Dynamics of p_t). *If $p_\varepsilon > 0$, then the precision of the private forecasts, p_t, is such that $p_\varepsilon t = H(p_t) - H(p_0)$ where $H(p) = 2\alpha\beta \log p + \beta^2 p - \alpha^2/p$,*

With knowledge of p_t, we proceed by characterizing the dynamics of the entire system.

Proposition 2 (Dynamics of Precisions)**.** *If $p_\varepsilon > 0$ then (i) the precision of the private forecast monotonically converges to infinity, (ii) the ratio $p_t/(p_t + P_t)$ monotonically converges to $p_\varepsilon/(p_\varepsilon + P_\varepsilon)$, and (iii) as $t \to \infty$ the precision of the total beliefs, $p_t + P_t$, is such that*

$$p_t + P_t = \left(\frac{p_\varepsilon}{p_\varepsilon + P_\varepsilon}\right)^2 (p_\varepsilon + P_\varepsilon)\, t + 2\left(\frac{P_\varepsilon}{p_\varepsilon} p_0 - P_0\right) \log(t) + R_t, \tag{27}$$

where R_t is some bounded function.

Proof. In the appendix. ☐

To understand the asymptotic result stated in the proposition, note that part (ii) of the proposition shows that in the limit, as time goes to infinity, the precision of the signals generated by the average action converges. Hence, by equation (23), the sum of the precisions of the endogenously generated private and public signals converges in the limit to

$$\lim_{t\to\infty} \left(\frac{p_t}{p_t + P_t}\right)^2 (p_\varepsilon + P_\varepsilon) = \left(\frac{p_\varepsilon}{p_\varepsilon + P_\varepsilon}\right)^2 (p_\varepsilon + P_\varepsilon) \tag{28}$$

As can be seen from equation (27), this is the coefficient on t in the asymptotic expansion of the precision of total beliefs. In other words, when time goes to infinity, it is as if agents repeatedly observe independent signals centered around x with precision given by (28).

4.2 Cross-sectional Beliefs

In the present normal-quadratic framework, the distribution of beliefs in the population is also normal and can be characterized in closed form. To see this, recall that the action of agent i is

$$a_{it} = \frac{p_t}{p_t + P_t} \hat{x}_{it} + \frac{P_t}{p_t + P_t} \hat{X}_t\ , \tag{29}$$

and that an agent's action is the same as her belief. Then, equations (19) and (18) imply that cross-sectional beliefs are normally distributed, and therefore entirely characterized by their average and dispersion. Taking expectations conditional on x, it follows that

$$E\left[a_{it}\,|\,x\right] = E\left[A_t|x\right] = x\left(1 - \frac{\bar{P}}{p_t + P_t}\right) \,, \tag{30}$$

and so the rate at which the expected average belief converges to the truth is the same as the rate at which the variance of an agent's belief, $1/(p_t + P_t)$, approaches zero.

The cross-sectional dispersion of beliefs, or actions, also can be computed. Given that the second term in the right hand side of equation (29) is common across all agents, the cross-sectional dispersion is simply the variance of the first term,

$$\theta_t \equiv E[(a_{it} - A_t)^2] = \frac{p_t}{(p_t + P_t)^2} \,,$$

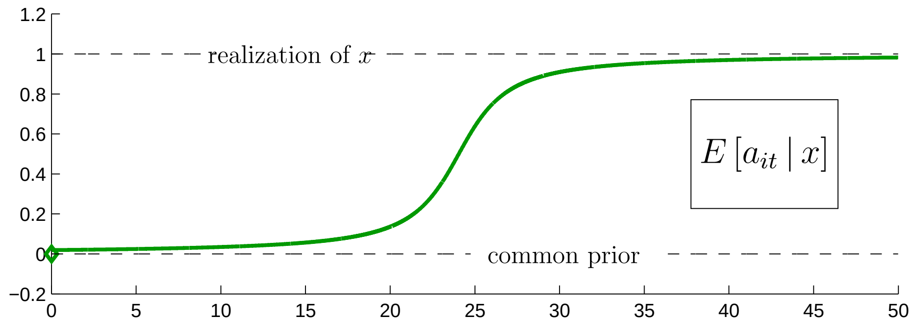
where the cross-sectional variance of $\hat{x}_{it}$ is $1/p_t$. Note that the dynamics of the average belief, or action, are given by

$$\begin{aligned} \frac{dE\left[a_{it}|x\right]}{dt} &= x\bar{P}\frac{\dot{p}_t + \dot{P}_t}{(p_t + P_t)^2} = x\bar{P}(p_\varepsilon + P_\varepsilon)\left(\frac{p_t}{(p_t + P_t)^2}\right)^2 \\ &= x\bar{P}(p_\varepsilon + P_\varepsilon)\theta_t^2. \end{aligned} \tag{31}$$

Hence, the expected action, $E[a_{it}|x]$, monotonically approaches x and its changes are proportional to the square of the cross-sectional dispersion, θ_t. Similarly, the dynamics of the cross-sectional dispersion are given by

$$\frac{d\theta_t}{dt} = \frac{\dot{p}_t}{(p_t + P_t)^2} - \frac{2p_t\left(\dot{p}_t + \dot{P}_t\right)}{\left(p_t + P_t\right)^2} = \frac{\dot{p}_t\,(p_\varepsilon + P_\varepsilon)}{(p_t + P_t)^2}\left(\frac{p_\varepsilon}{p_\varepsilon + P_\varepsilon} - 2\frac{p_t}{p_t + P_t}\right), \tag{32}$$

where we used that $\dot{P}_t = P_\varepsilon/p_\varepsilon \dot{p}_t$. Recall that Proposition 2 showed that $p_t/(p_t + P_t)$ converges to $p_\varepsilon/(p_\varepsilon + P_\varepsilon)$ monotonically. Together with the fact that $\dot{p}_t$ is positive at each time, this implies that the cross-sectional dispersion is eventually decreasing. It follows also that the cross-sectional dispersion will have a hump shape as long as it is increasing at time zero, that is as long as $p_\varepsilon/(p_\varepsilon + P_\varepsilon) > 2p_0/(p_0 + P_0)$.



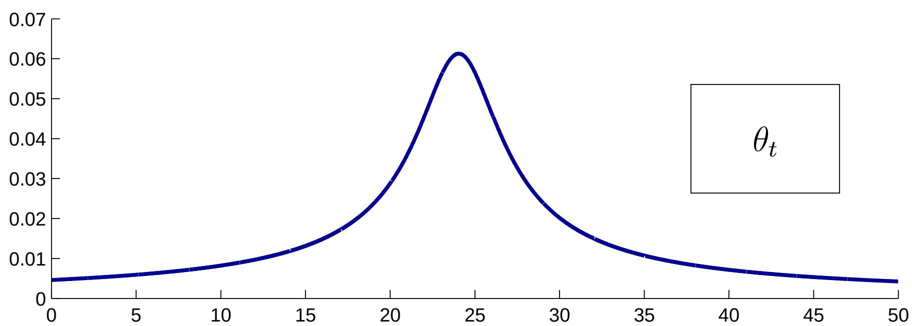


Figure 1: This is a particular example of the dynamics of cross-sectional beliefs when the path of the mean is S-shaped and the dispersion of belief is hump-shaped. We choose $\bar{P} = 2.04$, $\pi_0 = \Pi_0 = 0.02$, $p_\varepsilon = P_\varepsilon = 10$. In order to calculate the time path of (p_t, P_t), we use the Euler method (see Judd (1999)) with step size $h = 0.01$ to discretize the ODEs (24) and (25).

Finally, note that if the cross-sectional dispersion, θ_t, is hump-shaped, then it follows from equation (31) that the path of $E[a_{it}|x]$ is S-shaped. Moreover, the inflexion time of the S corresponds to the point of highest dispersion of actions. We have just proved the following:

Proposition 3 (S-shaped diffusion)**.** *The path of $E[a_{it}|x]$ monotonically approaches x as time goes to infinity. If $2p_0/(p_0 + P_0) < p_\varepsilon/(p_\varepsilon + P_\varepsilon)$ then there exists a $t_0 > 0$ such that $|dE[a_{it}|x]/dt|$ is increasing for all $t < t_0$ and decreasing for all $t > t_0$. Otherwise, $|dE[a_{it}|x]/dt|$ is decreasing for all t.*

As shown in the proposition and illustrated in Figure 1, the model is able to generate an S-shaped learning curve. The convex part of the S arises because of an information-snowballing effect as long as the initial private information is sufficiently dispersed, i.e. its precision is sufficiently small. In that case, as agents learn privately, they become increasingly confident about their private information. As a result, the weight they assign to their private information becomes larger and larger and both the private and the public signal become more and more informative: learning is faster and faster at the beginning. Eventually, as agents learn the truth, learning must slow down and the learning curve becomes concave.

The hump shape of the dispersion of belief is more standard, and arises because of the private learning channel. Initially, agents beliefs are concentrated close to their common prior. As time goes by, agents learn privately, independently of each other. Some become optimistic about the state, and others pessimistic, which means that the dispersion of beliefs increases at the beginning. However, this dispersion must converge to zero as agents asymptotically learn the truth.

5 The Impact of Public Information

Recently, several authors[7] have provided conditions for the release of public information in a market to actually reduce welfare. In this literature the models are essentially static, and rely on payoff externalities to generate negative effects of public information releases. In our dynamic model, in contrast, possible negative effects of public information can arise from a dynamic learning externality. Namely, we show in this section that, in our model, the provision of additional public information eventually reduces what agents know, and in some cases indeed reduces welfare.

5.1 Comparative Statics

The dynamics of the precision of the private forecast are governed by

$$\dot{p}_t = \left(\frac{p_t}{P_0 + p_t + P_\varepsilon / p_\varepsilon (p_t - p_0)} \right)^2 p_\varepsilon \ . \tag{33}$$

Hence, either an increase in p_0 or an increase in p_ε, increases $\dot{p}_t$ for any given $p_t > p_0$, and increases the precision of the private forecast at all times. The same can be said

[7]See Morris and Shin (2002), Hellwig (2005), and Angeletos and Pavan (2005))

of the effect of p_0 and p_ε on P_t. Hence, improving private information, either through an increase in p_0 or p_ε, unambiguously increases the amount known at all times. In contrast, the following two propositions show that that improving public information, either through an increase in P_0 or P_ε, has the exact opposite effect:

Proposition 4. *Let $p_\varepsilon > 0$. Consider two initial levels $P_0' > P_0$ of public information, and their associated paths (p_t', P_t') and (p_t, P_t) of private and public precisions. Then $t > 0$, $p_t' < p_t$, and $p_t' + P_t' < p_t + P_t$ for t large enough. If $P_\varepsilon > 0$, then $P_t' < P_t$ for all t large enough.*

Proof. Note first p_0 is unaffected (the initial condition of the public forecasts which is given by the precision of the initial private signals). By equation (26), which describes the law of motion for p_t, we see that an increase in P_0 increases α and reduces strictly $\dot{p}_t$ for all p_t. Hence p_t falls for all t. For the second part, note that equation (27) and using that $P_t = (P_t + p_t)(\beta - 1)/\beta + \alpha/\beta$, implies that

$$P_t = \frac{p_\varepsilon^2 P_\varepsilon}{(p_\varepsilon + P_\varepsilon)^2} t - 2\frac{(P_0 p_\varepsilon - p_0 P_\varepsilon) P_\varepsilon}{p_\varepsilon (p_\varepsilon + P_\varepsilon)} \log(t) + R_t, \tag{34}$$

where R_t is some bounded function. Note that P_0 has no effect in the highest order term, but has a negative effect in the $\log(t)$ term as long $P_\varepsilon > 0$. Hence increasing P_0 by any amount will eventually reduce P_t as well. Similarly, it can be shown for $p_t + P_t$ using equation (27) directly. $\square$

The proposition shows that a more precise initial public signal eventually will reduce the amount known. This result relies on the endogeneity of learning and on the dynamic interaction between the public and the private precisions. As the initial public forecast becomes more precise, agents put more weight on it, $P_t/(p_t + P_t)$, when choosing their actions. At the same time, they put less weight, $p_t/(p_t + P_t)$, on their private forecast, reducing the amount of private information that can be inferred from their actions. Hence, both the public and the private signals become less informative about the state of the world, x, and learning is impaired: the precision p_t of the private forecasts is reduced at all times. The proposition is even stronger: although it initially has a positive effect, public information eventually will reduce the total precision $p_t + P_t$ of agents' belief.

Improvements in the precision of the public information flow (increases in P_ε) have an adverse effect on private learning and reduce the precision of the private forecast at all times. As before, increases in P_ε have opposite effects on the precision of public information. However, unlike before, either effect can dominate asymptotically. Independent of its effect on the precision of the public forecast, an increase in P_ε will always decrease the precision of the total forecast:

Proposition 5. *Let $p_\varepsilon > 0$. Consider two different levels $P'_\varepsilon > P_\varepsilon$ of precision for the public information flow. Let (p'_t, P'_t) and (p_t, P_t) be the respective paths of public and private precisions. Then, $p'_t < p_t$ for all $t > 0$, and $p'_t + P'_t < p_t + P_t$ for all t large enough.*

Proof. Given that p_0 is constant, from (equation (26)), an increase in P_ε, increases β and strictly reduces $\dot{p}_t$ for all p_t. From equation (27), P_ε has a negative effect on the highest order term, and hence increasing it will eventually reduce $p_t + P_t$. □

In both propositions we require $p_\varepsilon > 0$: when agents learn through private interactions, any improvement in the public information eventually reduces the **total** amount learned. In the next section we show that $p_\varepsilon > 0$ is in fact a necessary condition for obtaining such effects: that is the two results no longer hold in the absence of a private learning channel, when $p_\varepsilon = 0$.

5.2 Public Learning without Private Learning: Vives [1997]

In this section we consider the effect of increasing public information when $p_\varepsilon = 0$: agents are initially asymmetrically informed about the state of the world, but subsequently learn (only) from observation of public signals. This is the assumption in most of the herding literature surveyed by Chamley (2004). These herding models are based on a learning externality similar to ours, and they illustrate the negative impacts of public signals when actions lie in discrete spaces (or when payoffs are non-regular, as in Banerjee (1992)) and beliefs are bounded. In our model, actions instead lie in a continuous space and beliefs are unbounded. Unlike the herding models, in our model when $p_\varepsilon = 0$, releasing public information increases the amount known at all time and thus is welfare improving.

Note that we are not the first to analyze such an environment. Indeed, when $p_\varepsilon = 0$, our model becomes a continuous-time version of the discrete-time model of

Vives (1997), although he did not study the effects of a release of public information. Our continuous-time formulation in this case has a very simple closed-form solution, allowing for a complete analysis of the properties of the equilibrium.

Proposition 6 (Vives Case)**.** *When $p_\varepsilon = 0$, the precision of the total beliefs, $p_t + P_t$, is*

$$p_t + P_t = p_0 + P_t = 3^{1/3} \left(p_0^2 P_\varepsilon t + \frac{(p_0 + P_0)^3}{3} \right)^{1/3} \tag{35}$$

Proof. Let $y_t = P_t + p_0$. Given that $p_t = p_0$ we have from (25) that $\dot{y}_t = \frac{1}{y_t^2} p_0^2 P_\varepsilon$. It is now easy to verify that (35) is the solution to this differential equation given initial condition $y_0 = p_0 + P_0$. □

This proposition is the continuous-time counterpart of Vives' intriguing result on asymptotic learning speed: when agents observe only public signals of the average action, the precision of total beliefs goes to infinity at rate $t^{1/3}$. This is one order of magnitude slower than the rate at which the beliefs would converge if agents were to observe noisy signals of x in every period with i.i.d. noise – in that case, the precision of belief would go to infinity at the linear rate t. Note that, when $p_\varepsilon > 0$, equation (27) shows that the precision of belief also goes to infinity at a linear rate. Hence, when learning is based only on public signals, the speed of learning is reduced dramatically.

What is the impact of public information on learning when $p_\varepsilon = 0$? As can be seen directly from equation (35), increases in the initial precision of the public signal, or in the precision of the public noise, always make the agents' beliefs more precise: they increase the amount known at all times. Hence, public information is benign in the $p_\varepsilon = 0$ case.[8]

The more general $p_\varepsilon > 0$ case yields the opposite result because an increase in P_0 changes the dynamics of the private forecasts, which eventually creates a negative feedback into the dynamics of the public forecast. In the $p_\varepsilon = 0$ case, there is no such negative feedback because p_t remains constant and equal to p_0 at all times.

[8] In the discrete-time version of the Vives, this is not always correct. However, a similar result holds if we perform the comparative statics from the second period onwards.

To understand this point more generally, suppose that p_t were not constant but followed some *exogenously given*, piecewise-continuous time path.[9] Suppose as well that agents still learn from a public signal centered around the average action, so that the precision of the public forecast solves equation (25).

Proposition 7. *Let p_t follow some exogenous piecewise-continuous time path, and P_t solve equation (25). Let $P'_0 > P_0$, then $P'_t > P_t$ for all t.*

Proof. Suppose the converse. Given continuity of P_t and P'_t, there exists some t such that $P_t = P'_t$. But this implies that the paths of P_t and P'_t are the same at all times, a contradiction. □

The proposition clarifies the result for the Vives' case: when there is no feedback from the public precision, P_t, to the private precision, p_t, better initial public information always increases the precision of the agent's belief at all times.

Morris and Shin (2005) point out that public information may reduce the informational content of macro aggregates. The same result obtains here: a release of public information has the side effect of making the public signals less informative, in that $(p_t/(p_t + P_t))^2$ is reduced at all times. However, as this proposition shows, that mechanism alone is not sufficient for public information to reduce welfare. Indeed, the reason that the informativeness of the public signals is reduced is precisely because agents know more at all times; that is P_t, has increased, while p_t remained constant.

The next section shows that when the dynamics of p_t are impaired by changes in the public information, the social value of public information changes dramatically.

5.3 The Social Value of Public Information

When the private learning channel is active, $p_\varepsilon > 0$, public information may or may not be socially beneficial depending on the trade-off between a short-term gain and a long-term loss. Public information initially improves the precision of agents, generating a short-term gain. The long-term loss, as shown in Proposition 4 comes from a learning externality: public information eventually reduces the amount known

[9] This could be generated, for instance, by the repeated observation of independent exogenous signals centered around x, with precision $\dot{p}_t$.

by everyone. 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 Morris and Shin (2002), we conclude that even in the absence of a payoff externality, more public information can be welfare reducing.

We take our welfare criterion to 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,

$$W = -E\left[\int_0^\tau (a_{it} - x)^2 \, dt\right], \tag{36}$$

which can be rearranged as,

$$\begin{aligned} W &= -E\left[\int_0^\infty \int_0^\tau (a_{it} - x)^2 \lambda e^{-\lambda\tau} \, dt \, d\tau\right] = -E\left[\int_0^\infty (a_{it} - x)^2 \int_t^\infty \lambda e^{-\lambda\tau} \, d\tau \, dt\right] \\ &= -E\left[\int_0^\infty e^{-\lambda t} (a_{it} - x)^2 \, dt\right], \end{aligned} \tag{37}$$

where the random time τ is geometrically distributed: $\lambda e^{-\lambda t}$ is the probability density that the economy ends at time $t > 0$. Because $a_{it} = E[x \mid \mathcal{G}_{it}]$ and $E\left[(a_{it} - x)^2 \mid \mathcal{G}_{it}\right] = 1/(p_t + P_t)$, an application of Fubini's Theorem implies that (37) can be written

$$W = -\int_0^\infty \frac{e^{-\lambda t}}{p_t + P_t} \, dt. \tag{38}$$

Public information increases the total precision $p_t + P_t$ of agents beliefs in the short run but, as shown in Proposition 4, it decreases $p_t + P_t$ in the long run. Hence, because (38) is the present value of $1/(p_t + P_t)$, it is natural to conjecture that as long as λ is close enough to 0, public information reduces welfare. Although intuitive, this result is not obvious 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 $1/(p_t + P_t)$ converges to zero, the welfare flows are vanishingly small. In order to analyze the impact of increasing P_0 on W, we first note that, from equation (24),

$$\dot{p}_t = \left(\frac{p_t}{\alpha + \beta p_t}\right)^2 p_\varepsilon \tag{39}$$

where recall that $\alpha = P_0 - P_\varepsilon / p_\varepsilon p_0$ and $\beta = 1 + P_\varepsilon / p_\varepsilon$. The welfare can be rewritten as,

$$W = -\int_0^\infty \left(\frac{\alpha + \beta p_t}{p_t^2} \right) \frac{1}{p_\varepsilon} e^{-\lambda t} \left(\dot{p}_t \, dt \right) . \tag{40}$$

Now, from Proposition 2, we have

$$t = \frac{H(p_t) - H(p_0)}{p_\varepsilon} , \tag{41}$$

for $H(p) \equiv 2\alpha\beta \log p + \beta^2 p - \alpha^2/p$. Given that p_t monotonically approaches infinity through time, we can change the integrating variable from t to p in equation (40). We find that

$$W = -\frac{1}{p_\varepsilon} \int_{p_0}^\infty \left(\frac{\alpha + \beta p}{p^2} \right) e^{-\lambda \left(\frac{H(p) - H(p_0)}{p_\varepsilon} \right)} \, dp. \tag{42}$$

The welfare function W depends on the initial precision P_0 of public information only through $\alpha = P_0 - P_\varepsilon / p_\varepsilon p_0$. A marginal increase in the precision of the initial public signal decreases social welfare if and only if $\partial W / \partial \alpha < 0$. Based on this remark we show

Theorem 2 (Welfare Cost of Public Information)**.** *Let $p_\varepsilon = \delta \psi_\varepsilon$, and $P_\varepsilon = \delta \Psi_\varepsilon$, for some $\psi_\varepsilon > 0$ and $\Psi_\varepsilon \geq 0$, as well as some common scaling factor $\delta > 0$. Then for all p_0 and P_0, there exists an $\eta > 0$ such that $0 < \lambda/\delta < \eta$ implies $\partial W / \partial P_0 < 0$.*

Proof. In the appendix □

Theorem 2 tells us that a *marginal* increase in public information reduces welfare, as long as the state is revealed in a sufficiently long time ($1/\lambda$ is high enough), or whenever the amount learned by unit of time, measured by δ, is high enough. In contrast with Morris and Shin (2002), we find that an increase in public information can reduce welfare even when payoffs induce no coordination motives.

The Theorem does not imply, however, that welfare is a monotonically decreasing function of P_0. Indeed, an infinite increase in precision would reveal the state of the world and would clearly improve welfare. By continuity, one might expect that

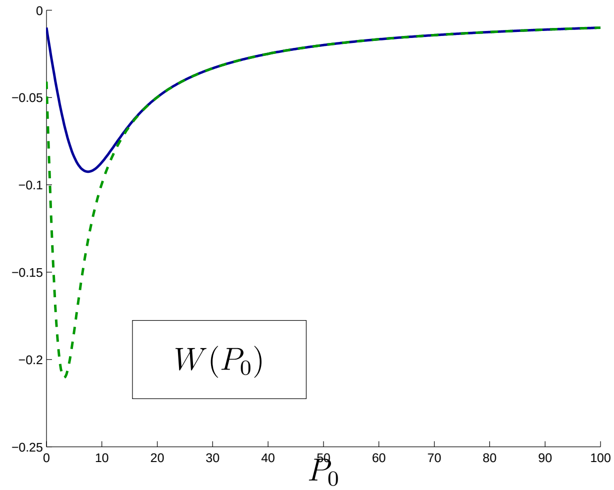


Figure 2: Here welfare is a function of the precision, P_0, of public information. We choose $p_0 = 0.05$, $p_\varepsilon = 1$, and $P_\varepsilon = 0.01$. The blue solid line is for an intensity $\lambda = 0.001$, and the green dotted line is for an intensity $\lambda = 0.005$.

a sufficiently large release of public information also would improve welfare. This intuition is confirmed by the numerical calculation of Figure 2: it shows that welfare is a non-monotonic function of P_0. It first decreases but eventually increases if P_0 is large enough.

In the case where $p_\varepsilon = 0$, and only the public channel is active for learning, a simple application of our results in Proposition 6 shows that an increase in public information always increases welfare, given that it increases the precision of agents' information at all times. This confirms our claim that the channel of information diffusion is fundamental to understanding the social value of public information.

6 Implications and Extensions

In this section we explore two different applications and extensions of our baseline model.

6.1 Technology Diffusion

In many interesting economic situations, it is reasonable to assume that firms and households learn not only from observing others' actions but also from their own experience. Such cases arise naturally, for example, if agents observe their realized payoffs which are noisy signals of the underlying state. One might wonder whether the results obtained in the baseline model generalize to this type of environment. In this subsection, we modify our baseline model to allow for the possibility of learning from experience: a new technology of uncertain productivity has arrived and agents learn about it from their own use, and as before, from observing how others use it.

Suppose that agent i owns a technology that converts labor input, denoted by a_{it}, into units of a final good of price 1. The agent incurs a quadratic loss, $a_{it}^2/2$, from supplying labor. The productivity of this technology is assumed to be the sum of a time-varying and random idiosyncratic component, plus an unknown but constant component. We assume that this constant component, denoted by x, is common across the technologies of all agents in the economy. The accumulated output up to time τ is given by,

$$\int_0^\tau \left(x dt + \frac{1}{\tilde{p}_\varepsilon} d\tilde{\omega}_{it} \right) a_{it} \ ,$$

where the term in brackets is the sum of x and the idiosyncratic productivity process. We assume that $\tilde{\omega}_{it}$ is a standard Wiener process.

For simplicity, we remove public learning from the model: $P_\varepsilon = 0$. However, we still assume that agents privately can observe a signal centered around the average labor supply, with precision p_ε. Unlike before, we now let the agents also observe their current output processes: agents are learning from others, but also from their own use of this technology.

Agents optimally choose their labor decision, $a_{it} = E[x|\mathcal{G}_{it}]$. A linear equilibrium is obtained by decomposing the beliefs at any time into a public and a private forecast. Given that there is no public learning, the precision of the public forecast is a constant,

P_0. The precision of the private forecast, p_t, follows the ODE

$$dp_t = \left(\frac{p_t}{p_t + P_0}\right)^2 p_\varepsilon dt + \tilde{p}_\varepsilon dt \ , \tag{43}$$

where the first term is the same as before, and the second term captures the knowledge that arises from observing your own output. The following proposition characterizes the dynamics of precision:

Proposition 8. *Define*

$$\tilde{H}(p) \equiv \frac{P_0\sqrt{p_\varepsilon}(p_\varepsilon - \tilde{p}_\varepsilon)}{\sqrt{\tilde{p}_\varepsilon}(\tilde{p}_\varepsilon + p_\varepsilon)^2} \arctan\left(\frac{P_0\tilde{p}_\varepsilon + (\tilde{p}_\varepsilon + p_\varepsilon)p}{P_0\sqrt{p_\varepsilon}}\right) + \frac{p_\varepsilon\sqrt{\tilde{p}_\varepsilon}P_0}{(\tilde{p}_\varepsilon + p_\varepsilon)^2} \log\left(P_0^2\tilde{p}_\varepsilon + 2P_0\tilde{p}_\varepsilon p + (\tilde{p}_\varepsilon + p_\varepsilon)p^2\right) + \frac{p}{p_\varepsilon + \tilde{p}_\varepsilon} \ ,$$

then, the solution to equation (43) is implicitly given by $\tilde{H}(p_t) - \tilde{H}(p_0) = t$, and the asymptotic expansion of p_t is

$$p_t = (\tilde{p}_\varepsilon + p_\varepsilon)t - 2\frac{P_0 p_\varepsilon}{p_\varepsilon + \tilde{p}_\varepsilon}\log(t) + R_t,$$

where R_t is some bounded function.

Proof. In the appendix. □

Note that if $\tilde{p}_\varepsilon = 0$, this collapses to our baseline asymptotic expansion for $P_\varepsilon = 0$. As can be seen from the asymptotic expansion, an increase in P_0, will eventually reduce the precision p_t by any finite amount, and a version of Proposition 4 holds.

6.2 Investment

In this subsection we introduce a market setting into our baseline mode and show that the price that arises in equilibrium generates a public signal centered around the average action in the population. In doing this, we also extend the baseline model by introducing a particular payoff externality: the payoffs to any individual will depend on the equilibrium price, which in turn is affected directly by the actions of others.

Consider an economy with two goods: a capital good and a final good. Capital goods are produced at a marginal cost that is increasing in the aggregate capital stock. In particular, if K_t is the aggregate stock of capital good at time t then the cost of producing an extra unit of capital at time t is assumed to be cK_t, for some $c > 0$. The capital good sector is competitive, implying that the price of the capital good at time t is cK_t.

There are a continuum of final-goods producers who accumulate capital until some random time τ. At that point, the final-goods market opens, and the final-goods producers transform their capital stock into final goods using a linear technology with marginal product x. We assume the same information structure as in our baseline model: at time zero, final-goods producers initially are asymmetrically informed about the state x of technology, then learn about it over time through public and private channels to be described below, until the state is realized at time τ.

Let a_{it} be the amount invested at time t by firm i. Agents' profits are realized at the same time as x:

$$\int_0^\tau \left[x a_{it} - cK_t a_{it} - \frac{a_{it}^2}{2}\right] dt,$$

where the second term, $cK_t a_{it}$ is the cost of purchasing the investment and the final term, $a_{it}^2/2$, is a quadratic adjustment cost.

Let $A_t = \int_0^1 a_{it} di$ be the aggregate level of investment at time t. The total capital stock evolves according to

$$dK_t = A_t \, dt + \frac{1}{P_\varepsilon} dW_t \tag{44}$$

where the term $\frac{1}{P_\varepsilon} dW_t$ captures random external demand shocks for capital.

Profit maximization by final-goods producers implies that $a_{it} = E_{it}(x) - cK_t$. Plugging this back into the previous equation, the law of motion for the capital stock becomes

$$dK_t = \left(\int_0^1 E_{it}(x) di - cK_t\right) dt + \frac{dW_t}{\sqrt{P_\varepsilon}} \; .$$

The change in the price of the capital good, cdK_t, is an endogenous public signal about the state of technology, x. Given that the price cK_t of the capital good is known by everyone, observing cdK_t is equivalent to observing

$$dZ_t = \left(\int_0^1 E_{it}(x)di\right) dt + \frac{dW_t}{\sqrt{P_\varepsilon}} \ . \tag{45}$$

Similarly to our baseline model, here we assume that agents also observe a private signal centered around the average action, i.e. A_t up to some idiosyncratic noise,

$$\left(\int_0^1 E_{jt}(x)di - cK_t\right) dt + \frac{d\omega_{it}}{\sqrt{p_\varepsilon}} \ .$$

As before, given that the price cK_t of the capital good is known by everyone, this signal is observationally equivalent to

$$dz_{it} = \left(\int_0^1 E_{it}(x)di\right) dt + \frac{d\omega_{it}}{\sqrt{p_\varepsilon}} \ . \tag{46}$$

The information structure is characterized by equations (45) and (46), which are the equations in our baseline model: therefore, the evolution of precisions in our baseline model also characterizes the evolution of precisions in this market setting, and our previous results in Section 4 concerning the dynamics and comparative statics remain unchanged.

However, in principle, the welfare function is different because of the presence of the cost of capital in the profits of the final-goods producers. Specifically, let us define the welfare function as the *ex ante* profit of a final-goods producer,

$$W = E_0\left[\int_0^\infty e^{-\lambda t}\left(xa_{it} - \frac{a_{it}^2}{2} - cK_t a_{it}\right) dt\right] = \int_0^\infty E_0\left[e^{-\lambda t}\left(xa_{it} - \frac{a_{it}^2}{2} - cK_t a_{it}\right)\right] dt \ .$$

Optimality implies $a_{it} = E_{it}(x) - cK_t$. Substituting into the welfare function, we find

$$\tilde{W} = \int_0^\infty e^{-\lambda t} E_0\left[-(x - E_{it}(x))^2 + \frac{(x - cK_t)^2}{2}\right] \ .$$

The first term inside the brackets is the same welfare flow as in our baseline model. We then need to compute $\int_0^\infty e^{-\lambda t} E_0 (x - cK_t)^2 dt$,

Proposition 9. *The welfare function, $\tilde{W}$, is*

$$\tilde{W} = -\frac{\lambda}{\lambda + 2c} \int_0^\infty \frac{e^{-\lambda t}}{p_t + P_t} \, dt + \frac{1/\bar{P} + c^2(K_0^2 + P_\varepsilon/\lambda)}{\lambda + 2c} \, dt.$$

Proof. In the appendix. □

The last two terms reflect initial conditions that do not matter for our welfare exercise, and the first term is the same as before, except for the multiplicative constant. Thus, the welfare result in Proposition 2, which obtained for the baseline model, also holds in this market setting.

7 Conclusion

This papers analyzes how private information diffuses among a continuum of agents who learn from both public and private observations of each others' actions. We provide conditions such that agents learn the truth along an S-shaped curve, reflecting an information-snowballing effect. We show 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. Further work may address the optimal timing of public information release in this economy.

A Proofs

A.1 Proof of Proposition 1

Equations (15)-(16) follows from a direct application of one-dimensional Kalman filtering formula (see, for instance, Oksendal (1985), pages 85-105). In order to derive equation (19), we multiplying both sides of (15) by P_t. We find

$$\begin{aligned}
& P_t d\hat{X}_t = \Pi_t \left[\left(x - \hat{X}_t \right) dt + \frac{dW_t}{\sqrt{\Pi_t}} \right] \\
\Rightarrow \quad & P_t d\hat{X}_t + dP_t \left(\hat{X}_t - x \right) = \sqrt{\Pi_t} dW_t \\
\Rightarrow \quad & d \left[P_t \left(\hat{X}_t - x \right) \right] = \sqrt{\Pi_t} dW_t \\
\Rightarrow \quad & P_t \left(\hat{X}_t - x \right) - P_0 \left(\hat{X}_0 - x \right) = \int_0^t \sqrt{\Pi_u} dW_u \\
\Rightarrow \quad & \hat{X}_t = \frac{P_0}{P_t} \hat{X}_0 + \left(1 - \frac{P_0}{P_t} \right) x + \frac{1}{P_t} \int_0^t \sqrt{\Pi_u} dW_u
\end{aligned} \tag{47}$$

where the second line follows from the fact that $dP_t = \Pi_t$. 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 $Z_0 = x + W_0/\sqrt{\Pi_0}$ and starting from the common prior that x is normally distributed with a mean of zero and a precision $\bar{P}$. Therefore, $P_0 = \bar{P} + \Pi_0$ and

$$\hat{X}_0 = \left(1 - \frac{\bar{P}}{\bar{P} + \Pi_0} \right) x. \tag{48}$$

Equation (19) then follows from plugging (48) back into (47). Equation (18) follows from exactly the same algebraic manipulation with one difference. In the case of the private forecast, $\hat{x}_{i0}$ and p_0 are the posterior mean and precision at time zero, after observing the private signal $z_{i0} = x + \omega_{i0}/\sqrt{\pi_0}$, and a totally diffuse prior. (That is the precision of the prior is equal to zero.) Therefore, we have that $p_0 = \pi_0$ and $\hat{x}_{i0} = z_{i0}$.

A.2 Proof of Corollary 1

The result follows easily by noticing that, under assumption H, the signals that generated the private and the public forecast are independent conditional on x. Hence, an agent's total forecast will be a linear combination of the public and his private forecasts, with weights given by their respective precisions.

A.3 Proof of Theorem 1

The only thing left to check is hypothesis (H).

We show that the filtration generated by Z and $\tilde{Z}$ are the same, and also that the filtration generated by $(\tilde{Z}, \tilde{z}_i)$ and (Z, z_i) are the same. First, after plugging Equation (22) into Equation (5) and (4), and after using the formula for Π_t and π_t, we find that

$$\begin{aligned} dZ_t &= \frac{P_t}{p_t+P_t}\hat{X}_t\,dt + \frac{p_t}{p_t+P_t}d\tilde{Z}_t \\ dz_{it} &= \frac{P_t}{p_t+P_t}\hat{X}_t\,dt + \frac{p_t}{p_t+P_t}d\tilde{z}_{it} \end{aligned}$$

Since, by construction, $\hat{X}$ is adapted to the filtration generated by $\tilde{Z}$, this implies that the filtration $\mathcal{F}_t$ generated by Z_t is included in the filtration $\tilde{\mathcal{F}}_t$ generated by $\tilde{Z}$. It also shows that the filtration $\mathcal{G}_{it}$ generated by (Z_t, z_{it}) is included in the filtration $\tilde{\mathcal{F}}_t \cup \tilde{\mathcal{F}}_{it}$ generated by $(\tilde{Z}_t, \tilde{z}_{it})$. To show the reverse inclusions, first rearrange the above equation into

$$d\tilde{Z}_t = \frac{p_t+P_t}{p_t}\left(dZ_t - \frac{P_t}{p_t+P_t}\hat{X}_t\,dt\right) \tag{49}$$

$$d\tilde{z}_{it} = \frac{p_t+P_t}{p_t}\left(dz_{it} - \frac{P_t}{p_t+P_t}\hat{X}_t\,dt\right) \tag{50}$$

Now, we also know from Proposition 1 that

$$d\hat{X}_t = \frac{\Pi_t}{P_t}\left[\left(x-\hat{X}_t\right)dt + \frac{dW_t}{\sqrt{\Pi_t}}\right] = \frac{\Pi_t}{P_t}\left(-\hat{X}_t\,dt + d\tilde{Z}_t\right)$$

After plugging equations (49) in the equation above and rearranging, we find:

$$d\hat{X}_t = \frac{\Pi_t}{P_t}\left\{-\left(1+\frac{P_t}{p_t}\right)\hat{X}_t\,dt + \frac{p_t+P_t}{p_t}\,dZ_t\right\}$$

Therefore, $\hat{X}$ is adapted to the filtration generated by Z. Together with (49), this means that $\tilde{Z}$ is adapted to the filtration $\mathcal{F}_t$ generated by Z. Together with (49) and (50) this implies that $(\tilde{Z}, \tilde{z}_i)$ is adapted to the filtration $\mathcal{G}_{it}$ generated by (Z, z_i).

A.4 Proof of Proposition 2

Part (i): Proved in the text.
Part (ii): The dynamics of p_t are given by

$$\dot{p}_t = \left(\frac{1}{\beta + \alpha/p_t}\right)^2 p_\varepsilon$$

This implies that p_t is increasing through time. Letting $\kappa = 1/\beta^2$ if $\alpha < 0$ and $\kappa = 1/(\beta+\alpha/p_0)^2$ otherwise, it follows that

$$\dot{p}_t \geq \kappa p_\varepsilon$$

where $\kappa > 0$, and,

$$p_t = p_0 + \int_0^t \dot{p}_\tau d\tau \geq p_0 + \kappa p_\varepsilon t,$$

and hence p_t converges to infinity as t goes to infinity.
Part (iii): Let's compute the dynamics of the ratio p_t/P_t. Given that $P_t = (\beta-1)p_t + \alpha$, and given

that p_t tends to infinity, it follows,

$$\lim_{t\to\infty}\frac{p_t}{P_t}=\lim_{t\to\infty}\frac{1}{\beta-1+\alpha/p_t}=\frac{1}{\beta-1}=\frac{p_\varepsilon}{P_\varepsilon}$$

The short run dynamics of p_t/P_t are given by

$$\frac{d}{dt}\log\frac{p_t}{P_t}=d\frac{\log p_t-\log P_t}{dt}=\frac{\dot p_t}{p_t}-\frac{\dot P_t}{P_t}$$

Using $p_t=\left(\frac{p_t}{p_t+P_t}\right)^2 p_\varepsilon$, and $\dot P_t=\left(\frac{p_t}{p_t+P_t}\right)^2 P_\varepsilon$, we get that

$$d\frac{\log p_t-\log P_t}{dt}=\left(\frac{p_t}{p_t+P_t}\right)^2\left(\frac{p_\varepsilon}{p_t}-\frac{P_\varepsilon}{P_t}\right)$$

The ratio p_t/P_t is strictly increasing whenever $p_t/P_t<p_\varepsilon/P_\varepsilon$ and strictly decreasing when $p_t/P_t>p_\varepsilon/P_\varepsilon$. So p_t/P_t monotonically approaches $p_\varepsilon/P_\varepsilon$.
Part (iv): Let $y_t\equiv p_t+P_t$. Note that $y_t=\alpha+\beta p_t$ and

$$\dot y_t=\left(\frac{y_t-\alpha}{y_t}\right)^2\frac{p_\varepsilon+P_\varepsilon}{\beta^2}$$

The solution of this differential equation is given by,

$$G(y_t,t)\equiv 2\alpha\log(y_t-\alpha)+y_t-\frac{\alpha^2}{y_t-\alpha}-p_\varepsilon^2/(P_\varepsilon+p_\varepsilon)t=C_0 \tag{51}$$

which after some re-arranging implies part (iv).
Part (v): First note that $\lim_{t\to\infty}y_t=\infty$ which implies by equation (51) that

$$\lim_{t\to\infty}\frac{y_t}{t}=p_\varepsilon^2/(P_\varepsilon+p_\varepsilon). \tag{52}$$

Note that

$$\left(y_t-\frac{p_\varepsilon^2}{p_\varepsilon+P_\varepsilon}t+2\alpha\log(t)\right)=-2\alpha\log\left(\frac{y_t}{t}-\frac{\alpha}{t}\right)+\frac{\alpha^2}{y_t-\alpha}+C_0 \tag{53}$$

is indeed bounded as all function on the right-hand side are, and because y_t/t is also bounded away from zero.

A.5 Proof of Theorem 2

We take the derivative of W with respect to α. We find

$$\begin{aligned}\frac{\partial W}{\partial\alpha}&=-\frac{1}{p_\varepsilon}\int_{p_0}^{\infty}\left\{\frac{1}{p^2}-\frac{\alpha+\beta p}{p^2}\frac{\lambda}{p_\varepsilon}\frac{\partial}{\partial\alpha}\left(H(p)-H(p_0)\right)\right\}e^{-\lambda/p_\varepsilon(H(p)-H(p_0))}\,dp\\&=-\frac{1}{p_\varepsilon}\int_{p_0}^{\infty}\left\{\frac{1}{p^2}-\frac{\alpha+\beta p}{p^2}\frac{\lambda}{p_\varepsilon}\left[-2\alpha\left(\frac{1}{p}-\frac{1}{p_0}\right)+2\beta\log\left(\frac{p}{p_0}\right)\right]\right\}e^{-\lambda/p_\varepsilon(H(p)-H(p_0))}\,dp.\end{aligned} \tag{54}$$

We now integrate the first term $\int_{p_0}^{\infty}(-1/p^2)e^{-\lambda/p_\varepsilon(H(p)-H(p_0))}\,dp$ of integral (54) by part, noting that $-1/p^2 = d/dp\,(1/p - 1/p_0)$. This gives

$$\begin{aligned}
& -\int_{p_0}^{\infty} \frac{1}{p^2} e^{-\lambda/p_\varepsilon(H(p)-H(p_0))}\,dp \\
= & \left[\left(\frac{1}{p}-\frac{1}{p_0}\right) e^{-\lambda/p_\varepsilon(H(p)-H(p_0))}\right]_0^{\infty} + \int_{p_0}^{\infty}\left(\frac{1}{p}-\frac{1}{p_0}\right) H'(p) e^{-\lambda/p_\varepsilon(H(p)-H(p_0))}\,dp \\
= & \int_{p_0}^{\infty}\left(\frac{1}{p}-\frac{1}{p_0}\right) H'(p) e^{-\lambda/p_\varepsilon(H(p)-H(p_0))}\,dp, \qquad (55)
\end{aligned}$$

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

$$\begin{aligned}
& \int_{p_0}^{\infty} \frac{\alpha+\beta p}{p^2}\frac{\lambda}{p_\varepsilon}\left[-2\alpha\left(\frac{1}{p}-\frac{1}{p_0}\right)+2\beta\log\left(\frac{p}{p_0}\right)\right] e^{-\lambda/p_\varepsilon(H(p)-H(p_0))}\,dp \\
= & \int_{p_0}^{\infty} \frac{H'(p)}{H'(p)}\frac{\alpha+\beta p}{p^2}\frac{\lambda}{p_\varepsilon}\left[\frac{2\alpha(p-p_0)}{pp_0}+2\beta\log\left(\frac{p}{p_0}\right)\right] e^{-\lambda/p_\varepsilon(H(p)-H(p_0))}\,dp \\
= & \int_{p_0}^{\infty} \left[\frac{p}{\alpha+\beta p}\right]^2 \frac{\alpha+\beta p}{p^2}\left[\frac{2\alpha(p-p_0)}{pp_0}+2\beta\log\left(\frac{p}{p_0}\right)\right]\frac{\lambda}{p_\varepsilon} H'(p) e^{-\lambda/p_\varepsilon(H(p)-H(p_0))}\,dp \\
= & \int_{p_0}^{\infty} \frac{1}{\alpha+\beta p}\left[\frac{2\alpha(p-p_0)}{pp_0}+2\beta\log\left(\frac{p}{p_0}\right)\right]\frac{\lambda}{p_\varepsilon} H'(p) e^{-\lambda/p_\varepsilon(H(p)-H(p_0))}\,dp, \qquad (56)
\end{aligned}$$

where the third line follows from the fact that $H'(p) = \left[(\alpha+\beta p)/p\right]^2$. Plugging (55) and (56) into the above equation (54) gives:

$$\frac{\partial W}{\partial \alpha} = \frac{1}{p_\varepsilon}\int_{p_0}^{\infty} \Phi(p,p_0)\frac{\lambda}{p_\varepsilon} H'(p) e^{-\lambda/p_\varepsilon(H(p)-H(p_0))}\,dp, \qquad (57)$$

where

$$\Phi(p,p_0) = \left[\frac{1}{p}-\frac{1}{p_0}\right] + \frac{1}{\alpha+\beta p}\left[\frac{2\alpha(p-p_0)}{pp_0}+2\beta\log\left(\frac{p}{p_0}\right)\right].$$

Now since $\Phi(p,p_0) \to -1/p_0$ as $p \to \infty$, there exists some p^* such that $\Phi(p,p_0) < -1/(2p_0)$ for all $p > p^*$. Letting $M^* = \sup_{p\in[p_0,p^*]}\Phi(p,p_0)$, equation (57) implies that

$$\begin{aligned}
\frac{\partial W}{\partial \alpha} = & \frac{1}{p_\varepsilon}\int_{p_0}^{p^*} \Phi(p,p_0)\frac{\lambda}{p_\varepsilon} H'(p) e^{-\lambda/p_\varepsilon(H(p)-H(p_0))}\,dp + \frac{\lambda}{p_\varepsilon}\int_{p^*}^{\infty} \Phi(p,p_0)\frac{\lambda}{p_\varepsilon} H'(p) e^{-\lambda/p_\varepsilon(H(p)-H(p_0))}\,dp \\
\leq & \frac{1}{p_\varepsilon}\left\{M^*\int_{p_0}^{p^*}\frac{\lambda}{p_\varepsilon} H'(p) e^{-\lambda/p_\varepsilon(H(p)-H(p_0))}\,dp - \frac{1}{2p_0}\int_{p^*}^{\infty}\frac{\lambda}{p_\varepsilon} H'(p) e^{-\lambda/p_\varepsilon(H(p)-H(p_0))}\,dp\right\} \\
= & \frac{1}{p_\varepsilon}\left\{M^*\left(1-e^{-\lambda/p_\varepsilon(H(p)-H(p_0))}\right) - \frac{1}{2p_0}e^{-\lambda/p_\varepsilon(H(p)-H(p_0))}\right\}
\end{aligned}$$

The term inside the curly brackets is negative as long as λ is small enough, and we are done with λ.

For the effect of δ, note that given $\psi_\varepsilon > 0$ and Ψ_ε, δ does not affect either α or β, hence H and Φ are unchanged. So δ would only appear as dividing λ, and the result follows.

A.6 Proof of Proposition 8

We can rewrite the solution as,

$$a_0 + (\tilde{p}_\varepsilon + p_\varepsilon)t = p_t + P_0 \frac{\sqrt{p_\varepsilon}(p_\varepsilon - \tilde{p}_\varepsilon)}{\sqrt{\tilde{p}_\varepsilon}(p_\varepsilon + \tilde{p}_\varepsilon)} \arctan\left[\frac{P_0\tilde{p}_\varepsilon + (p_\varepsilon + \tilde{p}_\varepsilon)p_t}{P_0\sqrt{p_\varepsilon \tilde{p}_\varepsilon}}\right] + \\ + \frac{P_0 p_\varepsilon}{p_\varepsilon + \tilde{p}_\varepsilon} \log\left[P_0^2\tilde{p}_\varepsilon + 2P_0\tilde{p}_\varepsilon p_t + (p_\varepsilon + \tilde{p}_\varepsilon)p_t^2\right] \ .$$

Given that p_t goes to ∞, it follows that $p_t/t \to (p_\varepsilon + \tilde{p}_\varepsilon)$ as t goes to ∞. Now,

$$p_t - (p_\varepsilon + \tilde{p}_\varepsilon)t + \frac{P_0 p_\varepsilon}{p_\varepsilon + \tilde{p}_\varepsilon}\log(t^2) = a_0 - P_0 \frac{\sqrt{p_\varepsilon}(p_\varepsilon - \tilde{p}_\varepsilon)}{\sqrt{\tilde{p}_\varepsilon}(p_\varepsilon + \tilde{p}_\varepsilon)} \arctan\left[\frac{P_0\tilde{p}_\varepsilon + (p_\varepsilon + \tilde{p}_\varepsilon)p_t}{P_0\sqrt{p_\varepsilon \tilde{p}_\varepsilon}}\right] \\ - \frac{P_0 p_\varepsilon}{p_\varepsilon + \tilde{p}_\varepsilon} \log\left[P_0^2\tilde{p}_\varepsilon/t^2 + 2P_0\tilde{p}_\varepsilon p_t/t^2 + (p_\varepsilon + \tilde{p}_\varepsilon)(p_t/t)^2\right] \ .$$

Note that, since $\arctan(\cdot)$ and p_t/t are bounded, and since p_t/t is bounded away from zero, the right-hand side is indeed a bounded function.

A.7 Proof of Proposition 9

Let $\mu_t = x - cK_t$, and $\gamma_t = p_t/(p_t + P_t)$, then

$$d\mu_t = c\left[(1-\gamma_t)\left(x - \hat{X}_t\right) - \mu_t\right]dt - c\frac{dW_t}{\sqrt{P_\varepsilon}}$$

and, by Ito's lemma

$$\begin{aligned} & d\left(\mu_t^2\right) = 2\mu_t d\mu_t + d\mu_t d\mu_t = 2\mu_t d\mu_t + \frac{c^2}{P_\varepsilon}dt \\ \Rightarrow\ & d\left(\mu_t^2\right) = 2\mu_t c\left(1-\gamma_t\right)\left(x - \hat{X}_t\right)dt - 2c\mu_t^2 dt - \frac{2\mu_t c}{\sqrt{P_\varepsilon}}dW_t + \frac{c^2}{P_\varepsilon}dt \\ \Rightarrow\ & \mu_t^2 = \mu_0^2 + \int_0^t 2\mu_s c\left(1-\gamma_s\right)\left(x - \hat{X}_s\right)ds - 2c\int_0^t \mu_s^2 ds - \int_0^t \frac{2\mu_s c}{\sqrt{P_\varepsilon}}dW_s + \frac{c^2}{P_\varepsilon}t. \end{aligned}$$

Taking expectations on both sides, we obtain

$$E_0\mu_t^2 = E_0\mu_0^2 + \frac{c^2}{P_\varepsilon}t + 2c\int_0^t \left(1-\gamma_s\right)E_0\mu_s\left(x - \hat{X}_s\right)ds - 2c\int_0^t E_0\mu_s^2 ds \tag{58}$$

To compute $E_0\left[\mu_t\left(x - \hat{X}_t\right)\right]$, note the following:

$$\begin{aligned} E_0\left[\mu_t\left(x - \hat{X}_t\right)\right] &= E_0\left[\left(x - cK_t\right)\left(x - \hat{X}_t\right)\right] \\ &= E_0\left[\left(x - \hat{X}_t\right)\left(x - \hat{X}_t\right)\right] + E_0\left[\left(\hat{X}_t - cK_t\right)\left(x - \hat{X}_t\right)\right], \end{aligned}$$

but, by definition of the conditional expectation $\hat{X}_t$, the residual $x - \hat{X}_t$ is orthogonal to $\hat{X}_t$ and K_t. Therefore, the second term is zero and we find that

$$E_0\left[\mu_t\left(x - \hat{X}_t\right)\right] = E_0\left[\left(x - \hat{X}_t\right)\left(x - \hat{X}_t\right)\right] = \frac{1}{P_t}.$$

Let $\bar{\mu}_t \equiv E_0 \mu_t^2$. Then plugging the above finding back into (58), and keeping in mind that $\gamma_t = p_t/(p_t + P_t)$, we obtain

$$\bar{\mu}_t = \bar{\mu}_0 + 2c \int_0^t \frac{1}{p_s + P_s} ds - 2c \int_0^t \bar{\mu}_s ds + \frac{c^2}{P_\varepsilon} t \ .$$

Differentiating,

$$\dot{\bar{\mu}}_t = \frac{2c}{p_t + P_t} + \frac{c^2}{P_\varepsilon} - 2c\bar{\mu}_t \ .$$

The above differential equation can now be solved for $\bar{\mu}_t$:

$$\begin{aligned}
\bar{\mu}_t &= e^{-2ct}\bar{\mu}_0 + \int_0^t e^{-2c(t-s)} \left(\frac{2c}{p_s + P_s} + \frac{c^2}{P_\varepsilon} \right) ds \\
&= e^{-2ct}\bar{\mu}_0 + \frac{c}{2P_\varepsilon}\left(1 - e^{-2ct}\right) + \int_0^t e^{-2c(t-s)} \frac{2c}{p_s + P_s} ds.
\end{aligned}$$

Note that, $\lim_{c \to 0} A_t = 0$. The welfare function is, then,

$$\tilde{W} = \int_0^\infty e^{-\lambda t} \left(-\frac{1}{p_t + P_t} + e^{-2ct}\bar{\mu}_0 + \int_0^t e^{2c(\tau - t)} \left(\frac{2c}{p_\tau + P_\tau} + \frac{c^2}{P_\varepsilon} \right) d\tau \right) dt \tag{59}$$

Now switching the order of the two integrals, we find

$$\int_0^\infty e^{-\lambda t} \int_0^t e^{2c(\tau - t)} \frac{2c}{p_\tau + P_\tau} d\tau \, dt = \int_0^\infty e^{-\lambda \tau} \frac{2c}{r + 2c} \frac{1}{p_\tau + P_\tau} \, d\tau.$$

Plugging back and integrating the other terms, we find:

$$\tilde{W} = -\frac{\lambda}{\lambda + 2c} \int_0^\infty \frac{e^{-\lambda t}}{p_t + P_t} dt + \frac{A_0}{\lambda + 2c} + \frac{c}{2P_\varepsilon} \frac{2c}{\lambda(\lambda + 2c)}.$$

Now recall that $\bar{\mu}_0 = E_0[(x - cK_0)^2] = E_0[x^2 - 2cK_0 x + c^2 K_0^2] = 1/\bar{P} + c^2 K_0^2$, and we are done.

References

Amador, Manuel and Pierre-Olivier Weill, "Learning by Matching," 2006. 9

_ and _ , "Learning from Prices: Central Bank Comunication and Welfare," 2007. 9

Angeletos, George-Marios and Alessandro Pavan, "Efficient Use of Information and Welfare Analysis in Economies with Complementarities and Asymmetric Information," 2005. 5, 20

Araujo, Luis and Andrei Shevchenko, "Price Dispersion, Information and Learning," *Journal of Monetary Economics, Forthcoming*, 2006. 6

_ and Braz Camargo, "Information, Learning, and the Stability of Fiat Money," *Journal of Monetary Economics, Forthcoming*, 2006. 6

Bala, Venkatesh and Sanjeev Goyal, "Learning from Neighbours," *Review of Economic Studies*, 1998, *65*, 595–621. 6

Banerjee, Abhijit and Drew Fudenberg, "Word-of-mouth learning," *Game and Economic Behavior*, 2004, *46*, 1–22. 6

Banerjee, Abhijit V, "A Simple Model of Herd Behavior," *The Quarterly Journal of Economics*, August 1992, *107* (3), 797–817. 5, 22

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch, "A Theory of Fads, Fashion, Custom, and Cultural Change in Informational Cascades," *Journal of Political Economy*, October 1992, *100* (5), 992–1026. 5

Blouin, Max R. and Roberto Serrano, "A Decentralized Market with Common Values Uncertainty: Non-Steady States," *Review of Economic Studies*, 2001, *68*, 323–346. 6

Burguet, Roberto and Xavier Vives, "Social Learning and Costly Information Acquisition," *Economic Theory*, 2000, *15*, 185–205. 6

Chamley, Christophe, *Rational Herds*, Cambridge: Cambridge University Press, 2004. 3, 9, 22

DeMarzo, Peter, Dimitri Vayanos, and Jeffrey H. Zwiebel, "Persuasion Bias, Social Influence, and Uni-Dimensional Opinions," *Quaterly Journal of Economics*, 2003, *118*, 909–968. 6

Duffie, Darrell and Gustavo Manso, "Information Diffusion in Large Population," 2006. 6

Fernández, Raquel, "Culture as Learning: The Evolution of Female Labor Force Participation over a Century," 2007. New York University. 3

Fogli, Alessandra and Laura Veldkamp, "Nature or Nurture? Learning and Female Labor Force Participation," 2007. Federal Reserve Bank of Minneapolis Staff Report 386. 3

Gale, Douglas and Shachar Kariv, "Bayesian Learning in Social Networks," 2003. Working Paper, New York University. 6

Green, Edward J., "Eliciting Traders' Knowledge in "Frictionless" Asset Markets, Working Paper," 1991. 6

Hellwig, Christian, "Heterogeneous Information and the Welfare Effects of Public Information Disclosures," 2005. 5, 20

Jovanovic, Bojan and Yaw Nyarko, "A Bayesian Learning Model Fitted to a Variety of Empirical Learning Curves," *Brooking Papers: Microeconomics*, 1995, pp. 247–304. 3

Judd, Kenneth L., *Numerical Methods in Economics*, Boston: MIT Press, 1999. 19

Katzman, Brett, John Kennan, and Neil Wallace, "Output and Price Level Effects of Monetary Uncertainty in a Matching Model," *Journal of Economic Theory*, 2003, *108*, 217–255. 6

Kircher, Philipp and Andrew Postlewaite, "Strategic Firms and Endogenous Consumer Emulation, Working Paper," 2006. 6

Kiyotaki, Nobuhiro and Randall Wright, "On Money as a Medium of Exchange," *Journal of Political Economy*, 1989, *97*, 927–954. 6

Morris, Stephen and Hyun Song Shin, "The Social Value of Public Information," *American Economic Review*, 2002, *92*, 1521–1534. 4, 20, 25, 26

_ and _ , "Central Bank Transparency and the Signal Value of Prices," *Brookings Papers on Economic Activity*, 2005, *2*, 1–66. 5, 24

Oksendal, Bernt, *Stochastic Differential Equations*, Berlin: Springer, 1985. 33

Protter, Philip, *Stochastic Integration and Differential Equations*, New York: Springer-Verlag, 1990. 7

Smith, Lones and Peter Norman Sorensen, "Rational Social Learning by Random Sampling," 2005. Working Paper, University of Michigan. 6

Trejos, Alberto and Randall Wright, "Search, Bargaining, Money, and Prices," *Journal of Political Economy*, 1995, *103* (1), 118–141. 6

Vives, Xavier, "How Fast do Rational Agents Learn?," *Review of Economic Studies*, 1993, *60*, 329–347. 2, 5

— , "Learning from others: a Welfare Analysis," *Games and Economic Behavior*, 1997, *20*, 177–200. 2, 5, 23

Wallace, Neil, "Short-Run and Long-Run Effects of Changes in Money in a Random-Matching Model," *Journal of Political Economy*, 1997, *105*, 1293–1307. 6

Wolinsky, Asher, "Information Revelation in a Market with Pairwise Meetings," *Econometrica*, 1990, *58*, 1–23. 6