Unjamming the Public Signal: optimal public communication decisions in learning environments

Marcello Miccoli*

Version 1.1 Nov 2010
PRELIMINARY AND INCOMPLETE

Abstract

In a setting of diffused imperfect knowledge among agents about the fundamentals of the economy, a benevolent public authority learns about the fundamentals by observing agents' actions. In any period the public authority decides how much of its knowledge to communicate in a public signal. While public communication improves current welfare, it may hamper the authority's learning process. The public signal crowds out private information from agents' actions, which is the valuable source of learning for the public authority, leading to future less precise public communication: this is the jamming effect of public communication. Simulations of the model show that the welfare maximizing communication strategy over time can deviate from full disclosure of the public authority's information. In particular, it can be optimal to *garble* the public message, that is, add noise to the public signal.

*Stanford University, 344 Serra Mall, Stanford, CA 94305 (miccoli@stanford.edu). Thanks to Manuel Amador, Aaron Bodoh-Creed, Carlos Lever, Bob Hall, Nir Jaimovich and Juuso Toikka for useful discussions and suggestions. All remaining errors are mine. I gratefully acknowledge financial support from the B.F. Haley and E.S. Shaw Fellowship.

1 Introduction

Since the seminal paper of Morris and Shin (2002), *Social Value of public infomartion* economists have been debating the welfare benefits of public communication. Morris and Shin, in this and in a companion article (Morris and Shin (2005), have argued that public communication can be welfare decreasing since it implies over-reliance on public information with respect to what is efficient. They apply their argument to Central Banks communication and derive the implication that transparency may actually be welfare decreasing.

The message spurred a sequence of paper on the subject. Svensson (2006) criticizes the results of Morris and Shin (2002), saying that for plausible values of parameters, actually their model is pro transparency, not con. Angeletos and Pavan (2007) have shown that Morris and Shin's hypothesis of complementarities in actions among agents is indeed fundamental for having negative effects of public information. However in a different context, Amador and Weill (2010), have shown that public information can have negative welfare effects because it can reduce the informational value of prices.

Morris and Shin (2002)'s model and following analyses have been static, and characterized by an exogenous source of public information. This approach misses what are two fundamental characteristics of public information. The first characteristics is that information flow is usually bi-directional between public authorities and agents, thus public information is never purely exogenous. If we think about the operate of a Central Bank, for instance, their source of information are economic data, which are decisions of agents and as such incorporate their private information and public information about the state of the world.

The second characteristics of information is that timing, that is, when information is available to agents, matters. Agents in the real world take decisions continuosly, and availability of information in some periods and not in other is welfare relevant. Consider a firm trying to choose production according to information about the business cycle. If investment choices are irreversible, then whether information about the businnes cycle is available or not at the time of the decision becomes a fundamental determinant of the profitability of choices.

This work will try to approach the problem of the social value of public communication by tackling the two above mentioned characteristics. We build a model in which agents have diffused information about an evolving state of the world. A public authority learns about the evolving state of the world by observing with measurement error an aggregate statistics of the agents actions and can communicate to the the agents with a public signal.

There is therefore an informational flow from the agents to the public authority through aggregate statistics, and from the public authority to the agents through the public signal.

The model tackles then the timing characteristics with the normative questions: when and how much of the public knowledge should the public authority reveal every period? The analysis will show that precise communication in any period might not be optimal. Optimal communication plan involve periods of no communication or *garbling*, that is, the knowledge of the public authority is communicated with noise.

The reason for the result lies in the circularity of information flow. Public communication at any point in time crowds out private information from agents' actions, therefore reducing the informativeness of aggregate statistics to the agency. This effect can be vicious, leading to an actual decrease in the quantity of information future public communication can reveal. This is the *jamming effect*: public communication blocks the arrival of new information to the public authority. By not communicating or garbling in some periods the agency increases the informative value of the aggregate statistics, and can communicate with more precision in the future. The choice of public communication therefore entails computing the marginal cost of not revealing precisely today with the marginal benefit of revealing more precise information in the future.

The implication of the model is that transparency might backfire. In particular when deciding a public communication strategy it is important to realize that the source of public information, how this is collected, has to be taken into account in order to avoid negative feedbacks of communication.

The model is abstract, but it can be applied to a variety of situations. We can think of the decision actors as firms trying to forecast the business cycle. Each firm has to choose production choices given their information about the cycle. The public authority, in this case the Federal Reserve or the Bureau of Labor Statistics decide then when and how much of its information to release. Miccoli (2010) studies the informational flow between financial institutions choosing portfolio compositions and a public authority set up to watch over their decisions.

However the model can also be applied to non-economic situations. In May 2000, the National Highway Traffic Safety Administration contacted the Ford company regarding a problem of tires exploding with some their models and causing rollovers of the vehicles involved. During the events, it was not clear which models of tires and vehicles were affected. This is clearly a case in which agents' actions, the driving of Ford vehicles, affected the collection of information and the learning of the facts, which model of cars and tires are subject to explosion. The prediction of the analysis in this work thus imply that extensive public communication about the few known facts could have led to a decrease in the accident rate, however it might have prevented effective learning about the car models and types of tires involved. Instead, no or sparse public communication, even though it decreased safety during the periods of silence, would have led to accumulation of observations and the discovery of the precise source of the tires explosion.[1]

The structure of the paper is the following. Next section briefly discuss the related literature. Then section (3) presents the model and discusses the source of the trade-off between present and future public communication. Section (4) set up the dynamic programming for the model, explores numerical solutions, and provides an interpretation of the results with an easy two periods model. Section (5) concludes.

2 Related Literature

Apart from the works already cited in the introduction, this work is connected to the literature on speed of convergence of beliefs, started by Vives (1993) and expanded by Amador and Weill (2009). Vives (1993) shows that the presence of a public signal of an aggregate of agents' actions decreases the speed of converge of rational expectations to the truth. Amador and Weill (2009) add to the public signal a private endogenous signal, coming from mouth-to-mouth discussions among agents. They show that the releasing a public signal at the beginning of the interaction among agents can lead to welfare losses. What distinguishes this work from these previous two papers is the question. In Vives and Amador and Weill's work the public signal is always present, and as such, it cannot be modified. Their analysis captures the informational flow but does not capture the timing question.

Vives (1997) can be considered a first attempt at moulding public information in such a way to increase welfare. The mechanism in Vives (1997) and in this work is similar. As in Vives' analysis, also in this model the public signal creates an informational externality: agents rely too much with respect to what is optimal, and this creates under-production of future information. However in Vives' work the public signal is thought as coming from market interactions, hence the object of analysis are cooperative approaches to information use that could increase welfare. Here the public signal is the explicit decision of an actor, and through its decisions, it tries to solve the informational externality.

[1] I thank Mitchell Polinsky for providing me with this example.

3 Model

The model has two actors: agents and a public authority, which we will call as the agency. This section explains objectives and choices of both of them, and how the information flow between them depends on their choices.

3.1 Agents

Time is infinite, $t = 0, 1, 2, \ldots$, and there is a continuum $i \in [0, 1]$ of agents in each period t. Agents are short lived, i.e. at the beginning of every period t there is a new generation of individuals which takes an action and dies at the end of the period.[2] The economy is characterized by a state θ_t which follows a random walk process determined by:

$$\theta_t = \theta_{t-1} + \epsilon_t \tag{1}$$

where $\epsilon \sim N(0, \sigma^2)$. θ_t is not known by the agents, however each agent i in every period t wants to choose an action to be as close as possible to the unknown parameter θ_t. We are going to assume the simplest possible functional form that delivers this behaviour description. Hence the objective of the agent is given by:

$$\min_{a_{i,t}} \mathbb{E}_{i,t} \left[(a_{i,t} - \theta_t)^2 \right] \tag{2}$$

where the $E_{i,t}[\cdot]$ defines the expections operator conditional on the information available to agent i at time t about θ_t. Given the quadratic objective the optimal choice is to set:

$$a_{i,t} = \mathbb{E}_{i,t}[\theta_t] \tag{3}$$

3.1.1 Information available to Agents

In every period t agents have an uninformative prior over θ_t. But every agent i in period t has at his disposal two types of information about θ_t. The first type of information is private, coming from a private signal observed only by agent i at the beginning of t:

$$\theta_{i,t} = \theta_t + \eta_{i,t} \tag{4}$$

[2] This is equivalent to assuming that individuals have no memory, i.e. they don't carry information from the past. The assumption of memoryless agents is done for tractability of the problem. Analysis without this assumption can be done only in much less general settings. See for instance Miccoli (2010).

where $\eta_{i,t} \sim N(0, p^{-1})$ is identically and independently distributed (i.i.d.) in the cross section of agents and across time. p, the inverse of the variance, is defined as the precision of the signal $\theta_{i,t}$ about θ_t. As it is customary in the literature, we will measure information in units of precision.

In addition to the private information, the agents has available public information, coming from communication common to all agents. The structure of the public signal is endogenous in the model, and depends on the information collection process of the agency explained below. For now we assume the following conjecture, which will be proved true later.

Conjecture 1. *The public signal in period t, m_t is, conditional on θ_t, normally distributed, with $\mathbb{E}[m_t|\theta_t] = 0$ and $\mathbb{V}ar[m_t|\theta_t] = P_t^{-1}$. Moreover $m_t|\theta_t$ is independent of $\eta_{i,t}$ for all t, for all i.*

With the public and private information available at their disposal, the agents update their uninformative prior over θ_t. Given the normality assumptions and indipendence of signals, standard Bayesian updating implies that the posterior mean about θ_t will be a convex combination of private and public information, with weights given by the relative precision of signals, and the posterior variance is the inverse of the sum of the precisions of the two signals:

$$\mathbb{E}_{i,t}[\theta_t] = \frac{p}{p+P_t}\,\theta_{i,t} + \frac{P_t}{p+P_t}\,m_t$$
$$\mathbb{V}ar_{i,t}[\theta_t] = \frac{1}{p+P_t}.$$

Given the FOC (3) the agents' action will reflect their mean posterior about θ_t. Note that the variance of the posterior is common to all agents, since the agents' information sets differ by the realization of the private signal $\theta_{i,t}$, but this signal is assumed to have common precision across the agents.

3.2 Agency

In the economy there is a public agency, with the authority to observe aggregate of the agents' actions and to send public messages to all agents about the unknown parameter θ_t. The object of the agency is to maximizie the *ex-ante aggregate welfare* of the agents. Both the content of the public signal and the object of the agency will be defined in the next two paragraphs.

3.2.1 Collection of Information by the Agency

In period $t = 0$, the Agency has an uninformative prior over θ_0. However at the end of every period t the Agency has the authority to observe, with measurement error, an aggregate of the actions of the agents. At the end of the period t it receives a signal:

$$\tilde{S}_t = \int a_{i,t}\, di + \nu_t \tag{5}$$

where ν_t is the measurement error, distributed normally with 0 mean and precision ψ_ν, i.i.d across time, and independent of $\eta_{i,t}$, for all i, t. Substituting in the optimal action of the agents, and assuming the convention that the laws of large number holds, so that $\int \eta_{i,t}\, di = 0$, we obtain that $\tilde{S}_t = \frac{p}{p+P_t}\theta_t + \frac{P_t}{p+P_t}m_t + \nu_t$. The Agency uses this signal in order to update its information about θ_t. However, the only relevant part for the learning of the Agency is the aggregate of private information of the agents, since m_t, being sent by the same Agency, is already part of its information set. Therefore the signale $\tilde{S}_t$ is observationally equivalent to a signal:

$$S_t \equiv \frac{p+P_t}{p}\left[\tilde{S}_t - \frac{P_t}{p+P_t}m_t\right] = \theta_t + \frac{p+P_t}{p}\ \nu_t \tag{6}$$

The Agency uses the signal S_t and its prior on θ_t in order to form a posterior on θ_t. Given a history of signals $S^{t-1} = \{S_1, S_2, \ldots, S_{t-1}\}$, all normally distributed, let $\mu_{A,t}$ and α_t be respectively the mean and variance of the Agency's prior at time t of θ_t, conditional on the observation of history S^{t-1}. Then the update of the agency at time t is determined by:

$$\mathbb{E}[\theta_t|S^t] = \frac{\alpha_t^{-1}}{\alpha_t^{-1} + \left(\frac{p}{p+P_t}\right)^2 \psi_\nu}\ \mu_{A,t} + \frac{\left(\frac{p}{p+P_t}\right)^2 \psi_\nu}{\alpha_t^{-1} + \left(\frac{p}{p+P_t}\right)^2 \psi_\nu}\ S_t \tag{7}$$

$$\mathbb{V}ar[\theta_t|S^t] = \frac{1}{\alpha_t^{-1} + \left(\frac{p}{p+P_t}\right)^2 \psi_\nu} \tag{8}$$

These define the posterior of the Agency on θ_t at the end of period t. This estimate will be carried on into next period $t+1$. We need to transform it however into an estimate of θ_{t+1}. Given that θ_t follows a martingale, then $\mu_{A,t+1} = \mathbb{E}[\theta_{t+1}|S^t] = \mathbb{E}[\theta_t|S^t]$, and

$$\alpha_{t+1} = \mathbb{V}ar[\theta_{t+1}|S^t] = \mathbb{V}ar[\theta_t + \epsilon_t|S^t] = \frac{1}{\alpha_t^{-1} + \left(\frac{p}{p+P_t}\right)^2 \psi_\nu} + \sigma^2 \tag{9}$$

3.2.2 Diffusion of Information by the Agency

We are now ready to verify the conjecture about the structure of the public signal. At the beginning of every period t, conditional on observed history S^t, the agency sends a public signal with its mean belief about θ_t:

$$m_t = \mu_{A,t} + \phi_t \tag{10}$$

where $\phi_t \sim N(0, \Phi_t^{-1})$ independent across time and of everything else. Φ_t, the precision of the signal of the agency conditional on its mean beliefs, is the object of choice of the agency. The agency can continuously change the precision in order to reveal more or less information at its disposal. If $\Phi_t \to 0$ then the agency is sending no message, if $\Phi_t \to \infty$, the agency is exactly revealing its mean belief. For any value in between the agency is *garbling*, that is, sciently obfuscating its knowledge in the public message.

Given the properties of the Bayesian updating, $\mu_{A,t}$ is a martingale, and so $\mu_{A,t}$ is an unbiased signal of θ_t.[3] Since by assumption ϕ_t has mean zero, then m_t is an unbiased signal of θ_t. Note that m_t is function only of the measurement error ν_t and of the noise added by the agency ϕ_t. Hence it is orthogonal to the noise present in the private signal $\theta_{i,t}$. Also, given the independence assumption of ϕ_t, the variance of m_t conditional on θ_t is determined by the sum of the variance of the belief of the agency α_t, and the variance of the noise, $1/\Phi_t$, the agency decides to tack on its message. Therefore by letting

$$\frac{1}{\alpha_t + 1/\Phi_t} \equiv P_t \tag{11}$$

we have verified the conjecture on the structure of the public signal.

Equation (9) and (11) fully specify the evolution of the precision of public knowledge among the agents given the choice of communication of the agency. Note that if $\Phi_t \to \infty$ then $P_t \to 1/\alpha_t$, that is, if the agency is perfectly communicating its mean beliefs, the precision of public knowledge is the same as the precision of the Agency. Conversely, if $\Phi_t \to 0$, then $P_t \to 0$, that is, when the agency is not communicating there is no public knowledge.

[3] On the properties of Bayesina updating, see for instance Vives (2008).

3.2.3 Welfare Objective

The Agency has the objective to maximize the *aggregate ex-ante* welfare W of the agents, where each future generation is discounted at rate $\beta < 1$:

$$W = -\sum_{t=0}^{\infty} \beta^t \int_0^1 \mathbb{E}_{i,t}\left[(a_{i,t} - \theta_t)^2\right] \, di \tag{12}$$

Note that agents were minimizing a welfare loss function. Instead of defining the problem over the welfare loss we defined welfare to be the negative of the welfare loss. By substituting in W the optimal choice $a_{i,t} = \mathbb{E}_{i,t}[\theta_t]$, we obtain $\mathbb{E}_{i,t}\left[(\mathbb{E}_{i,t}[\theta_t] - \theta_t)^2\right] = \mathbb{V}ar_t[\theta_t] = \frac{1}{p+P_t}$. The ex-ante indirect utility function is just the posterior variance of beliefs of the agents about θ_t, as argued before common across agents. The result that ex-ante we need only to care about precision of information and not realization of signals is a consequence of the linear-quadratic setting with normality assumption. Generically it need not be the case.

The problem the agency wants to solve is:

$$\max_{\{\Phi_t\}_{t=1}^{\infty}} \sum_{t=1}^{\infty} \beta^t \left[-\frac{1}{p+P_t}\right] \tag{13}$$

$$\text{s.t.} \quad \alpha_{t+1} = \frac{1}{\alpha_t^{-1} + \left(\frac{p}{p+P_t}\right)^2 \psi_\nu} + \sigma^2$$

$$P_t = \frac{1}{\alpha_t + 1/\Phi_t}$$

$$\alpha_1 = 1/\psi_\nu$$

Note that in period $t = 0$, there cannot be meaningful communication on the part of the Agency, since both agents and Agency share the same uninformative prior. It's only after $t = 0$ that the Agency has different information than the agents, since it has observed S_0, therefore choosing how much the agency wants to reveal of her knowledge becomes a meaningful problem. Given that in $t = 0$ there is no public communication, the signal to the agency is $S_0 = \theta_0 + \nu_0$, therefore the variance of the belief of the agency at the beginning of period $t = 1$ is determined only by this signal, hence $\alpha_1 = 1/\psi_\nu$, pinning down the initial condition for the infinite horizon sequential problem.

3.2.4 The trade-off between providing more information today or tomorrow

Public communication is welfare improving in any period: if the agency is communicating with high precision in period t, agents have more knowledge about the unknown parameter θ_t, therefore they will choose more precise actions, increasing their welfare. This is shown by the ex-ante welfare of the agents being increasing in P_t.

However note that, the more precise the signal the agency sends in period t, the less the agency will know about θ_{t+1} in period $t+1$, since by equation (9), $\frac{\partial \alpha_{t+1}}{\partial P_t} > 0$. This is so because agents will effectively use the more precise public signal and will rely heavily on it. However learning for the agency is higher when the agents only use their private information, since the public message is already in the information sets of the agency. Remember from equation (6) that the precision of signal the agency receives, conditional on θ_t is given by $\left(\frac{p}{p+P_t}\right)^2 \psi_\nu$. As we see, higher P_t implies a less precise signal to the agency, since agents use less of their private knowledge, which is the source of new information for the agency. Therefore more precise public communication today implies less precise information arriving to the agency next period.

This effect has a simple implication: when the agency decides how much to communicate tommorow, for any communication strategy it wants to follow, it will be able to provide less information. This is clear since $P_{t+1} = \frac{1}{\alpha_{t+1}+1/\Phi_{t+1}}$, and therefore $\frac{\partial P_{t+1}}{\partial \alpha_{t+1}} < 0$ for any Φ_{t+1}.

But then communicating more precisely in any period t does not have a negative effect on communication only in $t+1$. Given that the agency has memory, for any communication plan it decides in the future, communicating today will always imply less information in the future, and therefore less precise public communication in the future, that is:

$$\frac{\partial P_{t+s}}{\partial \Phi_t} < 0, \quad \forall s \geq 1 \tag{14}$$

The trade-off between the present public communication and future public communication is clear. By sending a very precise signal today the agency helps current individuals, however it prevents precise public communication tomorrow and therefore reduces welfare for agents tomorrow. When optimally choosing the agency has to trade-off the welfare benefits to current generation of agents to the welfare cost to future agents.

Whether or not the agency would like in some period t not to send a precise public signal depends on the marginal benefit to current generation of agents of one additional unit of precision in public signal versus the marginal cost to future generations. Let $w(P_t) = -\frac{1}{p+P_t}$

define the ex-ante aggregate welfare of agents in generation t. If

$$\frac{\partial w(P_t)}{\partial \Phi_t} > - \sum_{s=t+1}^{\infty} \beta^{s-t} \frac{\partial w(P_s)}{\partial \Phi_t} \tag{15}$$

then the agency would like to increase the precision of the public message. However it is easy to check that $w(P_t)$ is concave in P_t, therefore there are decreasing marginal returns in units of precision of the public signal. Conversely the costs will be convex, i.e., there are increasing marginal cost of precision of public information in t. This implies that there can be a limit to the public information the agency wants to provide in any period t. If equation (15) holds with equality, then there is a positive level of noise in public signal which is optimal. The agency will find it optimal to garble its message today in order to preserve efficient learning and effecive future public communication.

The trade-off implied by public communication today or tomorrow does not depend on the asssumption of a new generation of agents being born in any period. Miccoli (2010) shows that this effect is also present when agents are long lived and have memory. What is important is the circularity of the information flow between agents and the agency. Note also that the result does not depend on any not optimal decision on the part of the agents. When agents receive the public signal, each of them is individually too small to affect on its own the precision of information collected in the future by the agency. Therefore they optimally fully rely on the public information. However, relying on the public signal implies that the agency is not able to fully learn, hence there is under-production of future information.

4 The optimal communication plan

4.1 Dynamic Programming

In order to solve the problem we are going to use a dynamic programming approach and then numerical simulations. It is clear that the prior precision of the knowledge about θ_t the agency has at any point in time α_t can be used as state variable: it summarizes all past decisions of communications of the agency and tracks the evolution of θ_t.

By rearranging constraint (9) and substing it into the objective function, we can rewrite the problem only in terms of the state:

$$\max_{\{\alpha_{t+1}\}_{t=1}^{\infty}} \sum_{t=1}^{\infty} \beta^t \left\{ - \left[\frac{1}{p^2 \psi_\nu} \left(\frac{1}{\alpha_{t+1} - \sigma^2} - \frac{1}{\alpha_t} \right) \right]^{1/2} \right\} \tag{16}$$

$$\text{s.t.} \quad \sigma^2 + \frac{1}{1/\alpha_t + \psi_\nu} \leq \alpha_{t+1} \leq \sigma^2 + \frac{1}{1/\alpha_t + \left(\frac{p}{p+1/\alpha_t} \right)^2 \psi_\nu} \tag{17}$$

$$\alpha_1 = 1/\psi_\nu$$

Before setting up the dynamic programming problem, we can already note some properties that the optimal solution will surely have, since they are implied by the constraint set.

Proposition 1. *In any optimal solution $\{\alpha^*_{t+1}\}_t$, for any t, $\alpha^*_{t+1} > \sigma^2$*

Proof. Let $l(\alpha_t)$ and $h(\alpha_t)$ define respectively the lower and higher bound of α_{t+1} from equation (17). As $\alpha_t \to 0$, then both $l(\alpha_t)$ and $h(\alpha_t)$ tend to σ^2, reaching their lower bound. Hence for any α_t, $\alpha_{t+1} > \sigma^2$ for all t. This will imply in particular that the optimal solution α^*_{t+1} has to be greater than σ^2, for all t. □

This proposition reflects the assumption that θ_t is a random walk, hence for any choice of the public communication plan by the agency, in any period the agency cannot perfectly know θ_t. If the agency in some period t has very good knowledge about θ_t, then even if it does not send a public signal, it will know less about θ_{t+1}, just by the random walk assumption. Therefore variance of beliefs will always be bounded away from 0.

Proposition 2. *The problem admits a steady state region $[\alpha', \alpha'']$, $\alpha' > \sigma^2$, that is, if for some t, $\alpha_t \in [\alpha', \alpha'']$, then $\alpha_{t+1} \in [\alpha', \alpha'']$. Moreover, there exists T such that, for all $t > T$, $\alpha^*_{t+1} \in [\alpha', \alpha'']$*

Proof. Let $l(\alpha_t)$ and $h(\alpha_t)$ define respectively the lower and higher bound of α_{t+1} from equation (17). Showing that the boundaries of the constraint set have slope less then 1, together with the proof of proposition (1) which says that $l(\alpha_t), h(\alpha_t) \to \sigma^2$ as $\alpha_t \to 0$, is sufficient to prove the existence of the steady state region.

$$\frac{\partial h(\alpha_t)}{\partial \alpha_t} = \frac{1}{\alpha_t^2} \left(1/\alpha_t + \left(\frac{p}{p + 1/\alpha_t} \right)^2 \psi_\nu \right)^{-2} \left[1 - 2 \frac{p^2}{(p + 1/\alpha_t)^3} \psi_\nu \right],$$

$$\frac{\partial l(\alpha_t)}{\partial \alpha_t} = \frac{1}{\alpha_t^2} \left(1/\alpha_t + \psi_\nu \right)^{-2}$$

It's easy to show that $\frac{\partial h(\alpha_t)}{\partial \alpha_t} < 1$ for all parameter values, since:

$$\frac{\partial h(\alpha_t)}{\partial \alpha_t} < 1 \iff \left[1 - 2\frac{p^2}{(p+1/\alpha_t)^3}\psi_\nu\right] < \left(1 + \left(\frac{p}{p+1/\alpha_t}\right)^2 \psi_\nu \alpha_t\right)^2$$
$$\iff -2\frac{p^2}{(p+1/\alpha_t)^3}\psi_\nu < \left(\frac{p}{p+1/\alpha_t}\right)^4 \psi_\nu^2 \alpha_t^2 + 2\left(\frac{p}{p+1/\alpha_t}\right)^2 \psi_\nu \alpha_t$$

Also $\frac{\partial l(\alpha_t)}{\partial \alpha_t} < 1$ since

$$\frac{\partial l(\alpha_t)}{\partial \alpha_t} < 1 \iff \frac{1}{\alpha_t^2} < (1/\alpha_t + \psi_\nu)^2 \iff 0 < 2\frac{\psi_\nu}{\alpha_t} + \psi_\nu^2$$

Since the slope of $l(\alpha_t)$ is less than 1 and $l(0) = \sigma^2 > 0$, we can define $\alpha' : l(\alpha') = \alpha'$. Note also that $\frac{\partial l(\alpha_t)}{\partial \alpha_t} > 0$ for all $\alpha_t, \psi_\nu < \infty$ hence $\alpha' > \sigma^2$.

For the higher bound of the steady state region, we need to consider that $h(\alpha_t)$ achieves its maximum in α^* when $1/2 = \frac{p^2}{(p+1/\alpha^*)^3}\psi_\nu$. Therefore we will define $\alpha'' = h(\alpha^*)$ if $h(\alpha^*) > \alpha^*$, otherwise $\alpha'' : h(\alpha'') = \alpha''$. Given that $\frac{\partial l(\alpha_t)}{\partial \alpha_t}, \frac{\partial h(\alpha_t)}{\partial \alpha_t} < 1$, if $\alpha_t \in [\alpha', \alpha'']$, then $\alpha_{t+1} \in [\alpha', \alpha'']$.

The fact that the boundaries of the constraint set have slope less than 1 also implies that for whatever starting point, the optimal solution will eventually converge to the region $[\alpha', \alpha'']$. □

The proposition says that eventually the beliefs of the agency will converge to a well defined region and never move from there. The intuition for this is straightforward. Even if the agency starts with a very large prior variance for θ_t, then either by communicating or not communicating, it will receive a relatively precise signal about θ_t, since agents will not rely a lot on the public signal if they receive it. Therefore the agency learns and prior variance for θ_{t+1} can be smaller than prior variance for θ_t. On the other hand, by proposition (1) we know that perfect knowledge will never be achieved, and if the agency starts with very good knowledge about θ_t, for any communication decision, it will know less about θ_{t+1}. Therefore in any optimal policy it will never be the case that either the agency knows less and less, nor that it has perfect knowledge.

It must be noted that the previous proposition does not say anything about the behaviour of the optimal solution inside the steady state region nor out of it. It says that the agency will converge to a well specified interval of belifs, either by learning if it start with low knowledge, or by 'forgetting', if the knowledge about θ_t is not that useful for knowing

θ_{t+1}. But convergence to this region happens for whatever public communication decision.

We are now going to rewrite the sequential problem (16) as a functional equation problem. Let $y \equiv \alpha_{t+1}$, $x \equiv \alpha_t$ and $F(x,y) \equiv \left\{ - \left[\frac{1}{p^2 \psi_\nu} \left(\frac{1}{y-\sigma^2} - \frac{1}{x} \right) \right]^{1/2} \right\}$. Then the functional problem is defined as:

$$v(x) = \max_{y \in \Gamma(x)} F(x,y) + \beta v(y) \tag{18}$$

where $\Gamma(x)$ is the correspondence defined by the constraint (17). The following proposition establishes equivalence between the solutions of the two problems, and justifies the use of numerical simulations to find the solutions.

Proposition 3. *Let $x \in X \subseteq \mathbb{R}$ and the operator T on the space of continuous and bounded functions $C(X)$ be defined as $(Tf)(x) = \max_{y \in \Gamma(x)} F(x,y) + \beta f(y)$. Then T maps $C(X)$ into itself, T is a contraction and has a unique fixed point $v \in C(X)$; and for all $v_0 \in C(X)$*

$$||T^n v_0 - v|| \leq \beta^n ||v_0 - v||, \quad n = 0, 1, 2, \ldots \tag{19}$$

Moreover the solutions in terms of value and optimal plans of the functional equation problem (18) coincide witht the solutions of the sequential problem (16).

Proof. The proof follows standard theorems from Stokey and Lucas (1989). The only requirement is to show that $F(x,y)$ is bounded. This is done noticing that $x, y > 0$. In fact $x = \mathbb{V}ar[\theta_t | S^{t-1}] = \mathbb{V}ar[\theta_{t-1} | S^{t-1}] + \sigma^2$. Even if there is perfect knowledge of θ_{t-1} such that $\mathbb{V}ar[\theta_{t-1} | S^{t-1}] = 0$, as long as $\sigma^2 > 0$ then $x > 0$. The same will be true for y, hence $F(x,y)$ is bounded.

It is also clear that $\Gamma(x)$ is not-empty, compact valued, and continuous. Then we can apply Theorem 4.6, page 79 of Stokey and Lucas (1989) to prove the first part of the statement. The equivalence of solutions between the sequential problem and the functional problem is given by Theorems 4.2 - 4.5. □

Unfortunately, there are not standard arguments that can be made to prove properties of the value function $v(x)$ or of the optimal policy function $g(x)$, defined as $g(x) = \{y \in \Gamma(x) : v(x) = F(x,y) + \beta v(y)\}$. The reason is that the constraint correspondence $\Gamma(x)$ is highly irregular, not satisfying properties of monotonocity or convexity for all parameters values. For instance the higher bound of the constraint set can be monotonic or not monotonic in its argument x, depending on parameter values. We will therefore rely on the numerical approach to find the optimal solution.

4.2 The jamming effect of the public signal

Before relying on the numerical simulations, we will analyze in this section a particular case. We want to analyze the evolution of the precision of public knowledge if the agency is communicating in every period with no garbling. This will shed insights on the peculiar effects of public signals on knowledge of the agents.

When the agency is always precisely communicating, there is no distinction between public knowledge of the agents and knowledge of the agency. The agency becomes irrelevant in the story and the model becomes one in which agents are receiving a public signal of the aggregate of previous period actions. However public information in this case is somewhat special, because it is not lost after one period: agents transmit public knowledge to the following generation. Note that this special case allows us to focus precisely on the effects of public information when the source of this information is the observation of an aggregate of the agents' actions in the economy.

Proposition 4. *Suppose $\Phi_t \to \infty$, for all t. Then the evolution of precision of public knowledge is given by $P_{t+1} = \left\{ \sigma^2 + \left[P_t + \left(\frac{p}{p+P_t} \right)^2 \psi_\nu \right]^{-1} \right\}^{-1}$. If $\frac{\psi_\nu}{p} > \frac{1}{2}$, then there exists P' such that P_{t+1} is decreasing in P_t for $P_t \leq P'$ (weakly when $P_t = P'$), and increasing for $P_t > P'$. Otherwise P_{t+1} is monotonically increasing in P_t.*

Proof. By equation (11), if $\Phi_t \to \infty$, then $1/\alpha_t = P_t$. By using the higher bound of the constraint set defined in equation (17), the equation describing the evolution of P_{t+1} is found.

It can be easily checked that the sign of $\frac{\partial P_{t+1}}{\partial P_t}$ depends on the sign of $\frac{1}{2} - \frac{p^2}{(p+P_t)^3}\psi_\nu$. Let $f(P_t) \equiv \frac{p^2}{(p+P_t)^3}\psi_\nu$, then $lim_{P_t \to \infty} f(P_t) \to 0$ and $f(0) = \frac{\psi_\nu}{p}$. Since $f'(P_t) < 0$, if $\frac{\psi_\nu}{p} > \frac{1}{2}$, then there exists P' such that $f(P_t) > \frac{1}{2}$ for $P_t < P'$, otherwise $f(P_t) < \frac{1}{2}$ for all P_t.

Therefore, if $\frac{\psi_\nu}{p} > \frac{1}{2}$, $\frac{\partial P_{t+1}}{\partial P_t} < 0$ for $P_t < P'$ (with equality when $P_t = P'$), and $\frac{\partial P_{t+1}}{\partial P_t} > 0$ for $P_t > P'$. Otherwise $\frac{\partial P_{t+1}}{\partial P_t} > 0$. □

The proposition highlights two effects of public information. In one case, when $\frac{\psi_\nu}{p} \leq \frac{1}{2}$, the public signal always increases knowledge about θ_t. An increase in the level of precision of public knowledge today leads to an increase in public knowledge tomorrow, in this sense public knowledge is self-sustaining. However this is not true anymore when $\frac{\psi_\nu}{p} > \frac{1}{2}$. In this case, for small values of P_t, an increase in P_t actually makes P_{t+1} decrease. What is happening is that for sufficiently low values of precision of public knowledge, an increase in P_t makes agents rely more on the public signal than on private knowledge, therefore learning through the public channel is impaired. Future public signals cannot communicate more information and public knowledge decreases.

The effect is clearer if we consider the special case of a constant θ, that is $\sigma^2 = 0$. Then the evolution of public knowledge becomes:

$$P_{t+1} = P_t + \left(\frac{p}{p + P_t}\right)^2 \psi_\nu \tag{20}$$

From previous equation we can see that an increase in P_t has two effects. The first one is a positive increase on P_{t+1} given by P_t. This is the *memory effect*: if agents know more today they will know more tomorrow. The second one is the negative effect coming from the aggregate signal of past actions. This is the *jamming effect*: an increase in P_t implies that agents rely more on their private knowledge and therefore the precision of the aggregate signal, $\left(\frac{p}{p+P_t}\right)^2 \psi_\nu$, decreases. The proposition says that for low values of P_t, the jamming effect dominates the memory effect: more public information today can actually decrease future public information.

This effect happens only when $2\psi_\nu > p$, that is, when the measurement error of the aggregate signal is sufficiently more precise than the private information of the agents. If, on the other hand, private knowledge was sufficiently more precise, then an increase in P_t would always have a negative effect on the precision of the public signal. Only now it would not be sufficient to overcome the memory effect on public information, and therefore more public information today increases future public information.

This explanation has been done for the stark case of the agency always fully communicating, and therefore it is as if agents had memory. However the memory effect is present as long as the agency is communicating something, because the agency is the storage for public knowledge, and so, as long as it communicates, it communicates also past public knowledge.

Note also that a decrease in public knowledge per se does not imply a decrease in the welfare criterion, given that we have to evaluate an infinite sum of utilities. However we will show that this effect is welfare relevant through the numerical solutions.

4.3 Numerical Solutions

We are going to analyze the optimal policy solution as parameter values $\{\beta, \sigma^2, \psi_\nu, p\}$ change. For a particular application in mind we would like to calibrate parameters,[4] how-

[4] Fogli and Veldkamp (2007) have a learning model in which they calibrate parameters similar to ours in the context of women learning about the effect of their working choice on the achievements of their children.

ever here we are interested in finding some properties of the solution. In the following figures we are going to represent the policy function that solves the functional problem (18).

In all figures, on the x-axis there is the current value of the state variable, that is α_t, the variance of the prior beliefs of the agency about θ_t, and on the y-axis the future (one period ahead) value of the state, α_{t+1}. The dashed lines represent the higher and lower bound of the constraint set defined by equation (17). Therefore the higher bound defines the value of the state tomorrow when the agency is precisely communicating in period t, that is $\Phi_t \to \infty$, and the lower bound defines the value of the state tomorrow when the agency is not communicating in period t, that is $\Phi_t = 0$. The dashed-point line indicates the 45 line. The thick red line between the higher and lower bound indicates the policy function.

4.3.1 Optimal absence of public communication

In figure (1), $\{\beta, \psi_\nu, p\}$ are fixed respectively to $\{0.95, 1, 1\}$ and σ^2, the term representing the innovations of the random walk process, takes 3 different values in the three different panels. The first panel in figure (1) is drawn for $\sigma^2 = 1$. As we can see the policy function is always on the higher bound, hence the optimal policy is always to communicate with full precision. On the dynamics of the evolution of the state, depending on where we start, we can learn or 'forget', but the convergence to the steady state is monotonic.

In the second panel of figure (1), $\sigma^2 = 0.01$, that is, there are 'smaller' innovation in the parameter θ_t, or, put differently, the term θ_t is relatively more constant over time. Now the policy function is completely different, and it lies for some values of the state α_t over the lower bound of the constraint set. When this happens it is actually optimal to not communicate, that is, the agency does not send a public message, $\Phi_t = 0$.

This is the starkest solution the agency can put into effect in order to prevent the underproduction of information. By not sending the public message in that period the agency is able to observe with more accurate precision the private information of the agents. When it has acquired sufficient information, α_t has become smaller, then it will start communicating with full precision. Doing so it allows to achieve higher welfare than communicating for every value of the state. Eventually, however, it will be optimal to communicate, and the variance of prior beliefs converges to a long run steady state level. By not communicating, the agency is effectively preventing the jamming effect to happen: no present public communication will entail more precise future public signals. Note that by not communicating in some period, the agency is actually damaging the agents alive in that period. However the welfare benefits to later generations are higher than the welfare costs of not communicating for one period, and therefore aggregate welfare is higher.

In the third panel of figure (1), $\sigma^2 = 0.001$, that is, there are even smaller innovations. The policy function changes in a well defined way: there is a larger range of the current value of the state that prescribes for not communication. The reason for this change is the policy function when σ^2 changes is clear. The smaller the σ^2, the more constant is θ_t over time. This in particular implies that the information collected in any period becomes more valuable for the future. If $\sigma^2 \to \infty$, then the parameter θ_t would be completely different every period, hence the information collected by the agency would have no implication for future periods θ_t. In such case, always communicating is the optimal policy. However, if $\sigma^2 = 0$, then θ_t is constant over time, therefore being able to collect more precise information in any one period has greater consequences over future public communication. This is why, when σ^2 decreases, it can be optimal for the agency to be silent in order to efficiently learn, and the more so, the smaller is σ^2.

In figure (2), $\{\sigma^2, \psi_\nu, p\}$ are fixed respectively to $\{0.03, 1, 1\}$, while β, the value at which the agency discounts future generations, changes. In the first panel $\beta = 0.9$. Here, as in the first panel of figure (1), the optimal policy is always to communicate. In the second and third panel, β respectively takes the value of 0.95 and 0.99. In these panels we see that the optimal policy also prescribes to be silent for some value of the state. However the reason is different with respect to the previous case when we were letting σ^2 change. In this case the agency is valuing more the welfare of future generations, hence it will try to help them more by shutting down current communication, and acquire more information for the future. Note in fact that a higher β increases the marginal benefits as defined in equation (15). Therefore, the higher the weight on future generations, the more likely (in terms of larger range of the current value of the state) it is that the agency will choose to be silent.

4.3.2 Optimal garbling of public communication

In figure (3), $\{\beta, \sigma^2, \psi_\nu\}$ are fixed respectively to $\{0.95, 0.1, 1\}$, while, the precision of the public signal of the agents, changes. In the first panel $p = .01$. The policy function says that, for some values of the state, it is optimal to tack finite noise to the public message, that is $\Phi_t^* \in (0, \infty)$. The agency will choose to communicate, but in order not to let the agents rely too much on the public signal and therefore create under-production of information, it will decide to add noise to its message. The agency is *garbling* its knowledge to the individuals. This is a striking feature of the optimal communication plan: better than just not communicating, as in previous instances, the agency provides the public signal in order to partially improve the welfare of the current generation of agents, but garbles the message in order to efficiently communicate in the future.

As we can see from the other two panels, the policy function has non monotonic behaviour in p. As p increases the range of values for α_t that imply a garbling of the public message first increases, but then it is completely absent, leading again to an optimal policy of precise communication. The intuition for this is that, when p is sufficiently precise, agents already have large private information, hence public communication cannot affect welfare too much. In this case there are no welfare advantage of garbling or shutting down the public signal.

In figure (4), $\{\beta, \sigma^2, p\}$ are fixed respectively to $\{0.95, 0.1, 0.1\}$, while ψ_ν, the precision of the measurement error of the signal to the agency, changes. As ψ_ν changes from 0.01, in the first panel, to 100, in the third panel, we see that the optimal policy function shows non monotonic behaviour. In the first and last panel the optimal solution is always to fully communicate, while in the middle panel, there is a range of values for the state for which it is optimal to garble. The intuition for this result is the following: when ψ_ν is sufficiently small (large), the precision of the signal to the agency is small (large) for any communication policy. Therefore there is not too much knowledge that can be acquired by shutting down communication. On the other hand, it is for intermediate values of ψ_ν that the trade off between present and future public communication becomes relevant in terms of welfare.

4.4 A simple two periods analysis

The infinite horizon intertemporal model is too complex to be analyzed analytically, hence the use of numerical simulations. However a simple two periods model can be used in order to replicate analytically the insights of the numerical simulations.

We are going to consider the same structure of the economy, but now $t = 1, 2$. Assume also that $\sigma^2 = 0$, so that θ_t is constant across the two periods. The agency sends a public signal in period t before observing the actions for that period. We are going to assume that in $t = 2$ the agency fully reveals her mean beliefs. The agency can decide however on the precision of the public signal it sends in $t = 1$, defined by Φ. At the beginning of period $t = 1$ the agency has prior beliefs over the variance of the θ given by α.

Aplying the model specified in (16) to this two periods case, we want the agency to solve the following problem:

$$\min_{\Phi} \frac{1}{p+P_1} + \beta \frac{1}{p+P_2} \tag{21}$$

$$\text{s.t.} \quad P_1 = \frac{1}{\alpha + 1/\Phi}$$

$$P_2 = P_1 + \left(\frac{p}{p+P_1}\right)^2 \psi_\nu$$

Let the problem be represented by $\min_\Phi WL(\alpha, \beta, \psi_\nu, p, \Phi)$ and let $\Phi^* = \arg\min\ WL(\alpha, \beta, \psi_\nu, p, \Phi)$. The following two propositions shed lights on the simulations' results.

Proposition 5. *Necessary condition for $\Phi^* < \infty$ is $\beta > 3$.*

Proof. Let $y \equiv \frac{p^2}{(p+P_1)^3}\psi_\nu$. Then since Φ can take any value between 0 and infinity, $y \in [\frac{p^2}{(p+1/\alpha)^3}\psi_\nu, \frac{\psi_\nu}{p}]$. Notice that $y \to \frac{p^2}{(p+1/\alpha)^3}\psi_\nu$ as $\Phi \to \infty$, while $y \to \frac{\psi_\nu}{p}$ as $\Phi \to 0$. The problem can therefore be rewritten as:

$$\min_y \left(\frac{y}{p^2\psi_\nu}\right)^{1/3} \left(1 + \frac{\beta}{1+y}\right)$$

Let $f(\cdot)$ define the objective. Equalizing to zero the first order condition $\frac{\partial f(\cdot)}{\partial y}$ gives a quadratic equation in y: $y^2 + 2(1-\beta)y + (\beta+1) = 0$, whose roots are $y_{1,2} = (\beta - 1) \pm \sqrt{\beta(\beta-3)}$. Therefore necessary condition for having real roots is $\beta > 3$. It can be checked that in this case both y_1 and y_2 are positive.

If roots are not real, from the first order condition, it can be checked that when $\beta \leq 3$, $\frac{\partial f(\cdot)}{\partial y} > 0$ (with equality when $y = 2$), therefore the unique global minimizer is $y = \frac{p^2}{(p+1/\alpha)^3}\psi_\nu$, that is, $\Phi^* \to \infty$.

On the other hand when $\beta > 3$, $\frac{\partial f(\cdot)}{\partial y} > 0$ for $y < y_1$, $\frac{\partial f(\cdot)}{\partial y} < 0$ for $y_1 < y < y_2$, and $\frac{\partial f(\cdot)}{\partial y} > 0$ for $y > y_2$. Therefore the function exhibits a local minimum in y_2. If $\frac{p^2}{(p+1/\alpha)^3}\psi_\nu < y_2 < \frac{\psi_\nu}{p}$, then, depending on the value of the $f(\cdot)$ at the extrema of the range of y, y_2 can be also the global minimizer. Since y_2 belongs to the interior of the range of y, then y_2 being the minimizer implies $\Phi^* < \infty$. $\square$

This proposition highlights a simple fact. It is necessary to have more than two periods in order for the trade-off between present and future communication to be welfare relevant. In the simple two periods setting with normal values for the discount factor β the trade-off is not welfare relevant: even by shutting down communication in the first period, the welfare benefits given by the informative gains in the second period are not enough to justify the welfare loss of not communicating in the first period. By raising the value of β higher than 1, the objective behaves as if we are adding more periods to the problem, hence the trade-off between present and future communication becomes welfare relevant.

Proposition 6. *Suppose $\beta > 3$, then:*

i. there exist p', ψ'_ν, α' and $\tilde{\alpha}$, $\alpha' < \tilde{\alpha}$, such that $\Phi^* \to \infty$ for $\alpha < \alpha'$, $0 < \Phi^* < \infty$ for $\alpha' < \alpha < \tilde{\alpha}$, and $\Phi^* \to \infty$ for $\alpha > \tilde{\alpha}$;

ii. there exist p'', ψ''_ν and α'' such that $\Phi^* \to \infty$ for $\alpha < \alpha''$ and $\Phi^* = 0$ for $\alpha > \alpha''$;

iii. there exist p''', ψ'''_ν such that, for all α, $\Phi^* \to \infty$;

where $\frac{\psi'''_\nu}{p'''} < \frac{\psi''_\nu}{p''} < \frac{\psi'_\nu}{p'}$.

Proof. From proposition (5) we know that if $\beta > 3$, there exists $y_{1,2} = (\beta - 1) \pm \sqrt{\beta(\beta - 3)}$ that solve the equation given by first order condition of the problem equalized to 0. Moreover y_1 is a local maximum and y_2 is a local minimum.

i. Fix $\beta > 3$ and let p', ψ'_ν such that $\frac{\psi'_\nu}{p'} > y_2(\beta)$. Since $\frac{p^2}{(p+1/\alpha)^3}\psi_\nu$ is monotonically increasing in α, it tends to 0 as $\alpha \to 0$ and it tends to $\frac{\psi_\nu}{p}$ as $\alpha \to \infty$ then for α sufficiently small the range of y is such that $\frac{p'^2}{(p'+1/\alpha)^3}\psi'_\nu < y_1(\beta) < y_2(\beta) < \frac{\psi'_\nu}{p'}$.
   $y_2(\beta)$ is a local minimum, in order to check whether it is a global minimum we need to compare $f(y_2(\beta))$ with $f(\frac{p'^2}{(p'+1/\alpha)^3}\psi'_\nu)$. Note that $f(\frac{p'^2}{(p'+1/\alpha)^3}\psi'_\nu) \to 0$ as $\alpha \to 0$, and $f(\frac{p'^2}{(p'+1/\alpha)^3}\psi'_\nu) \to f(\frac{\psi'_\nu}{p'})$ as $\alpha \to \infty$. On the other hand $f(y_2(\beta))$ is always strictly positive and independent of α, and by the choice of p', ψ'_ν, $f(y_2(\beta)) < f(\frac{\psi'_\nu}{p'})$.
   Therefore, using the sign of $\frac{\partial f(\cdot)}{\partial y}$, there exists α' such that $f(y_2(\beta)) = f(\frac{p'^2}{(p'+1/\alpha')^3}\psi'_\nu)$ and $\frac{p'^2}{(p'+1/\alpha')^3}\psi'_\nu \neq y_2(\beta)$ such that:
   - for $\alpha < \alpha'$, $f(\frac{p'^2}{(p'+1/\alpha)^3}\psi'_\nu) < f(y_2(\beta))$, therefore the global minimizer is $y = \frac{p'^2}{(p'+1/\alpha)^3}\psi'_\nu$;
   - for $\alpha = \alpha'$, both $y_2(\beta)$ and $\frac{p'^2}{(p'+1/\alpha)^3}\psi'_\nu$ are global minimizers;
   - for $\alpha > \alpha'$ then there exists $\tilde{\alpha} > \alpha'$ such that $\frac{p'^2}{(p'+1/\tilde{\alpha})^3}\psi'_\nu = y_2(\beta)$ such that
     - for $\alpha' < \alpha < \tilde{\alpha}$, $f(\frac{p'^2}{(p'+1/\alpha)^3}\psi'_\nu) > f(y_2(\beta))$, therefore the global minimizer is $y_2(\beta)$;
     - for $\alpha > \tilde{\alpha}$ the global minimizer is $\frac{p'^2}{(p'+1/\alpha)^3}\psi'_\nu$.

ii. Fix $\beta > 3$ and let p'', ψ''_ν such that $y_1(\beta) < \frac{\psi''_\nu}{p''} < y_2(\beta)$. Using the same analysis as before, for α sufficiently small $\frac{p''^2}{(p''+1/\alpha)^3}\psi''_\nu < y_1(\beta) < \frac{\psi''_\nu}{p''}$. $y_1(\beta)$ is a global maximum. The minimizer will depend on the value of the function $f(\cdot)$ at the extrema of the range of y, that is we need to compare $f(\frac{p''^2}{(p''+1/\alpha)^3}\psi''_\nu)$ and $f(\frac{\psi''_\nu}{p''})$. Given the signs of $\frac{\partial f(\cdot)}{\partial y}$, there exists α'' such that $f(\frac{p''^2}{(p''+1/\alpha'')^3}\psi''_\nu) = f(\frac{\psi''_\nu}{p''})$ and $\frac{p''^2}{(p''+1/\alpha'')^3}\psi''_\nu \neq \frac{\psi''_\nu}{p''}$ such that:

- for $\alpha < \alpha''$, $f(\frac{p''^2}{(p''+1/\alpha)^3}\psi''_\nu) < f(\frac{\psi''_\nu}{p''})$, therefore the global minimizer is $y = \frac{p''^2}{(p''+1/\alpha)^3}\psi''_\nu$;
- for $\alpha > \alpha''$, $f(\frac{p''^2}{(p''+1/\alpha)^3}\psi''_\nu) > f(\frac{\psi''_\nu}{p''})$, therefore the global minimizer is $y = \frac{\psi''_\nu}{p''}$;
- for $\alpha = \alpha''$, both $y = \frac{p''^2}{(p''+1/\alpha)^3}\psi''_\nu$ and $y = \frac{\psi''_\nu}{p''}$ are global minimizers.

iii. Fix $\beta > 3$, and let p''', ψ'''_ν such that $\frac{\psi'''_\nu}{p'''} < y_1(\beta)$. Then $\frac{\partial f(\cdot)}{\partial y} > 0$ and the global minimizer is $y = \frac{p'''^2}{(p'''+1/\alpha)^3}\psi'''_\nu$.

By reminding that $\Phi^* \to \infty$ as $y \to \frac{p^2}{(p+1/\alpha)^3}\psi_\nu$, and that $\Phi^* = 0$ if $y = \frac{\psi_\nu}{p}$, the statement is obtained. □

This proposition replicates all the results obtained through the numerical solutions. Notice that subcase *i.* replicates the case when there is finite amount of noise in the public signal, for instance the policy function in the first panel of figure (3). For low values of the state, the optimal policy is to precisely communicate, while for high values of the state, the optimal policy is to garble. As in the numerical solution, there is a discontinuity in the two period model solution when $\alpha = \alpha'$.

The discontinuity can be rationalized in this way. Suppose α is small, that is, precision of knowledge of the agency is large. We are in the case when the memory effect wins over the jamming effect. Public communication greatly improves welfare in the first period, and even though it leads to under-production of information in the second period, it still improves public information in the second period. As α increases, welfare in the first period decreases, since there is less information communicated by the public signal. But also welfare in the second period decreases, since agents are carrying less information from the first period.

When α increases more, the precision of public knowledge in first period, P_1 still decreases. But then the jamming effect offsets the memory effect, and now P_2 and P_1 move in opposite directions. A further increase in α leads to a decrease in P_1 which leads to an increase in P_2. At this point the agency has the option to noise up the public signal. Given that welfare in the first period is decreasing and convex in the variance of the public signal of the first period, adding one additional unit of noise will entail a small marginal cost to the welfare in first period. However marginal benefit in the second period of noising up the public signal are high for two reasons: we are at low levels of P_2, and the jamming effect greatly increases P_2. Therefore the agency will optimally increase the amount of noise in the public signal, raising P_2 until marginal benefits are equated to the marginal costs. The presence of the jamming effect entails the discrete jump between precisely communicating and adding finite amounts of noise. Depending on its strength, the optimal decision might be to add finite amount of noise or completely shut down public communication.

4.4.1 Garbling and the jamming effect

From a descriptive point of view, it is worthwhile understanding when garbling is preferred to shutting down public communication. From proposition (6) we know that garbling will be optimal with respect to no communication when the ratio $\frac{\psi_\nu}{p}$ is sufficiently large. By looking at the evolution of public knowledge, $\frac{\partial P_{t+1}}{\partial P_t} = 1 - 2\frac{1}{(1+P_t/p)^3}\frac{\psi_\nu}{p}$. A higher value of $\frac{\psi_\nu}{p}$ implies a higher jamming effect: a smaller decrease of the precision of public knowledge in the first period can lead to a larger increase of public knowledge in the second period. Therefore sending in the first period the public signal with some noise improves welfare for the first period, while the large jamming effect allows to obtain anyway large informational gains, and hence welfare improvements, for the second period. When there are large jamming effect there is no need to completely shut down communication in order to obtain future informational gains.

5 Conclusion

This work builds a model of informational flows between imperfectly informed agents and a benevolent public authority. It shows that public communications coming from the authority can have negative effects on future public communication, preventing efficient production of information. In order to prevent the negative effect optimal public communication strategies might involve silence or garbling, that is adding noise to the public signal.

The analysis could give a rationalization to FOMC statements. Several commentators have seen FOMC statements as expressly ambigous, blurring sometimes the message and the knowledge at their disposal. This behaviour might indeed be the optimal choice of the FOMC in order not to reveal too much of their knowledge and not to make agents take actions that are based only on common public information.

The general implication of this work is that trasparency can backfire. In particular, it is fundamental to ascertain what is the source of the public information before deciding on the communication plan. If information comes from observing aggregate actions, then public communication in any period might actually prevent effective learning. Therefore, transparency, or complete and constant public communication, might not be the optimal policy to follow.

The literature on the social value of public communication has always been abstract: the future of this stream of literature has to be in a more applied approach, where effects of public communication can be estimated in order to provide a quantification of its welfare relevance.

References

Amador, Manuel and Pierre-Olivier Weill, "Learning from Private and Public observations of others' actions," 2009. mimeo, Stanford University.

—— **and** —— , "Learning from Prices: Public Communication and Welfare," 2010. mimeo, Stanford University.

Angeletos, George-Marios and Alessandro Pavan, "Efficient Use of Information and Social Value of Information," *Econometrica*, 2007, *75* (4), pp. 1103–1142.

Chamley, C., *Rational herds: Economic models of social learning*, Cambridge Univ Pr, 2004.

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

Miccoli, Marcello, "Optimal Dynamic Public Communication," 2010. mimeo, Stanford University.

Morris, Stephen and Hyun Song Shin, "Social Value of Public Information," *The American Economic Review*, 2002, *92* (5), pp. 1521–1534.

—— **and** —— , "Central Bank Transparency and the Signal Value of Prices," *Brookings Papers on Economic Activity*, 2005, *2005* (2), pp. 1–43.

—— **,** —— **, and Hui Tong**, "Social Value of Public Information: Morris and Shin (2002) Is Actually Pro-Transparency, Not Con: Reply," *The American Economic Review*, 2006, *96* (1), pp. 453–455.

Stokey, N. and R. Lucas, *with E. Prescott: Recursive Methods in Economic Dynamics*, Harvard University Press, 1989.

Svensson, Lars E. O., "Social Value of Public Information: Comment: Morris and Shin (2002) Is Actually Pro-Transparency, Not Con," *The American Economic Review*, 2006, *96* (1), pp. 448–452.

Vives, X., *Information and learning in markets: the impact of market microstructure*, Princeton Univ Pr, 2010.

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

——, "Learning from Others: A Welfare Analysis,," *Games and Economic Behavior*, 1997, *20* (2), 177 – 200.

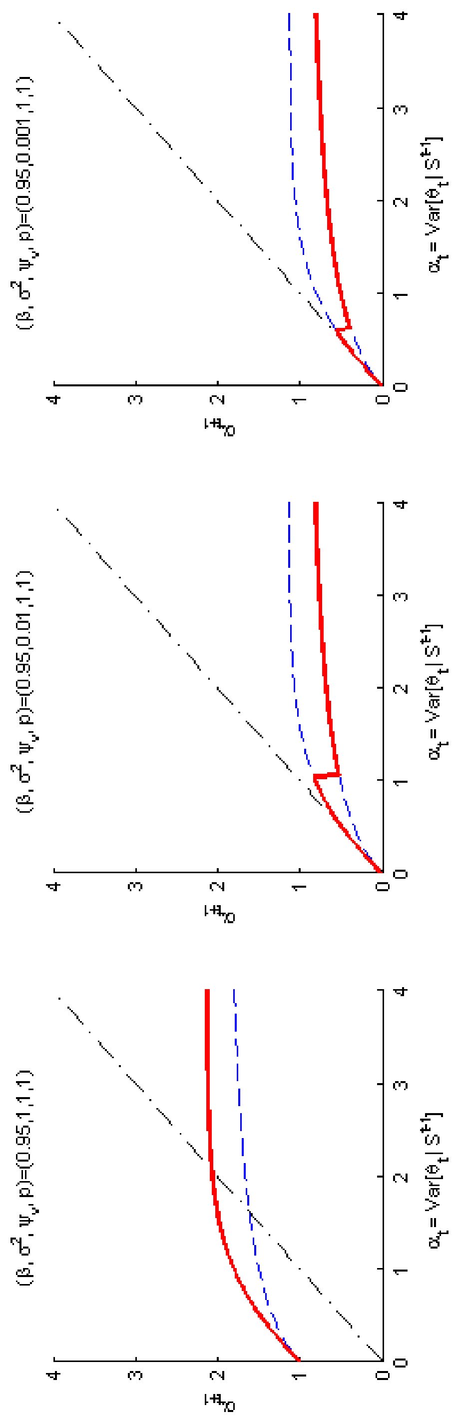

Figure 1: Polic function as function of σ^2.

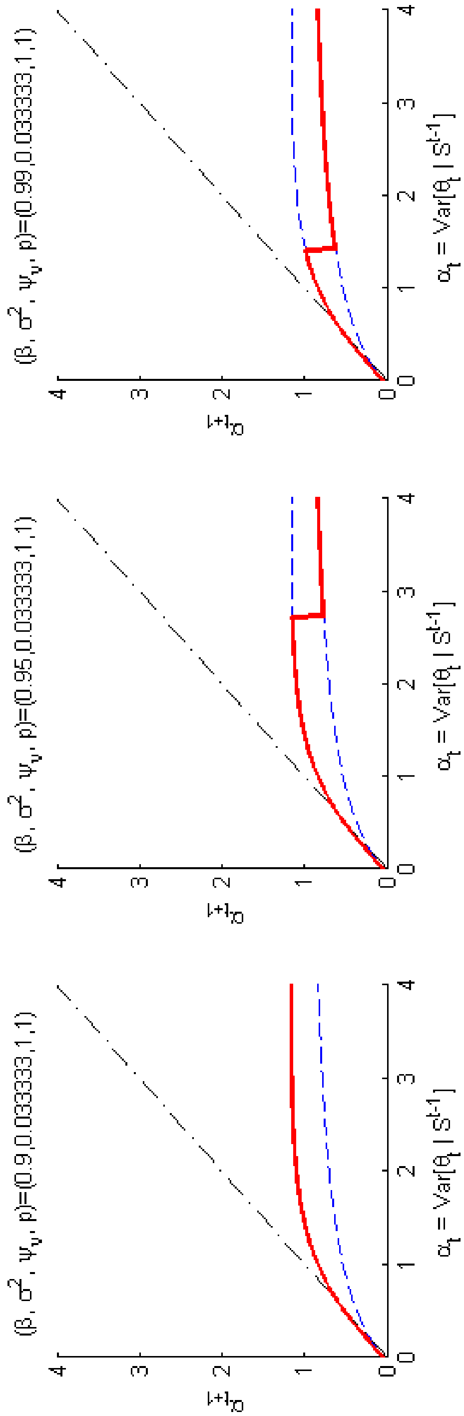

Figure 2: Policy function as function of β.

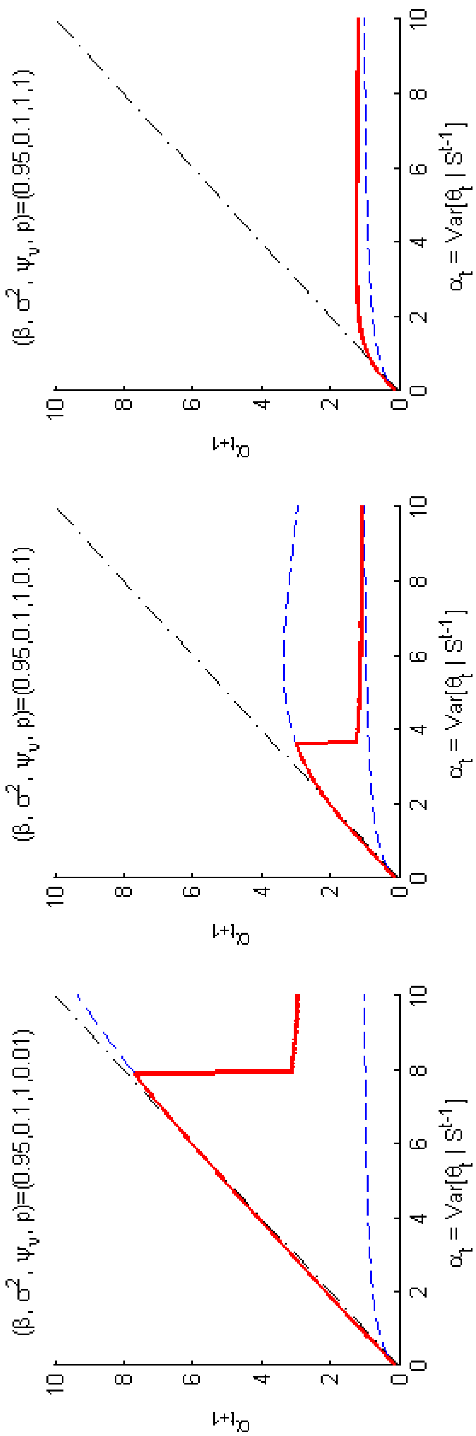

Figure 3: Policy function as function of p.

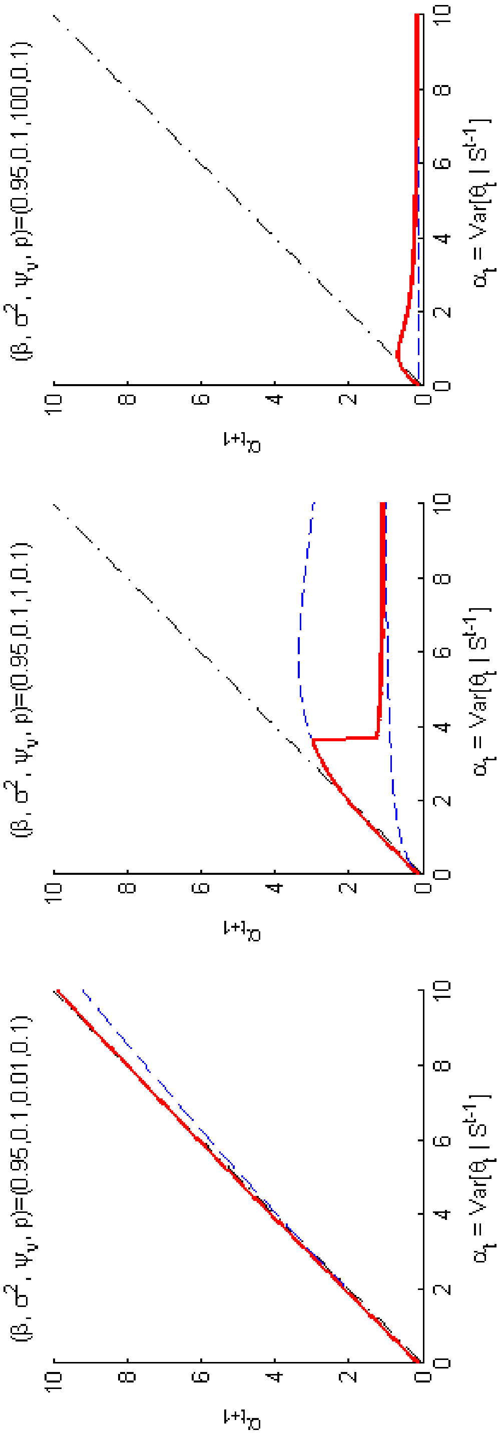

Figure 4: Policy function as function of ψ_ν.