Information Choice Technologies

Christian Hellwig, Sebastian Kohls and Laura Veldkamp*

January 17, 2012

Theories based on information costs or frictions have become increasingly popular in macroeconomics and macro-finance. The literature has used various types of information choices, such as rational inattention, inattentiveness, information markets or costly precision.[1] Using a unified framework, we compare these different information choice technologies and explain why some generate increasing returns and others, particularly those where agents choose how much public information to observe, generate multiple equilibria. The results can help applied theorists to choose the appropriate information choice technology for their application and to understand the consequences of that modeling choice.

1 The game

We convey our main intuition using a beauty contest game (as in Morris and Shin, 2002). Agents seek to take actions close to the true state and close to the average action of others. The agents choose what information to observe about the true state, before they play this game. Different information choice technologies are represented as different information cost functions and different constraints on the signal choice set.

We use a quadratic objective because of its tractability and because by quadratically approximating objectives, we can map many models into

*Hellwig: University Toulouse, Toulouse School of Economics, Manufacture de Tabacs, 21 Allees de Brienne, 31000 Toulouse, France, christian.hellwig@tse-fr.eu. Kohls: Northwestern University, 2001 Sheridan Rd., Evanston, IL 60208, sebastian.kohls@u.northwestern.edu. Veldkamp: New York University, Stern School of Business, 44 West Fourth St, rm 7-77, NY, NY 10012, lveldkam@stern.nyu.edu. Thanks to David Myatt for his helpful discussion. We are grateful for financial support from the European Research Council (starting grant agreement 263790).

[1]See e.g., Sims (2003), Reis (2006), Veldkamp (2006), Mackowiak and Wiederholt (2009), Matejka (2011), Myatt and Wallace (2011), and Veldkamp (2011) for many other examples.

this framework. In stage one, nature draws the state variable s from the distribution $N(\mu, \tau_s^{-1})$ and a series of signals about s. Agents choose signals to observe. In stage 2, agents observe their chosen signals and simultaneously pick an optimal action.

Specifically, a measure one continuum of agents, indexed by $i \in [0,1]$ choose an action $a_i \in \mathbb{R}$ to minimize the expected squared distance between their action and a target action, that is a weighted average of the average action $\bar{a} = \int a_i di$ and the unknown state s, minus any cost c of acquiring information, where c is denominated in units of expected utility:

$$u\left(a_i, \bar{a}, s\right) = -\left(a_i - r\bar{a} - (1-r)s\right)^2 - c. \tag{1}$$

If s is common knowledge, the best response is $a_i = (1-r)\, s + r\bar{a}$, and $a_i = \bar{a} = s$ constitutes the unique equilibrium. The coefficient $r < 1$ measures the complementarity/substitutability of agents' decisions. If $r > 0$, decisions are complementary: Best responses are increasing in the prices set by other agents. If $r < 0$, decisions are strategic substitutes. A higher r means more complementarity.

Denote the information set that includes chosen signals $\mathbf{s}$ as $\mathcal{I}_i$. The first order condition of (1) with respect to a_i yields: $a_i = E[r\bar{a} + (1-r)s|\mathcal{I}_i]$. Utility (1) is then simply a conditional variance $u\left(a_i, \bar{a}, s\right) = Var(r\bar{a} + (1 - r)s|\mathcal{I}_i) - c$. The variance of this sum can be decomposed into the individual variances and a covariance term:

$$E[u\left(a_i, \bar{a}, s\right)] = r^2 Var[\bar{a}|\mathcal{I}_i] \tag{2}$$

$$+2r(1-r)Cov[\bar{a}, s|\mathcal{I}_i] + (1-r)^2 Var[s|\mathcal{I}_i] - c.$$

Since 2 is the expected utility of an agent who acts optimally in the second-stage action game, it is the payoff function for the first-stage information choice game. Thus, to understand the value of any signal choices, it is sufficient to know what the information implies for three moments: the conditional variance of the state, the conditional variance of the average action and the covariance between the average action and the state.

A flexible signal structure. Suppose that nature selects a $k \times 1$ vector of common signal noises $\mathbf{u} \sim \mathcal{N}\left(0, I_k\right)$, independent of the state s. In addition, for each agent, nature selects an $l \times 1$ vector of idiosyncratic signal noises $\mathbf{v}^i$, which are independently and identically distributed across agents, $\mathbf{v}^i \sim \mathcal{N}\left(0, I_l\right)$, independent of s and $\mathbf{u}$. These shocks generate an $n \times 1$ vector of potentially observable signals $\mathbf{z}^i$:

$$\mathbf{z}^i = \mathbf{1}_n \cdot s + D\mathbf{u} + B\mathbf{v}^i \tag{3}$$

$\mathbf{1}_n$ is an $n \times 1$ vector of ones and D and B are diagonal $(n \times n)$ matrices of coefficients.[2]

Thus, we can express agent i's jth signal as $z_j^i = s + d_j u_j + b_j v_j^i$. This signal structure allows for arbitrary correlation in signals across agents. In particular, by setting either $\mathbf{d}$ or $\mathbf{b}$ equal to zero, we can allow for special cases where all signals are either purely private, i.e. with noise independent across agents, or common. If $d_j + b_j = \infty$, signal j is unobserved (or is uninformative).

The agent's cost of information is determined by a function $c(\mathbf{d}, \mathbf{b})$, which is decreasing in both arguments.

Bayesian updating. Each of agent i's signals is an unbiased predictor of the state s with variance $b_j^2 + d_j^2$. Bayes' Law for normal variables delivers posterior beliefs

$$E[s|\mathcal{I}_i] = \frac{\tau_s \mu + \sum_j (b_j^2 + d_j^2)^{-1} z_j}{\tau_s + \sum_j (b_j^2 + d_j^2)^{-1}} \tag{4}$$

$$Var[s|\mathcal{I}_i] = \frac{1}{\tau_s + \sum_j (b_j^2 + d_j^2)^{-1}}. \tag{5}$$

How agents update about the average action $\bar{a}$ depends on the properties of the signal they observe. We consider symmetric information choices. This implies that in equilibrium, all agents choose to observe signals with the same precision and therefore choose the same action rules, although, signal outcomes and realized actions may differ. Since the first-order condition and the Bayesian updating formula are both linear in signals and priors, $a_i = \gamma_0 \mu + \sum_j \gamma_j z_j^i$, where γ_0 denotes the weight on priors in actions, γ denotes the weight on the signal if only 1 signal is observed and γ_j denotes the weight on signal $j \geq 1$ when multiple signals are observed. Since v^i is independent across agents, $\bar{a} = \gamma_0 \mu + \sum_j \gamma_j (s + d_j u_j)$. Thus, the beliefs about average actions are summarized by

$$E[\bar{a}|\mathcal{I}_i] = \gamma_0 \mu + \sum_j \gamma_j (E[s|\mathcal{I}_i] + d_j E[u_j|\mathcal{I}_i]). \tag{6}$$

[2]For most of our results, D and B can be arbitrary $(n \times k)$ and $(n \times l)$ matrices with rank n. Such cross-signal correlation does not affect the key properties of the problem. The agent simply uses the inverse of the variance-covariance to undo the correlation and back out the underlying orthogonal shocks. But doing so make the problem less transparent. See Appendix A and Veldkamp (2011), chapter 3 for details. The mathematical appendix, containing derivations and proofs is posted on the authors' websites.

2 Comparing signal choices

Many frequently-used learning technologies can be described as special cases of (3), with some restriction on $\mathbf{d}$ and $\mathbf{b}$ choice. We discuss what this restriction implies for the three sufficient statistics and the information choice equilibria.

2.1 Full revelation (Inattentiveness)

Suppose an agent can chose one of two options: Observe no signal ($\mathbf{d}+\mathbf{b}=\infty$) or observe s exactly ($\mathbf{d}=\mathbf{b}=0$) at a cost c. This is a limiting case of either public or private information acquisition, as the precision tends to infinity. An example of this learning technology in the literature is the "inattentiveness" choice in Reis (2006), where agents choose dates at which agentes acquire full information.

For informed agents, the sufficient statistics are simple. Since they know the state and others' information sets, they can deduce average actions. Thus, $Var[s|s]=Var[\bar{a}|s]=Cov[\bar{a},s|s]=0$. Let α be the fraction of agents that choose to become informed. The first order condition tells us that uninformed agents should choose $a_i=\mu$ and informed agents should choose $a_i=(1-\gamma)\mu+\gamma s$ where $\gamma=(1-r)/(1-r\alpha)$. Then for uninformed agents, the sufficient statistics are $Var[s]=\tau_s^{-1}$, $Var[\bar{a}]=\gamma^2\tau_s^{-1}$ and $Cov[\bar{a},s]=\gamma\tau_s^{-1}$.

There are three types of possible equilibria: Either all agents, no agents, or some agents acquire full information. Which equilibrium prevails depends on the information cost c, the degree of complementarity r and prior precision τ_s.

Proposition 1 *With fixed costs of full revelation, and complementarity in actions ($r>0$), multiple equilibria exist if $c\in\left((1-r)^2\tau_s^{-1},\tau_s^{-1}\right)$.*

When there is strategic substitutability in actions ($r<0$), the game has a unique equilibrium. But when when actions are complements, information choice is also a complement. The combination of complementarity and the discrete nature of the choice (to learn or not to learn) generates multiple equilibria.

2.2 Private signals (rational inattention)

In many settings, signals about the state s are conditionally uncorrelated across agents ($D=0$). Suppose the agent observes a single signal. Then, $z^i=s+b^iv^i$ where $v^i\sim N(0,1)$ are independent across i. Each agent chooses b_i to maximize expected utility (2), subject to a cost-function $c(b_i)$ that is decreasing in b_i.

An example of this kind of learning technology is rational inattention (Sims, 2003), where all information is potentially available to an agent. But their limited information processing ability causes them to add noise to whatever they observe. Each agent creates their own noise, independent of any other agent.[3]

Setting $\mathbf{d} = 0$ in (5) and (6) reveals that the three summary statistics are

$$Var[s|\mathcal{I}_i] = 1/(\tau_s + (\mathbf{b}^i)^{-2}) \tag{7}$$

$$Var[\bar{a}|\mathcal{I}_i] = (1-\gamma_0)^2 Var[s|\mathcal{I}_i] \tag{8}$$

$$Cov[\bar{a}, s|\mathcal{I}_i] = (1-\gamma_0) Var[s|\mathcal{I}_i]. \tag{9}$$

If agent i observes more information, $Var[s|\mathcal{I}_i]$, $Var[\bar{a}|\mathcal{I}_i]$ and $Cov[s, \bar{a}|\mathcal{I}_i]$ all fall by the same proportion.

A unique information choice equilibrium. If other agents acquire more information, they put more weight on the more precise private signals when forming their actions. Thus, $(1-\gamma_0)$ rises. When actions are complements $(r > 0)$, this increases the marginal value of reducing $Var[s|\mathcal{I}_i]$ by acquiring information oneself. This is a complementarity in information acquisition. But this complementarity is not sufficiently strong to generate multiple equilibria (Hellwig and Veldkamp, 2009).

The choice of one signal's precision is unique. With two or more private signals and a cost function of the sum of the signal precisions, there will always be multiplicity. Learning from two signals with precisions τ_1 and τ_2 or with $\tau_1 + t \geq 0$ and $\tau_2 - t \geq 0$ leaves all the sufficient statistics and the information cost unchanged. Thus, an agent is indifferent between any signal precisions that have the same sum. So, multiple equilibria exist, but the distinction between these equilibria is not economically meaningful.

Rational inattention and cost concavity. When the state s and the signals are normally distributed, rational inattention dictates that the amount of information processed is $K = \frac{1}{2}\ln\left(|Var(s)|/|Var(s|\mathcal{I}_i)|\right)$. It then allows for any arbitrary cost function $c(K)$, or simply a bound on K.

Rational inattention has a form of diminishing marginal cost of precision in it. Here are examples of that property: 1) When s is a scalar, a one-unit increase in signal precision increases posterior precision $1/Var(s|\mathcal{I}_i)$ by one

[3]In principle, rational inattention allows agents to choose not only the precision of their private signal, but also the shape of this signal noise distribution. Since our objective function is quadratic, normal signals are optimal in this setting (Sims, 2003).

unit. That increase has a marginal cost proportional to $Var(s)$. This implies that learning about something unfamiliar (high $Var(s)$) is costly. 2) In a dynamic problem, if an agent learns more about s over time, $Var(s)$ falls. For a given amount of K, signal precision could grow over time. 3) When there are multiple risks, K depends on the determinant of the precision matrix $|Var(s|\mathcal{I}_i)^{-1}|$. If risks and signals are independent, this is a product of posterior precisions: $\prod_j(\tau_{sj}+b_j^{-2})$. Increasing the precision of signals that are already precise (high b_j^{-2}) increases the product by less (is less costly).

The fact that knowing more makes acquiring additional signals less costly represents a process of refined search. The amount of information K is approximately the number of binary signals required the transmit information of that precision (Sims, 2003). Suppose that the first binary signal tells the observer whether the outcome is above or below the median. The second signal, in conjunction with the first, tells the observer which quartile the outcome is in, and so forth. If the outcomes are uniformly distributed, each signal is reducing the standard deviation by half (increasing the precision 4-fold). Increasing precision *proportionately* is always equally costly. The fact that the interpretation of the second signal depends on the first illustrates how existing information helps agents interpret new information more effectively.

While cost concavity may have a realistic foundation, it can also generate multiple equilibria. When a non-convex cost is subtracted from a concave objective, multiple utility-maximizing choices may arise. See Myatt and Wallace (2011) for a proof and examples.

2.3 Public signals (information markets)

One way of modeling information frictions is by assuming that agents can purchase signals from an information market (Veldkamp, 2006). Typically, the producer of the signal has to pay a fixed cost to discover the signal. Once discovered, he can replicate it and sell it to others. Because the seller is selling exact replicas of the same signal, it is a purely public signal ($\mathbf{b} = 0$). Agents choose a set J^i of signals to purchase and observe. Hellwig and Veldkamp (2009) consider a setting with a large number of signals and take a limit as the signal precision (d_j^{-2}) and the cost per signal approach zero. This effectively eliminates the discreteness in the choice variable.

Multiple equilibria. A key property of this problem is that in a symmetric equilibrium, $Var[\bar{a}|\mathcal{I}_i] = 0$ and $Cov[s, \bar{a}|\mathcal{I}_i] = 0$. Only the state is still uncertain: $Var[s|\mathcal{I}_i] = 1/(\tau_s + \sum_{j \epsilon J^i} d_j^{-2})$. An agent who learns less public information than others would reduce $Var[\bar{a}|\mathcal{I}_i]$ by learning more. But an agent who learns more public information than others does not change $Var[\bar{a}|\mathcal{I}_i]$. This features creates a discontinuity in the marginal utility of information that is responsible for multiple equilibria (Hellwig and Veldkamp, 2009).

The intuition is that, when actions are complements, public information is more valuble because it can be used both to forecast the state and to directly forecast others' actions (reduce $Var[\bar{a}|\mathcal{I}_i]$). Thus, the marginal value of public information exceeds the marginal value of private information. But learning one additional increment of public information, beyond what others have learned, is effectively learning private information. It is potentially public because others can learn that bit of information, but it is effectively private because others have chosen not to learn it. If others observe that additional public signal, then learning the signal has a higher marginal value because it lowers $Var[\bar{a}|\mathcal{I}_i]$. Learning that signal becomes a best response. If others choose not to learn that signal, it is effectively private, has lower value, and therefore may not be optimal to learn.

2.4 Correlated signals

Finally, we consider signals with both public and private noise. First, we fix the amount of public noise and allow agents to choose private noise, as in Myatt and Wallace (2011). Then, we fix the amount of private noise and allow agents to vary the weight their signal places on public noise. As in Myatt and Wallace (2011), we interpret a lower b_j as "paying more attention" and a lower d_j as "clarifying" signal j.

Such information choices affect all three sufficient statistics: First, equation (5) gives us $Var[s|\mathcal{I}_i]$. It reveals that more attention to signal j (lower b_j) lowers the conditional variance of the state forecast. Reducing b_j has a larger effect if d_j is also small and vice-versa: Paying attention is more valuable when the signal is clear; clearer signals are more valuable if one can pay close attention to them. Second, the covariance of this average action with the state s is given by (9). It is proportional to $Var[s|\mathcal{I}_i]$. The third statistic, the conditional variance of the average action, depends on how agents forecast others' signals and on the weight they place on the jth signal in actions (γ_j):

$$Var[\bar{a}|\mathcal{I}_i] = \sum_j \gamma_j^2 b_j^2 \frac{\tau_s^{-1} + d_j^2}{\tau_s^{-1} + d_j^2 + b_j^2}. \tag{10}$$

If agents pay little attention to signal j (b_j is large), then signal j becomes a private signal. This increases uncertainty about $\bar{a}$ because the agent has little idea of what signals others observe. If agents pay lots of attention to signal j (b_j is small), it becomes public. As all entries of $\mathbf{b}$ go to zero, agent i knows $\bar{a}$ with certainty: $lim_{\mathbf{b}\to 0} Var[\bar{a}|\mathcal{I}_i] = 0$. In between these two extremes, there is a continuous monotonic mapping that increases $Var[\bar{a}|\mathcal{I}_i]$ as signal correlation falls.

Choosing attention to public signals. In order to characterize equilibria, it is useful to simplify the setting. Suppose agent i observes two signals: $z_j^i = s + d_j u_j + b_j v_j^i$ for $j \in \{1, 2\}$. The d's are exogenous. Agents can choose private precisions b_1^{-2}, b_2^{-2}, subject to a cost function $c(b_1, b_2)$.

Proposition 2 *Suppose that information costs are a function of the sum of private precisions: $c(b_1^{-2} + b_2^{-2})$. Then the equilibrium information choice is unique.*

See Myatt and Wallace (2011). When actions are complements, choices of signal precision are complements as well. But, just like in the private signal case, this complementarity is not strong enough to generate multiple equilibria. Because agents are choosing the amount of private noise, rather than whether to see the next increment of information, the choice is continuous and there is no kink in the marginal benefit curve.

Choosing signal clarity. Consider the same signal structure as the previous section. But instead of fixing $\mathbf{d}$ and choosing $\mathbf{b}$, we fix $\mathbf{b}$ and allow agents to choose $\mathbf{d}$. One way to interpret this technology is that agents choose from a continuum of news outlets that have the same news with some common noise. But some outlets achieve a higher signal-to-noise ratio than others. In addition, agents may add independent signal processing noise to whatever they read, but they cannot control this processing noise.

Proposition 3 *When $c(\mathbf{d})$ is a convex function, there is a unique symmetric equilibrium in the choice of signal clarity $\mathbf{d}$.*

One might think that the choice of $\mathbf{d}$ and choosing how much of the newspaper to read (II.C) would be isomorphic problems. Moreover, it is not the presence of private signal noise that explains why one problem has multiple equilibria and the other does not. If $B = 0$, proposition 3 still holds. Rather, the key difference is that one problem has a continuous marginal utility and the other does not.

Clarity vs. quantity of public information. One key distinction between these information choice technologies is that in the newspaper model, an agent can decompose his signal into information that others see and information they do not. The information others observe has a discretely different marginal utility than the additional information others have not observed. That discrete difference creates the kink in utility and multiple equilibria. In the signal clarity problem, there is no such decomposition. In fact, if $B = 0$, then an agent who observed two signals with different degrees of clarity could infer the public noise u and the true state s exactly.

A second property that distinguishes the two technologies is that more signal clarity can lower expected utility. A precise signal about s reveals little about u and thus tells the observer little about what others know and what they will do. In other words, it can raise $Var[\bar{a}|\mathcal{I}_i]$. In the newspaper model, an agent who learns more information never forgets his existing information and therefore cannot become more uncertain about $\bar{a}$. Therefore, more information always increases expected utility.

Finally, a clearer signal has the same state s and same noise u with different weights on them. Learning more newspaper information could be represented as choosing a signal with more precision. But as the precision changed, the noise u, and its correlation with the u in others' signals, would have to change as well.

3 Conclusion

Formulating a problem with information choice requires a learning technology. Which technology is appropriate depends on the type of data agents are acquiring. Inattentiveness is a useful way to describe facts that can be objectively known and easily transmitted, e.g., one's bank balance, a stock price, or an election outcome. Looking up the result might require effort, but it is not likely to be observed with noise. Everyone who observes it knows that other observers have seen the same signal. Rational inattention is a useful way to describe more subjective evaluations, such as the probability of crisis, an optimal price, or future productivity. Shown the same data, reasonable people might come to different conclusions. More cognitive effort might improve estimates. Similarly, public signal choice describes a situation where the signal may not be right, but once we see the announcement, we all know what we saw and we know that other observers saw the same thing. Correlated signals represent both the idea that the underlying signal may have error and that agents may disagree about how to interpret that signal.

In each case, there was also a similarity: When agents want to do what other do, they want to know what others know. They also want to know more when others know more. Both the choice of which signals to observe and the precision with which to observe those signals exhibit the same strategic motives in actions. But there are various sources of non-concavities, discreteness in choice variables, or discontinuities in marginal utilities that can arise, depending on the information choice technology. When coupled with complementarity in information acquisition, these features can generate multiple equilibria.

References

Hellwig, Christian, and Laura Veldkamp. 2009. "Knowing What Others Know: Coordination Motives in Information Acquisition." *Review of Economic Studies*, 76: 223–251.

Mackowiak, Bartosz, and Mirko Wiederholt. 2009. "Optimal Sticky Prices Under Rational Inattention." *American Economic Review*, 99(3): 769–803.

Matejka, Filip. 2011. "Rationally Inattentive Seller: Sales and Discrete Pricing." CERGE Working Paper.

Myatt, David, and Chris Wallace. 2011. "Endogenous Information Acquisition in Coordination Games." *Review of Economic Studies*, forthcoming.

Reis, Ricardo. 2006. "Inattentive Producers." *Review of Economic Studies*, 73(3): 793–821.

Sims, Christopher. 2003. "Implications of Rational Inattention." *Journal of Monetary Economics*, 50(3): 665–690.

Veldkamp, Laura. 2006. "Media Frenzies in Markets for Financial Information." *American Economic Review*, 96(3): 577–601.

Veldkamp, Laura. 2011. *Information Choice in Macroeconomics and Finance.* Princeton University Press.

A Derivations and Proofs

A.1 Sufficient statistics with correlated signal errors

This appendix shows how to compute the conditional state and average action variances and their covariance when the set of signals an agent observes are correlated with each other.

For each information choice vector χ, we construct a corresponding $m \times n$ matrix X of zeros and ones. The number of rows m is the number of ones in χ. If the j^{th} entry of χ is 1, then there is a row of X that is all zeros, except for a one in the j^{th} position. Agent i's information set is thus summarized by the vector of signals he observes: $X\mathbf{z}^i$.

We begin by computing posterior beliefs, conditional on an information choice and signal realizations. Agents are trying to forecast both the true state and the common signal noise which determines the average action of other agents. Therefore, we define a $k+1 \times 1$ vector of both variables $\omega = \begin{bmatrix} s & \mathbf{u}' \end{bmatrix}'$. This is the relevant state variable. It is normally distributed with mean 0 and $Var(\omega) = I_{k+1}$. The covariance of the observed signals and this state is $Cov(\omega, X\mathbf{z}^i) = [\mathbf{1}_m, XD]$. The distribution of posterior beliefs comes from standard formulas for the conditional distribution of bivariate normals. Conditional on observing $\mathcal{I}_i$, ω is normally distributed with posterior mean and variance-covariance matrix

$$E(\omega|\mathcal{I}_i) = Cov(\omega, X\mathbf{z}^i)'\, Var(X\mathbf{z}^i)^{-1}\, X\mathbf{z}^i \tag{11}$$

$$\Sigma(\chi) := Var(\omega|\mathcal{I}_i) = Var(\omega) - Cov(\omega, X\mathbf{z}^i)'\, Var(X\mathbf{z}^i)^{-1}\, Cov(\omega, X\mathbf{z}^i) \tag{12}$$

where $Var(X\mathbf{z}^i) = X\Gamma\Gamma'X'$ and $\Gamma = \begin{bmatrix} \mathbf{1}_n & D & B \end{bmatrix}$.[4]

A.2 Proof of proposition 1

Proof of claim 1: Suppose all agents acquire full information. In this case $\Pi(b_i) = 0$ and each agent has to pay c. Deviating from the equilibrium means that the agent receives an uniformative signal, i.e. $b_i \to \infty$, but does not have to pay c. From the definition we can easily see that as $b_i \to \infty$, $\tau_v \to 0$. In the full information equilibirum, the agent knows that $\bar{a} = s$ and thus $Var[\bar{a}|\mathcal{I}_i] = Cov[\bar{a}, s|\mathcal{I}_i] = Var[s|\mathcal{I}_i]$. Put differently, the agent only needs to forecast s, as he knows the action of the other players. Thus, after deviation, the agent's payoff becomes

$$\Pi_i = -Var[s|\mathcal{I}_i] \tag{13}$$

$$= -\tau_s^{-1} \tag{14}$$

This deviation is strictly profitable if and only if

$$\tau_s^{-1} < c$$

From there it follows that an equilibrium with full revelation is sustainable if and only if $c \leq \tau_s^{-1}$, which is what we wanted to show.

[4] The relevant variance and covariance matrices are obtained by writing $\left[\omega' \ \left(X\mathbf{z}^i\right)'\right]'$ as

$$\begin{bmatrix} \omega \\ X\mathbf{z}^i \end{bmatrix} = \begin{bmatrix} I_{k+1} & 0 \\ \multicolumn{2}{c}{X\Gamma} \end{bmatrix} \begin{bmatrix} \omega \\ \mathbf{v}^i \end{bmatrix}$$

and using the fact that if a vector x is distributed according to $x \sim \mathcal{N}(\mu, I)$, then $Cx \sim \mathcal{N}(C\mu, CC')$.

Proof of claim 2: Suppose we are in an equilibrium, in which no agent has acquired information. In this case, all agents know that no other agent has acquired information so that for each i we have $a_i = \bar{a} = \mathbb{E}[s] = \mu$. Thus, $Var[s|\mathcal{I}_i] = \tau_s^{-1}$ and $Var[\bar{a}|\mathcal{I}_i] = Cov[s, \bar{a}|\mathcal{I}_i] = 0$. Hence, in equilibrium, the payoff for each agent π_i is given by

$$\pi_i = -(1-r)^2 \tau^{-1}$$

If the agent were to deviate, he would fully learn s and incur a cost of c. He still knows that $\bar{a} = \mu$. Thus, he would play $a_i = (1-r)s + r\mu$. We see from there, as well as from the fact that $Var[s|\mathcal{I}_i] = Var[\bar{a}|\mathcal{I}_i] = Cov[s, \bar{a}|\mathcal{I}_i] = 0$ that $\Pi = 0$. thus, the agent's payoff after deviation will be $-c$. From there it follows that a deviation is strictly profitable if and only if

$$(1-r)^2 \tau^{-1} > c$$

Hence, an equilibrium in which no agent acquires any information is sustainable as long as
$c \geq (1-r)^2 \tau_s^{-1}$, which is what we wanted to show.

Proof of claim 3: Suppose that we are in an equilibrium in which a fraction $\alpha \in (0,1)$ of the agents is informed and a fraction $(1-\alpha)$ is not. Denote by $\mathcal{I}^I$ an informed agents' information set and by $\mathcal{I}^U$ the uninformed agents information set. The informed agents know the precise value of s, the uninformed agents have to rely on the prior alone. We conjecture that in equilibrium both, informed and uninformed agents play linear strategies of their signal and the prior:

$$a_i^I = \bar{a}^I = \gamma_1 \mu + (1-\gamma_1) s a_i^U = \bar{a}^U = \gamma_2 \mu \tag{15}$$

Where we have used the fact that there is no idiosyncratic noise in the signals for the informed agents, since s is fully revealed. Also note that $\gamma_1, \gamma_2 \in [0,1]$. Let us now find γ_1 and γ_2 in such an equilibrium. The average action is given by:

$$\bar{a} = (\alpha\gamma_1 + (1-\alpha)\gamma_2)\mu + \alpha(1-\gamma_1)s$$

From there it follows that

$$\begin{aligned} \mathbb{E}\left[\bar{a}|\mathcal{I}^U\right] & \quad = (\alpha + (1-\alpha)\gamma 2)\mu \\ \mathbb{E}\left[\bar{a}|\mathcal{I}^I\right] & = \bar{a} = (\alpha\gamma_1 + (1-\alpha)\gamma_2)\mu + \alpha(1-\gamma_1)s \end{aligned} \tag{16}$$

From there, using the expression for the optimal action of agents we get:

$$\begin{aligned} \gamma_2 \mu & = r(\alpha + (1-\alpha)\gamma_2)\mu + (1-r)\mu \\ \Rightarrow \quad \gamma_2 & = r(\alpha + (1-\alpha)\gamma_2) + (1-r) \\ & = \frac{\alpha r + (1-r)}{1 - r(1-\alpha)} = 1 \end{aligned}$$

With this, we can then solve for γ_1 and $(1-\gamma_1)$:

$$\begin{aligned} \gamma_1 \mu + (1-\gamma_1)s & = r(\alpha\gamma_1 + (1-\alpha))\mu + (r\alpha(1-\gamma_1) + (1-r)s \\ \Rightarrow \quad \gamma_1 & = r(\alpha\gamma_1 + (1-\alpha)) \\ & = \frac{(1-\alpha)r}{1-\alpha r} \\ \Rightarrow \quad (1-\gamma_1) & = \frac{1-r}{1-\alpha r} \end{aligned}$$

Let us now turn to the sufficient statistics in such an equilibrium. Trivially, we have $Var\left[s|\mathcal{I}^I\right] = Cov\left[\bar{a}, s|\mathcal{I}^I\right] = Var\left[\bar{a}|\mathcal{I}^I\right] = 0$. For the uninformed agents, we get:

$$Var\left[s|\mathcal{I}^U\right] = \tau_s^{-1} \tag{17}$$

$$Cov\left[\bar{a}, s|\mathcal{I}^U\right] = \alpha(1-\gamma_1)\tau_s^{-1} \tag{18}$$

$$Var\left[\bar{a}|\mathcal{I}^I\right] = \alpha^2(1-\gamma_1)^2\tau_s^{-1} \tag{19}$$

Denoting by Π^U and Π^I the payoff of uninformed and informed agents respectively, we know that a necessary condition for a mixed equilibrium is $\Pi^U = \Pi^I$ Since we have

$$\Pi^I = r^2 Var\left[s|\mathcal{I}^I\right] + 2(1-r)rCov\left[\bar{a}, s|\mathcal{I}^I\right] + (1-r)^2 Var\left[\bar{a}|\mathcal{I}^I\right] - c = -c \tag{20}$$

$$\begin{aligned}\Pi^U &= r^2 Var\left[s|\mathcal{I}^U\right] + 2(1-r)rCov\left[\bar{a}, s|\mathcal{I}^U\right] + (1-r)^2 Var\left[\bar{a}|\mathcal{I}^U\right] \\ &= (r^2\alpha^2(1-\gamma_1)^2 + 2(1-r)\alpha(1-\gamma_1) + (1-r)^2)\tau_s^{-1} \\ &= [r\alpha(1-\gamma_1) + (1-r)]^2\tau_s^{-1}\end{aligned} \tag{21}$$

The necessary condition for a mixed equilibrium then becomes:

$$[r\alpha(1-\gamma_1) + (1-r)]^2 = \tau_s c \tag{22}$$

$$r\alpha\frac{(1-r)}{1-\alpha r} + (1-r) = \sqrt{\tau_s c} \tag{23}$$

$$1-r = \sqrt{\tau_s c}(1-\alpha r) \tag{24}$$

$$\alpha = \frac{\sqrt{\tau_s c} - (1-r)}{r\sqrt{\tau_s c}} \tag{25}$$

Since we require $\alpha > 0$, we get the first condition:

$$\sqrt{\tau_s c} - (1-r) > 0 \Rightarrow c > (1-r)^2\tau_s^{-1}$$

From $\alpha < 1$ we then get the second condition:

$$\begin{aligned} \sqrt{\tau_s c} - (1-r) &< r\sqrt{\tau_s c} \\ \Rightarrow \quad r\sqrt{\tau_s c} &< 1 \\ \Rightarrow \quad c &< \tau_s^{-1} \end{aligned}$$

A.3 Proof of proposition 2

See Myatt and Wallace (2011).

A.4 Proof of proposition 3

Our objective (1) is equivalent to maximizing the following linear quadratic utility functions:

$$u = \bar{u} - r(a_i - \bar{a})^2 - (1-r)(a_i - s)^2 - C(\psi_i)$$

Where s is some state of nature. Agents have an improper prior over s. We can now consider any number n of signals, each signal $z_{il}, l = 1, 2...n$ for agent i is given by

$$z_{il} = s + \psi_{il}^{-.5} u_l + B_l v_{il}$$

Where u_l is signal specific with distribution $u_l \sim \mathcal{N}(0,1)$ and $v_{il} \sim \mathcal{N}(0,1)$ is private noise. v and u are independent to one another as well as to s.

Each agent chooses his action a_i as well as his information strategy $\{\psi_{il}\}_{l=1}^{n}$ and has an information cost function $C(\psi_i)$, where ψ_i denotes the vector of information choices. The agents private noise component is determined by B_l but is fixed for each signal and cannot be altered by the agent.

We now first conjecture that each agent chooses a linear strategy in the n signals, assigning weights γ_{il} to each signal l such that $\sum_{l=1}^{n} = 1$ and their strategy is given by

$$\begin{aligned} a_i &= \sum_{l=1}^{n} \gamma_{il} z_{il} \\ &= s + \sum_{l=1}^{n} \gamma_{il} \left(\psi_{il}^{-.5} u_l + B_l v_{il} \right) \end{aligned} \tag{26}$$

Proofs for this are standard.

Furthermore, in any symmetric equilibrium, all agents choose the same $\psi_l = \psi_{il}\ \forall i$ and the same $\gamma_l = \gamma_{il}\ \forall i$ for each signal l.

Since the idiosyncratic noise washes out in the aggregate and all agents choose the same precision, in equilibrium the average action is given by

$$\bar{a} = s + \sum_{l=1}^{n} \gamma_l \psi_l^{-.5} u_l$$

Maximizing utility is the same as minimizing the loss function

$$L(\gamma_i, \psi_i) = r\mathbb{E}\left[(a_i - \bar{a})^2\right] + (1-r)\mathbb{E}\left[(a_i - s)^2\right] + C(\psi_i)$$

Let us first derive the first term of this expression:

$$\begin{aligned} (a_i - \bar{a})^2 &= \left[s + \sum_{l=1}^{n} \gamma_{il} \left(\psi_{il}^{-.5} u_l + B_l v_{il} \right) - s - \sum_{l=1}^{n} \gamma_l \psi_l^{-.5} u_l \right]^2 \\ &= \left[\sum_{l=1}^{n} \gamma_{il} B_l v_{il} + \sum_{l=1}^{n} (\gamma_{il} \psi_{il}^{-.5} - \gamma_l \psi_l^{-.5}) u_l \right]^2 \end{aligned} \tag{27}$$

Since we assumed u_l and v_{il} to be independent, the expectation becomes:

$$\mathbb{E}\left[(a_i - \bar{a})^2\right] = \textstyle\sum_{l=1}^{n} \gamma_{il}^2 B_l^2 + \sum_{l=1}^{n} \left(\frac{\gamma_{il}}{\sqrt{\psi_{il}}} - \frac{\gamma_l}{\sqrt{\psi_l}} \right)^2 \tag{28}$$

The second term can be found by looking at

$$\begin{aligned} (a_i - s)^2 &= \left[s + \sum_{l=1}^{n} \gamma_{il} \left(\psi_{il}^{-.5} u_l + B_l v_{il} \right) - s \right]^2 \\ &= \left[\sum_{l=1}^{n} \gamma_{il} (B_l v_{il} + \psi_{il}^{-.5} u_l) \right]^2 \end{aligned} \tag{29}$$

Again, because of independence we get

$$\mathbb{E}\left[(a_i - s)^2\right] = \textstyle\sum_{l=1}^{n} \gamma_{il}^2 \left(B_l^2 + \frac{1}{\psi_{il}} \right) \tag{30}$$

Combining these results yields:

$$
\begin{aligned}
L(\gamma_i, \psi_i) &= r\left[\sum_{l=1}^{n}\gamma_{il}^2 B_l^2 + \sum_{l=1}^{n}\left(\frac{\gamma_{il}}{\sqrt{\psi_{il}}} - \frac{\gamma_l}{\sqrt{\psi_l}}\right)^2\right] + (1-r)\left[\sum_{l=1}^{n}\gamma_{il}^2\left(B_l^2 + \frac{1}{\psi_{il}}\right)\right] + C(\psi_i) \\
&= \sum_{l=1}^{n}\gamma_{il}^2\left(B_l^2 + \frac{(1-r)}{\psi_{il}}\right) + r\sum_{l=1}^{n}\left(\frac{\gamma_{il}}{\sqrt{\psi_{il}}} - \frac{\gamma_l}{\sqrt{\psi_l}}\right)^2 + C(\psi_i) \qquad (31)
\end{aligned}
$$

Regardless of the exact equilibrium we are at, local to equilibrium strategies (γ, ψ) the second term of the above expression becomes zero and we can disregard it for the first order conditions, since at this point changes to the information strategy or action only have second order effects.

The fact that the middle term has zero partial derivative with respect to ψ_{il} when $\psi_{il} = \psi_l$ means that $\partial^2 L/\partial\psi_{il}\partial\psi_l = 0$ in a symmetric equilibrium. In other words, the precision that others choose does not affect the marginal benefit of additional precision for an individual. Thus, there is a unique crossing point where marginal benefit and marginal cost are equal. This unique crossing point determines the unique level of ψ_l that can be sustained as a symmetric equilibrium.