Attention, Coordination, and Bounded Recall*

Alessandro Pavan†

Northwestern University

November 2023

Abstract

I consider a class of strategic interactions under incomplete information in which, prior to finalizing their actions (consumption, production, or investment decisions), agents choose the attention to allocate to a large number of information sources about exogenous events that are responsible for the incompleteness of information (the underlying fundamentals). I study what type of payoff interdependencies contribute to inefficiency in the allocation of attention. I then compare the results for the benchmark of perfect recall (in which the agents remember the content of individual sources) to those for bounded recall (in which the agents are unable to keep track of the influence of individual sources on posterior beliefs). More generally, the analysis illustrates the implications (for attention and usage of information) of a certain form of bounded rationality whereby the summary statistics the agents recall from the sources they pay attention to is distorted away from the optimal action towards the Bayesian projection of the exogenous fundamentals over the signals received.

Keywords: attention, endogenous information, strategic complementarity/substitutability, externalities, bounded rationality, efficiency, welfare, bounded recall.

JEL classification: C72, D62, D83, E50

---

*For comments and useful suggestions, I thank seminar participants at the University of California Berkeley, University of Cambridge, the Catholic University of Milan, George Washington University, the Cornell-Penn State Macroeconomics Conference, the University of Essex, the Federal Reserve Bank of Chicago, the University of Iowa, Northwestern University, the University of Toronto, the University of Wisconsin at Madison, Iowa State University, Stanford University, the University of Zurich, the Barcelona Summer Forum on Information and Market Frictions, the EEA-ESEM Meetings, the second conference on Rational Inattention, and the ESSET Conference in Gerzensee. I thank George Marios Angeletos, Simone Galperti, Filip Matejka, David Myatt, Isabel Trevino, Yeneng Sun, and Chris Wallace for detailed feedback. Matteo Li Bergolis, Henrique Roscoe de Oliveira, and Sebastian Kohls provided excellent research assistance. The usual disclaimer applies.

†Department of Economics, Northwestern University; alepavan@northwestern.edu

1 Introduction

Many socio-economic interactions occur under incomplete information about relevant fundamentals affecting preferences and technology. For example, firms make real and nominal decisions under limited information about the demand for their products and/or the cost of their inputs; consumers choose consumption bundles under limited information about their preferences and needs; traders choose portfolios under limited information about the profitability of relevant assets; voters choose candidates under limited information about their valiance and policy platforms.

The incompleteness of information may either reflect limits on what is known to society as a whole (the long-run profitability of stocks is unknown to anyone), or individual constraints on the amount of information that each single decision maker can process. Time and cognition are limited, implying that the information that individuals use for most of their decisions is less than what is in the public domain.

Furthermore, in most situations of interest, individuals experience difficulty in keeping track of the influence of individual sources of information on their posterior beliefs. For example, an investor reading a large number of articles about a critical event responsible for the profitability of his investment decisions (e.g., the possibility of a war, or the default of a sovereign borrower) may have difficulty remembering how his beliefs have been influenced by the specific sources he attained to. Such a difficulty is irrelevant when the decision maker recalls what action maximizes his payoff given the information received. However, because the sources contain information about primitive events and not directly about optimal actions, when attaining to a large number of sources, what the agent recalls may be distorted away from the optimal action towards the Bayesian projection of the primitive events of interest (the exogenous fundamentals) over the information received. Importantly, this form of bounded rationality is relevant only in games. In single decision-maker problems, whether the decision maker remembers the individual content of the sources or only the Bayesian projection of the fundamentals on the information received is irrelevant. This is because optimal actions are measurable in the Bayesian statistics. Instead, in strategic settings, optimal actions naturally differ from the Bayesian statistics. This is because the individual sources may contain information about the actions of others that is not summarized in the Bayesian projection of the exogenous fundamental on the individual signals received. In other words, in general, the Bayesian projection need not be an appropriate summary statistics of the individual signals when it comes to predicting the *joint distribution* of the exogenous fundamentals and the other agents' actions.

In this paper, I investigate the implications (for the allocation of attention and the usage of information) of a form of bounded rationality that is meant to reflect the difficulty described above. Specifically, I study how agents change their allocation of attention to a large number of information sources when they expect that they will not be able to recall the information received from individuals sources and that what they will recall may be distorted away from their optimal action towards the Bayesian summary statistics of the sources' information (namely, the projection of the fundamentals over the signals received from the individual sources). The difficulty in recalling individual signals (or equivalently, the optimal action) is particularly relevant in environments in which the number of information sources is large (see, e.g., Kahneman (1973, 2011), and Kahneman, Slovic and Tversky (1982) for studies documenting such a difficulty).

The class of strategic situations I consider accommodates for various payoff interdependencies leading to either complementarity or substitutability in actions and responsible for various inefficiencies in the allocation of attention and the equilibrium usage of information. I first study the case of perfect recall and relate possible inefficiencies in the equilibrium allocation of attention and the subsequent usage of information to primitive conditions. This part brings under the same unifying umbrella results that are scattered in the literature and that have been established for specific linear-quadratic-Gaussian applications. The contribution of this part is twofold: (a) it summarizes and generalizes what is known in the literature, and (b) it permits me to compare the equilibrium acquisition and usage of information under perfect recall with those under bounded recall which I study in the second part of the paper.

Formally, bounded recall is a measurability constraint on the agents' actions. Agents allocate attention to a large number of information sources about exogenous fundamentals that matter for their payoffs, but their actions are measurable in a summary statistics of the information received from the individual sources. The statistics is distorted away from the perfect-recall equilibrium action towards the Bayesian projection of the underlying fundamentals on the signals received. The reason why the Bayesian projection may distort the agents' recollection is twofold. First and foremost, the information contained in the individual sources is about the exogenous fundamentals and not the equilibrium actions. These fundamentals stand for primitive events such as the likelihood of a war, the default of a financial institution, or the collapse of a political establishment, that are not directly influenced by the agents' actions. As a result, the agents, often involuntarily, summarize what they learned about such primitive events before they distill implications for their optimal actions. Second, the agents may use the information they learned about the exogenous fundamentals for many decisions other than the one in the specific strategic situation under consideration. When this is the case, the Bayesian statistics may naturally distort recall away from the equilibrium action for the specific strategic situation of interest.

Agents are sophisticated in that they understand how their allocation of attention shapes the comovement between their actions, the fundamentals, and other agents' actions. They also anticipate the difficulty recalling the content of individual sources (equivalently, the optimal action). They then adjust their allocation of attention to best account for such difficulties. A prominent case is when the statistics the agents recall coincides with the Bayesian projection (equivalently, with the agents' posterior beliefs about the exogenous fundamentals). When this is the case, the measurability constraint may also reflect the fact that, in certain environments of interest, agents do not really learn individually from the sources of information they pay attention to. Instead, these sources jointly contribute to a single signal, which is the agents' posterior belief about the event they are learning about, as often assumed in the rational-inattention and information-design literatures.

Contrary to the recent literature on rational inattention, I model explicitly the different sources of information. This permits me to compare the benchmark of perfect recall with the case of bounded recall. The information sources are modeled as in Dewan and Myatt (2008) and Myatt and Wallace (2012), with each source defined by its accuracy and transparency. The *accuracy* of a source is the precision of its content, whereas its *transparency* is the rate of return of attention to the source, that is, the extent to which additional attention to the source leads to a marginal reduction in the idiosyncratic interpretation of its content.

The first part of the paper considers the benchmark of perfect recall. It shows that any source that receives positive attention in equilibrium is characterized by a ratio between its transparency and its marginal cost of attention exceeding a critical threshold. In the special case in which the attention cost depends only on the total amount of attention, the result implies that only the most transparent sources receive attention in equilibrium, thus extending a property first noticed by Myatt and Wallace (2012) to the more general payoff structure considered in the present paper. I then compare the equilibrium allocation of attention to the efficient allocation of attention (defined as the one that maximizes the ex-ante utility of a representative agent) and I identify primitive conditions (on payoffs) that are responsible for inefficiencies in the equilibrium allocation of attention and in the usage of information. Broadly speaking, these inefficiencies originate in the interaction of two forces: (i) the value that each agent assigns to reducing the dispersion of her actions around the mean action, relative to the value that the planner assigns to the same reduction; and (ii) the reduction in the dispersion of individual actions around the mean action that obtains when individual actions are determined by the equilibrium strategies, relative to the reduction that obtains when actions are determined by the efficient rule.[1]

The contribution of these preliminary results is in permitting a comparison to the case of bounded recall, which is the subject of the second part of the paper. As mentioned above, bounded recall is a constraint on the agents' actions that imposes that the latter be measurable in a statistics that is distorted away from the perfect-recall equilibrium action towards the Bayesian projection. The first insight is that, with bounded recall, the benefit that each agent assigns to an increase in the attention allocated to any given source combines the reduction in the dispersion of her action around the mean action (as in the benchmark with perfect recall), with a novel effect that stems from the change in the distribution of the agent's own average action around its complete-information counterpart. This second effect is absent under perfect recall, and has important implications for the equilibrium allocation of attention. Relative to the case of perfect recall, I show that agents reallocate their attention from sources of low and high endogenous publicity (these are sources of, respectively, low and high transparency) to sources of intermediate publicity.

[1]These results generalize insights in Colombo, Femminis and Pavan (2014) to a setting with an arbitrarily large number of information sources—Colombo, Femminis and Pavan (1014) consider a setting with a single source of private information.

To understand the result, it is best to consider the case where the summary statistics coincides with the agents' posterior beliefs about the exogenous fundamentals (formally, the Bayesian projection of the fundamentals on the individual signals). Observe that sources of low publicity are sources whose ratio between transparency and accuracy is low. These sources serve the agents well in predicting the exogenous fundamentals but are poor coordination devices, given that their interpretation is largely idiosyncratic. In the case of perfect recall, paying attention to such sources is justified by the possibility of letting the sensitivity of the equilibrium actions to these sources be different from the sensitivity to other sources, thus limiting the impact of the idiosyncratic interpretation of these sources on the dispersion of individual actions around the mean action. Such a possibility is severed under bounded recall, thus reducing the benefit of paying attention to these sources. Sources of high publicity, instead, are sources whose ratio between transparency and accuracy is high. These sources may be imprecise when used to predict the exogenous fundamentals but serve the agents well when used to predict the forecasts, and hence ultimately the actions, of other agents. With bounded recall, however, because the agents cannot respond to these sources differently from how they respond to other sources, paying a lot of attention to these sources may lead to a high volatility of an agent's own expected action around its complete-information counterpart. Because such a volatility contributes negatively to individual payoffs, the benefit of paying attention to these sources is dampened relative to the case of perfect recall. Sources of intermediate publicity, instead, are good *compromises*: they are decent forecasters of both the underlying fundamentals and other agents' actions. As a result, with bounded recall, the benefit of paying attention to these sources is higher than under perfect recall. Assessing the empirical support of this prediction is fascinating but difficult because of the complexity of measuring (or even proxying for) how agents allocate their attention to a large number of information sources. However, that, in many situations of interest, agents favor information sources with an intermediate transparency/accuracy ratio seems broadly consistent with heuristics (see also the survey on the influence of economics and finance news outlets and journalists in Ragas and Tran (2015)).[2]

I conclude by investigating how bounded recall affects the (in)efficiency of the equilibrium allocation of attention. Inefficiencies now originate not only in the discrepancy between the private and the social value of reducing the dispersion of individual actions around the mean action, but also in the discrepancy between the private and the social value of reducing the dispersion of individual *average actions* around their complete-information counterparts. Despite these novel effects, the key normative insights from the benchmark with perfect recall carry over to the case of bounded recall. In particular, economies in which agents value coordination more than the planner typically feature

[2] The article analyzes the results of a survey investigating the influence of various US financial and economics media outlets and individual journalists. According to the article, *The Wall Street Journal* and Andrew Ross Sorkin of *The New York Times* are widely perceived by peers as the most influential financial media outlet and journalist, respectively. In many ways, these sources are neither the most specialized (and hence accurate) nor the most accessible ones. Yet, they strike a good balance between accuracy and transparency and are perceived as highly influential.

an excessively high allocation of attention to sources of high endogenous publicity (which in turn are sources with a high transparency/accuracy ratio), and an excessively low allocation of attention to sources of low publicity. The opposite property holds in economies in which the planner values aligning individual actions more than the agents.

The rest of the paper is organized as follows. I review the pertinent literature below. Section 2 contains all results for the case of perfect recall, while Section 3 contains all the results for bounded recall. Section 4 concludes. All proofs are in the Appendix at the end of the document.

1.1 Related literature

The paper belongs to the recent literature on attention and information acquisition in coordination environments.[3] As mentioned above, the description of the information sources is the same as in Dewan and Myatt (2008) and Myatt and Wallace (2012). The payoff structure, instead, is the same as in Angeletos and Pavan (2007). This more flexible payoff structure facilitates the analysis of the normative questions of interest. In fact, the payoff structure in Myatt and Wallace (2012) abstracts from externalities that create a wedge between the equilibrium and efficient acquisition and usage of information. Specifically, Myatt and Wallace (2012) consider a "potential" game, where the potential function is social welfare (see, e.g., Monderer and Shapley (1996)). This specification is thus appropriate for positive analysis, but not for identifying sources of inefficiency in the equilibrium acquisition and usage of information. The characterization of the equilibrium acquisition and usage of information in Proposition 1 in the present paper is similar to the one in Myatt and Wallace (2012). The key difference between Section 2 of this paper and Myatt and Wallace (2012) is the welfare analysis. Related welfare results are discussed in Myatt and Wallace (2014) for Lucas-Phelps economies, Myatt and Wallace (2015) for Cournot competition, and Myatt and Wallace (2018) for asymmetric and differentiated Bertrand competition. The contribution of Section 2 in the present paper vis-a-vis these previous works is in unifying the results and bringing them under the same framework. This in turn facilitates the comparison with the case of bounded recall in the second part of the paper, which is the paper's core contribution.

As mentioned above, the payoff structure in the present paper is the same as in Angeletos and Pavan (2007) and in Colombo, Femminis, and Pavan (2014). This structure is quite flexible and has been used both in the coordination literature (see, among others, Bergemann and Morris (2013), Vives (2017) and the references therein) and in the rational-inattention literature (see Sims (2003, 2011) for an introduction to this literature). The information structure, however, is more flexible. The richer information structure in the present paper is instrumental to the analysis of bounded recall. Banerjee et al. (2020) also consider a coordination model with motivated beliefs in which the payoff structure is the same as in the present paper but in which the agents choose how to interpret

[3] For an analysis of general properties of monotone equilibria in Bayesian games of strategic complementarity, see the earlier work by Van Zandt and Vives (2007) and the more recent work by Amir and Lazzati (2016).

the precision of their private information. The motivation is similar to the one discussed in the aspirational utility literature: in choosing how to interpret the various sources of information, agents trade off the ability to align their actions with the fundamentals and the actions of others with the possibility to derive a higher aspirational utility by believing the precision of their information is higher than the true one. Linear-quadratic-Gaussian models have also been used in the network literature (see, among others, Calvo-Armengol, Marti', and Prat (2015), and Galeotti, Golub and Goyal (2020)) and in the organization-economics literature (see, among others, Alonso, Dessein, and Matouschek (2008), and Dessein, Galeotti, and Santos (2016)). Lambert, Martini, and Ostrovsky (2018) provide general equilibrium existence and uniqueness results for quadratic games with flexible information structures.

Related are also Hebert and La'O (2022) and Angeletos and Sastry (2023). The first paper considers a coordination setting in which payoffs are as in Myatt and Wallace (2012). The analysis identifies key properties of the cost of information responsible for non-fundamental volatility and inefficiency in the usage of information. The former property obtains when learning from public signals (which contain noise at the source) is cheaper than learning the fundamentals directly (the paper identifies the precise property of the cost functional that is responsible for such a cost saving and show that it is violated e.g. under entropy reduction). The second property obtains when the agents' attention affects other agents' ability to learn. In the absence of such a learning externality, because welfare is the potential of the game, information is collected and used efficiently. Angeletos and Sastry (2023), instead, study the validity of the welfare theorems in economies with rationally-inattentive agents. The key finding is that these theorems hold provided that there are no learning externalities. Because markets are complete in Angeletos and Sastry (2023), the inefficiencies in the collection and usage of information discussed in the present paper do not arise in the economies considered in that paper.

Pavan, Sundaresan and Vives (2023) study optimal policy interventions in economies in which markets are incomplete and agents acquire private information and then submit price-contingent schedules. Angeletos and La'O (2020) study monetary policy in economies with endogenous dispersed information. Colombo, Femminis, and Pavan (2023), instead, study optimal fiscal and monetary policy in economies with endogenous private information and investment spillovers.

Hellwig and Veldkamp (2009) are the first to study the relation between the complementarity/substitutability in actions and the complementarity/substitutability in information acquisition. The information structure in that paper is different from the one in the present paper in that it assumes that the publicity of each source is exogenous and that the attention allocated to each source is binary. This last property can favor equilibrium indeterminacy. In contrast, the (symmetric) equilibrium is unique in the present paper, as well as in most of the papers cited above. Chahrour (2014) studies optimal central bank disclosures in an economy in which processing information is costly and in which agents may mis-coordinate on which sources they pay attention to. Herskovic and Ramos (2020) study coordination and information acquisition in a model of network formation in which agents learn from peers. Llosa and Venkateswaran (2023) compare the equilibrium acquisition of private information to its efficient counterpart in three different specifications of the business cycle.

The works mentioned above consider economies with continuous payoffs. Information acquisition in games of regime change (where payoffs are discontinuous) is studied in Szkup and Trevino (2015), Yang (2015), and Morris and Yang (2022). The first paper considers a canonical information structure with a single perfectly private additive signal whose precision is determined in equilibrium. The second paper considers a flexible information structure with an entropy cost and shows how the possibility to learn asymmetrically across states (which is appealing in discontinuous games) may lead to equilibrium indeterminacy. The third paper, instead, shows how equilibrium (in)determinacy relates to the possibility of distinguishing nearby states at a finite cost. Contrary to the present paper, these works assume that the signals the agents receive are independent, conditional on the underlying fundamental. This assumption is relaxed in Denti (2023); that paper shows, among other things, how the endogenous correlation in the noise in the agents' signals may contribute to equilibrium uniqueness.

As anticipated above, the present paper is also related to the literature on rational inattention, as pioneered by Sims. See, e.g., Sims (2003, 2011) for an earlier overview of this literature, Maćkowiak and Wiederholt (2009) for an influential business-cycle application, Matejka and McKay (2012, 2015) for how rational inattention provides a foundation for the multinomial logit model, Hébert and Woodford (2016), and Stewart et al (2016), for extensions to dynamic problems, and Maćkowiak et al. (2022) for an overview of the more recent literature. Among these papers, the closest is Maćkowiak and Wiederholt (2012). That paper compares the equilibrium allocation of attention to its efficient counterpart, assuming that the decision makers can absorb any information as long as the reduction in entropy is below a given threshold. In contrast, in the present paper, the cost of attention is smooth. The information structure is also different and permits me to investigate which dimension (transparency versus accuracy) receives more weight in equilibrium, and whether the equilibrium weights are socially inefficient. Importantly, none of these works studies how the allocation of attention changes if the agents do not recall the individual signals, or the equilibrium actions.

The paper is also related to the literature that investigates the effects of bounded memory on individual decision making (see, e.g., Mullainathan (2002), Benabou and Tirole (2004), Wilson (2004), and Kocer (2010)). This literature does not investigate how bounded memory influences the allocation of attention in a strategic setting, or the discrepancy between the equilibrium and the efficient allocation of attention. The effects of bounded recall in settings with strategic interactions are, instead, examined in the literature on dynamic (and repeated) games with imperfect information (see, e.g., Mailath and Samuelson (2006) and the references therein). The formalization of bounded recall, as well as the questions addressed in that literature, are, however, different from the ones in the present paper.

Finally, the paper is related to the literature that investigates how boundedly rational agents may group together different information sets into analogy-based equivalence classes when computing best responses, as pioneered by Jehiel (2005)—see also Jehiel and Samet (2007), Jehiel and Koessler (2008), and Jehiel and Samuelson (2012). In the present paper, the coarsening of the information sets is the one corresponding to the equivalence classes defined by the agents' summary statistics of the individual signals, with the latter coinciding with the Bayesian projection when the agents recall only their posterior beliefs about the underlying fundamentals.

2 Perfect Recall

2.1 Environment

Agents, Information Sources, and Attention. The economy is populated by a measure-one continuum of agents, indexed by i and uniformly distributed over $[0, 1]$. Each agent i has access to $N \in \mathbb{N}$ sources of information about a primitive payoff variable θ which is responsible for the incompleteness of information (hereafter, the *exogenous fundamentals*). Depending on the application of interest, such variable parametrizes a technology shock, a demand shifter, or the profitability of a new investment opportunity. Agents share a common prior that θ is drawn from a Normal distribution with mean zero and precision $\pi_\theta \equiv \sigma_\theta^{-2}$ (σ_θ^2 is thus the variance of the distribution).[4] The information contained in each source $n = 1, ..., N$ is given by

$$y_n = \theta + \varepsilon_n,$$

where ε_n is normally distributed noise, independent of θ and of any ε_s, $s \neq n$, with mean zero and precision η_n. By paying attention $z^i \equiv (z_n^i)_{n=1}^N \in \mathbb{R}_+^N$ to the various sources, agent $i \in [0, 1]$ then receives $x^i \equiv (x_n^i)_{n=1}^N \in \mathbb{R}^N$ *private* signals, with each x_n^i given by

$$x_n^i = y_n + \xi_n^i,$$

where ξ_n^i is idiosyncratic noise, normally distributed, with mean zero and precision $t_n z_n^i$, drawn independently of θ, $\varepsilon \equiv (\varepsilon_n)_{n=1}^N$, and ξ_s^j, with $s = 1, ..., N$ for $j \neq i$, and $s = 1, ..., n-1, n+1, ..., N$ for $j = i$. The parameter $\eta_n \in \mathbb{R}_+$ measures the source's *accuracy*, whereas the parameter t_n its *transparency* (the extent to which a marginal increase in the *attention* z_n^i allocated to the source reduces the idiosyncratic interpretation of its content). One can think of z_n^i as the "time" or "effort" allocated to interpreting source n.

Actions and Payoffs. Let $k^i \in \mathbb{R}$ denote agent i's action, $K \equiv \int_j k^j dj$ the mean action in the population, and $\sigma_k^2 \equiv \int_j \left[k^j - K\right]^2 dj$ the dispersion of individual actions around the population mean action. Each agent's payoff is given by the (expectation of the) Bernoulli utility function

$$u\left(k^i, K, \sigma_k, \theta\right) - C(z^i),$$

[4] That the prior mean is zero simplifies the formulas, without any important effect on the results.

where $C(z^i)$ denotes the attention cost incurred by the agent. I assume that C is increasing, convex, and continuously differentiable.[5]

The function u is a second-order polynomial, which can be interpreted as an approximation of some more general function. I also assume that dispersion σ_k has only a second-order non-strategic external effect, so that $u_{k\sigma} = u_{K\sigma} = u_{\theta\sigma} = 0$ and that $u_\sigma(k, K, 0, \theta) = 0$, for all (k, K, θ).[6,7] The assumption that u is quadratic ensures the linearity of the agents' best responses and simplifies the analysis.

In addition to the above assumptions, I also assume that partial derivatives satisfy the following conditions: (i) $u_{kk} < 0$, (ii) $\alpha \equiv -u_{kK}/u_{kk} < 1$, (iii) $u_{kk} + 2u_{kK} + u_{KK} < 0$, (iv) $u_{kk} + u_{\sigma\sigma} < 0$, and (v) $u_{k\theta} \neq 0$. As shown in Angeletos and Pavan (2007), Condition (i) imposes concavity at the individual level, so that best responses are well defined. Condition (ii) implies that the slope of best responses is less than one, which in turn guarantees uniqueness of the equilibrium actions, for any given allocation of attention. Conditions (iii) and (iv) guarantee that the first-best allocation is unique and bounded. Finally, Condition (v) ensures that the fundamental θ affects equilibrium behavior, thus making the analysis non-trivial.

Timing. Agents simultaneously choose the attention they allocate to the various sources of information. Each agent then receives private signals x^i. Finally, agents simultaneously commit their actions, and payoffs are realized.

2.2 The equilibrium allocation of attention

First note that, under complete information about θ, the unique equilibrium features each agent taking the action $k^i = \kappa$ where $\kappa \equiv \kappa_0 + \kappa_1\theta$, with $\kappa_0 \equiv \frac{-u_k(0,0,0,0)}{u_{kk}+u_{kK}}$ and $\kappa_1 \equiv \frac{-u_{k\theta}}{u_{kk}+u_{kK}}$. Now consider the problem of an agent $j \in [0,1]$ who allocated attention z^j to the various sources of information and received the signals x^j. Optimality requires that, for any x^j, the agent's action satisfies[8]

$$k^j = \mathbb{E}[(1-\alpha)\kappa + \alpha K \mid z^j, x^j], \tag{1}$$

where $\alpha \equiv u_{kK/}|u_{kk}|$ measures the slope of individual best responses to aggregate activity.

[5]The assumption that C is convex need not be compatible with an entropy-based cost function (that is, a cost function increasing in the mutual information between $y \equiv (y_n)_{n=1}^N$ and $x^i \equiv (x_n^i)_{n=1}^N$), as in certain models of rational inattention. With that type of cost function, equilibrium uniqueness cannot be guaranteed for sufficiently high degrees of coordination. However, even in that case, social welfare continues to be concave in the allocation of attention, meaning that the efficient allocation of attention remains unique. Besides, all key results pertaining to (a) the comparison between the equilibrium allocation of attention and the efficient allocation of attention and (b) the comparison of the equilibrium with perfect recall and the one with bounded recall are established by looking at the *gross* private benefit of increasing the attention allocated to any given source. Because of this, all key results extend to a situation in which the attention cost is concave, even if in the latter case equilibrium uniqueness cannot be guaranteed.

[6]The notation u_k denotes the partial derivative of u with respect to k, whereas the notation u_{kK} denotes the cross derivative with respect to k and K. Similar notation applies to the other arguments of the utility function.

[7]In other words, u is additively separable in σ_k^2 with coefficient $u_{\sigma\sigma}/2$.

[8]This step follows from Angeletos and Pavan (2009)—Proposition 3.

Next, consider the agent's choice of attention. Suppose that all agents allocate attention z to the various sources of information. The (endogenous) *precision* of each source $s = 1, ..., N$ is then given by

$$\pi_s \equiv \frac{\eta_s z_s t_s}{z_s t_s + \eta_s}$$

and is increasing in the source's accuracy η_s, its transparency t_s, and the attention z_s allocated to the source. Denote by

$$\varphi_s^j \equiv \varepsilon_s + \xi_s^i$$

the total noise in the signal agent j receives from source s, and by

$$\rho_s \equiv corr(\varphi_s^j, \varphi_s^i) = \frac{z_s t_s}{z_s t_s + \eta_s}$$

the correlation in the noise among any two different agents $i, j \in [0,1]$, $i \neq j$. Following Myatt and Wallace (2012), I refer to ρ_s as the source's *endogenous publicity.* Finally, let

$$C_n'(z) \equiv \partial C(z)/\partial z_n$$

denote the marginal cost of increasing the attention allocated to the n-th source of information, starting from z. The following result is then true:

Proposition 1. *There exists a unique symmetric equilibrium. In this equilibrium, the attention $\hat{z}$ that each agent $i \in [0,1]$ allocates to the various sources of information is such that, for any source $n = 1, ..., N$ that receives strictly positive attention*[9]

$$C_n'(\hat{z}) = \frac{|u_{kk}|}{2} \frac{(\kappa_1 \gamma_n(\hat{z}))^2}{(\hat{z}_n)^2 t_n}, \tag{2}$$

where, for any z,

$$\gamma_n(z) \equiv \frac{\frac{(1-\alpha)\pi_n(z)}{1-\alpha\rho_n(z)}}{\pi_\theta + \sum_{s=1}^N \frac{(1-\alpha)\pi_s(z)}{1-\alpha\rho_s(z)}}, \text{ with } \pi_s(z) = \frac{\eta_s z_s t_s}{z_s t_s + \eta_s} \text{ and } \rho_s(z) = \frac{\pi_s(z)}{\eta_s}, \ s = 1, ..., N. \tag{3}$$

Given the equilibrium allocation of attention $\hat{z}$, the equilibrium actions are given by[10]

$$k^i = k(x^i; \hat{z}) = \kappa_0 + \kappa_1 \left(\textstyle\sum_{n=1}^N \gamma_n(\hat{z}) x_n^i \right), \text{ all } i \in [0,1], \text{ all } x^i \in \mathbb{R}^N. \tag{4}$$

[9] For any source that receives no attention

$$C_n'(\hat{z}) \geq \frac{|u_{kk}|}{2} \frac{(\kappa_1)^2 (1-\alpha)^2 t_n}{\left[\pi_\theta + \sum_{s=1}^N \frac{(1-\alpha)\eta_s \hat{z}_s t_s}{(1-\alpha)\hat{z}_s t_s + \eta_s}\right]^2} = \lim_{z_n \to 0^+} \frac{|u_{kk}|}{2} \frac{(\kappa_1 \gamma_n(\hat{z}_{-n}, z_n))^2}{(z_n)^2 t_n},$$

where $\hat{z}_{-n} \equiv (\hat{z}_1, ..., \hat{z}_{n-1}, \hat{z}_{n+1}, ..., \hat{z}_N)$.

[10] Here I follow the pertinent literature and, with abuse of notation, denote by $k(\cdot; \hat{z})$ the function mapping the individual signals x^i into the actions k^i, when the agents' attention is $\hat{z}$.

To understand the result, note that, when, given attention z, all agents follow the strategy $k(\cdot;z)$ in (4), in equilibrium, the dispersion of individual actions in the population is given by

$$Var[k - K \mid z,\ k(\cdot;z)] = \kappa_1^2 \sum_{s=1}^{N} \frac{\gamma_s^2(z)}{z_s t_s},$$

where $\kappa_1 \gamma_s(z)$ is the influence of each source s on the equilibrium actions. Differentiating $Var[k-K \mid z,\ k(\cdot;z)]$ with respect to z_n while keeping fixed the strategy $k(\cdot;z)$ as defined in (4) (for all agents, including agent i), then reveals that the *private benefit* of increasing the attention allocated to each source n (the right-hand side in (2)) is equal to

$$\frac{|u_{kk}|}{2} \frac{(\kappa_1 \gamma_n(z))^2}{(z_n)^2 t_n} = \frac{|u_{kk}|}{2} \left| \frac{\partial}{\partial z_n} Var[k - K \mid z,\ k(\cdot;z)] \right|. \tag{5}$$

In equilibrium, the marginal benefit that each agent assigns to paying more attention to any given source of information thus coincides with the marginal reduction in the dispersion of the individual's action around the mean action, weighted by the importance $|u_{kk}|/2$ that the individual assigns to such a reduction. Importantly, the reduction in dispersion is computed by holding fixed the strategy $k(\cdot;z)$ (from the usual envelope arguments, the agent expects the information to be used optimally once collected). As I show below, this interpretation helps understanding the sources of inefficiency in the equilibrium allocation of attention.

Also note that, fixing the equilibrium allocation of attention $\hat{z}$, the influence $\kappa_1 \gamma_n(\hat{z})$ that each source n exerts on the equilibrium actions (as per (4)) increases with the source's endogenous precision π_n and increases with the source's endogenous publicity ρ_n when agents value positively aligning their actions with the actions of others (i.e., when $\alpha > 0$), whereas it decreases when they value such an alignment negatively (i.e., when $\alpha < 0$). In turn, both the precision π_n and the publicity ρ_n of any given source increase with the source's accuracy η_n and with its transparency t_n. Finally, note that, when $\alpha \to 0$, the sensitivity of the equilibrium actions to each source of information converges to $\kappa_1 \delta_n$ with

$$\delta_n \equiv \frac{\pi_n}{\pi_\theta + \sum_{s=1}^{N} \pi_s}.$$

This limit corresponds to a single decision maker's problem, in which the relative influence of any two sources of information is given by their relative informativeness, as captured by the ratio between the two sources' precisions. In contrast, when $\alpha \to 1$, $\gamma_n \to 0$ for all $n = 1, ..., N$: as the agents' concern for aligning their actions with the actions of others grows large, they ignore all sources of information and base their actions on the common prior.

The following is then also true:

Corollary 1. *There exists a threshold $R > 0$ such that, in the unique symmetric equilibrium, for any source that receives strictly positive attention*

$$\frac{t_n}{C_n'(\hat{z})} > R,$$

whereas for any source that receives no attention $t_n/C'_n(\hat{z}) \leq R$.

When a source's transparency is low and/or the marginal cost of expanding the attention to the source is high, paying attention to the source is not worth the cost, given that the reduction in the idiosyncratic interpretation of the source's content is small. The source thus does not receive any attention in equilibrium.[11] Also note that the attention that the sources receive in equilibrium need not be monotone in their transparency, even when the marginal cost is constant across the sources. This is because, when transparency is high, a small amount of attention suffices to almost completely eliminate any idiosyncratic interpretation of a source's content; as a result, attention can be maximal for intermediate degrees of transparency.

In general, solving for the equilibrium allocation of attention in close form can be tedious at this level of generality. Fortunately, none of the results below requires arriving at close-form solutions. However, a special case where close-form solutions can easily be obtained is when the cost is linear and small enough that all sources receive positive attention in equilibrium.

Example 1. *Suppose that there exists $\bar{c} \in \mathbb{R}_{++}$ such that, for any $z \in \mathbb{R}^N_+$, $C(z) = \bar{c} \cdot \sum_{s=1}^N z_s$, and assume that $\bar{c}$ is sufficiently small that all sources receive strictly positive attention in equilibrium. The attention that each source receives is then given by*

$$\hat{z}_n = \frac{\eta_n}{\sqrt{t_n}(1-\alpha)} \left[\frac{(1-\alpha)\kappa_1 \sqrt{\frac{|u_{kk}|}{2\bar{c}}} + \sum_{s=1}^N \frac{\eta_s}{\sqrt{t_s}}}{\pi_\theta + \sum_{s=1}^N \eta_s} - \frac{1}{\sqrt{t_n}} \right]. \tag{6}$$

The example illustrates the general properties discussed above that attention is increasing in accuracy, but possibly non-monotone in transparency. It also shows that, under the assumed cost functional, as the value of coordination α increases, the attention allocated to sources of low transparency decreases, whereas the attention allocated to sources of high transparency increases.[12] Finally, it shows that the total amount of attention decreases with the coordination motive, α.[13]

2.3 The efficient allocation of attention

I now turn to the allocation of attention that maximizes the ex-ante utility of a representative agent. The analysis permits me to identify payoff interdependencies that, under perfect recall, are

[11] Corollary 1 extends to the more general environment under consideration a property first noticed in Myatt and Wallace (2012) that only those sources of sufficiently high transparency receive attention in equilibrium.

[12] Formally,

$$\frac{\partial \hat{z}_n}{\partial \alpha} < 0 \text{ if } \sqrt{t_n} \leq \left(\frac{\pi_\theta + \sum_{s=1}^N \eta_s}{\sum_{s=1}^N \frac{\eta_s}{\sqrt{t_s}}} \right) \text{ and } \frac{\partial \hat{z}_n}{\partial \alpha} > 0 \text{ if the previous inequality is reversed.}$$

[13] This is not immediate to see, but can be verified by differentiating $\hat{Z} \equiv \sum_n \hat{z}_n$ with respect to α and using the property that

$$\left(\sum_{s=1}^N \frac{\eta_s}{\sqrt{t_s}} \right)^2 \leq \sum_{s=1}^N \frac{\eta_s}{t_s} \sum_{s=1}^N \eta_s.$$

responsible for inefficiency in the equilibrium allocation of attention. These results, when applied to specific applications, may guide policy interventions aimed at increasing the efficiency of market interactions.

First, observe that, for any allocation of attention z, the efficient use of information consists in all agents following the unique strategy $k^*(\cdot; z)$ that solves the functional equation[14]

$$k\left(x; z\right) = \mathbb{E}\left[\left(1 - \alpha^*\right) \kappa^* + \alpha^* K \ \mid z, x\right] \text{ for all } x \in \mathbb{R}^N, \tag{7}$$

where $\kappa^* = \kappa_0^* + \kappa_1^* \theta$ is the first-best allocation[15], $K = \mathbb{E}\left[k\left(x; z\right) \ \mid z, \theta, \varepsilon\right]$ is the average action, and

$$\alpha^* \equiv \frac{u_{\sigma\sigma} - 2u_{kK} - u_{KK}}{u_{kk} + u_{\sigma\sigma}} \tag{8}$$

is the socially optimal degree of coordination (that is, the level of complementarity, or substitutability, that the planner would like the agents to perceive in order for the equilibrium of the economy to coincide with the efficient allocation.) Because (7) differs from the equilibrium optimality condition (1) only by the fact that α is replaced by α^* and κ by κ^*, it is then immediate that the efficient strategy takes the linear form

$$k^*(x; z) = \kappa_0^* + \kappa_1^* \left(\textstyle\sum_{n=1}^N \gamma_n^*(z) x_n\right), \tag{9}$$

where

$$\gamma_n^*(z) \equiv \frac{\frac{(1-\alpha^*)\pi_n(z)}{1-\alpha^*\rho_n(z)}}{\pi_\theta + \sum_{s=1}^N \frac{(1-\alpha^*)\pi_s(z)}{1-\alpha^*\rho_s(z)}} = \frac{\frac{(1-\alpha^*)\eta_n z_n t_n}{(1-\alpha^*)z_n t_n + \eta_n}}{\pi_\theta + \sum_{s=1}^N \frac{(1-\alpha^*)\eta_s z_s t_s}{(1-\alpha^*)z_s t_s + \eta_s}}$$

has the same structure as $\gamma_n(z)$ in (3), but with α^* replacing α.

Next note that, for any attention z, welfare under the efficient use of information $k^*(\cdot; z)$ can be expressed as

$$w^*(z) \equiv \mathbb{E}[u(\kappa^*, \kappa^*, 0, \theta)] - \mathcal{L}^*(z) - C(z),$$

where $\mathbb{E}[u(\kappa^*, \kappa^*, 0, \theta)]$ is expected welfare under the first-best allocation, whereas

$$\mathcal{L}^*(z) \equiv \frac{|u_{kk} + 2u_{kK} + u_{KK}|}{2} Var[K - \kappa^* \mid z, k^*(\cdot; z)] + \frac{|u_{kk} + u_{\sigma\sigma}|}{2} Var[k - K \mid z, k^*(\cdot; z)]$$

combines the welfare losses originating in the volatility of the average action K around its first-best counterpart with the losses originating in the dispersion of individual actions around the mean action.

I now turn to the efficient allocation of attention. Using the envelope theorem and observing that, holding fixed the strategy $k^*(\cdot; z)$, the volatility of the aggregate action around its complete-information counterpart $Var[K - \kappa^* \mid z, k^*(\cdot; z)]$ is independent of the allocation of attention, I have

[14]The characterization of the efficient use of information follows from steps similar to those in Angeletos and Pavan (2009). The contribution here is in the characterization of the efficient allocation of attention.

[15]The scalars κ_0^* and κ_1^* are given by $\kappa_0^* = \frac{u_k(0,0,0) + u_K(0,0,0)}{-(u_{kk} + 2u_{kK} + u_{KK})}$ and $\kappa_1^* = \frac{u_{k\theta} + u_{K\theta}}{-(u_{kk} + 2u_{kK} + u_{KK})}$, respectively.

that the *social benefit* of increasing the attention allocated to any source n (gross of its cost) is given by[16]

$$\frac{|u_{kk}+u_{\sigma\sigma}|}{2}\left|\frac{\partial}{\partial z_n}Var[k-K \mid z,k^*(\cdot;z)]\right| = \frac{|u_{kk}+u_{\sigma\sigma}|}{2}\frac{\left(\kappa_1^*\gamma_n^*(z)\right)^2}{\left(\hat{z}_n\right)^2 t_n} \tag{10}$$

where $\partial Var[k-K \mid z,k^*(\cdot;z)]/\partial z_n$ is computed holding fixed the efficient strategy $k^*(\cdot;z)$ that maps the signals x into individual actions. In other words, the social benefit of allocating more attention to any given source is given by the reduction in the dispersion of individual actions around the mean action that obtains when agents allocate more attention to that source, weighted by the social aversion to dispersion $|u_{kk}+u_{\sigma\sigma}|/2$. The following result then follows from the arguments above:

Proposition 2. *Suppose that the planner can control the use of information (i.e., can dictate to the agents the mapping from their signals to their actions). There exists a unique allocation of attention z^* that maximizes welfare. Under such an allocation, for any source n that receives strictly positive attention,*[17]

$$C_n'(z^*) = \frac{|u_{kk}+u_{\sigma\sigma}|}{2}\frac{\left(\kappa_1^*\gamma_n^*(z^*)\right)^2}{\left(z_n^*\right)^2 t_n},$$

where $\kappa_1^*\gamma_n^*(z^*)$ represents the influence of the source on the agents' actions under the efficient strategy $k^*(\cdot;z^*)$.

The following conclusion can then be established by comparing the private benefit (5) to the social benefit (10) of increasing the attention allocated to any given source of information:

Corollary 2. *Let $\hat{z}$ denote the equilibrium allocation of attention. Suppose that the planner can control the use of information. Then, starting from $\hat{z}$, forcing the agents to pay more attention to a source that receives strictly positive attention in equilibrium (i.e., for which $\hat{z}_n > 0$) increases welfare if*

$$|u_{kk}|(\kappa_1\gamma_n(\hat{z}))^2 < |u_{kk}+u_{\sigma\sigma}|(\kappa_1^*\gamma_n^*(\hat{z}))^2, \tag{11}$$

and decreases it if the inequality in (11) is reversed, where $\kappa_1\gamma_n(\hat{z})$ and $\kappa_1^*\gamma_n^*(\hat{z})$ denote, respectively, the sensitivity of the equilibrium and of the efficient actions to the n-th source of information, when

[16] As in the equilibrium case, the expression in (10) applies to sources that receive strictly positive attention (that is, for which $z_n > 0$). The marginal benefit of increasing the attention allocated to a source that receives zero attention is simply the limit of the right-hand side of (10) as $z_n \to 0^+$ which is equal to

$$\frac{|u_{kk}+u_{\sigma\sigma}|}{2}\frac{(\kappa_1^*)^2(1-\alpha^*)^2 t_n}{\left[\pi_\theta + \sum_{s=1}^N \frac{(1-\alpha^*)\eta_s z_s t_s}{(1-\alpha^*)z_s t_s+\eta_s}\right]^2}.$$

[17] As in the equilibrium case, for any source that receives no attention, the following condition must hold:

$$C_n'(z^*) \geq \frac{|u_{kk}+u_{\sigma\sigma}|}{2}\frac{(\kappa_1^*)^2(1-\alpha^*)^2 t_n}{\left[\pi_\theta + \sum_{s=1}^N \frac{(1-\alpha^*)\eta_s z_s t_s}{(1-\alpha^*)z_s t_s+\eta_s}\right]^2} = \lim_{z_n\to 0^+}\frac{|u_{kk}+u_{\sigma\sigma}|}{2}\frac{(\kappa_1^*\gamma_n^*(z_{-n}^*,z_n))^2}{\left(z_n\right)^2 t_n}.$$

the attention allocated to the various sources is $\hat{z}$. Likewise, forcing the agents to pay attention to a source n that receives no attention in equilibrium (i.e., for which $\hat{z}_n = 0$) increases welfare if

$$\frac{|u_{kk} + u_{\sigma\sigma}|}{2} \frac{(\kappa_1^*)^2(1-\alpha^*)^2 t_n}{\left[\pi_\theta + \sum_{s=1}^N \frac{(1-\alpha^*)\eta_s \hat{z}_s t_s}{(1-\alpha^*)\hat{z}_s t_s + \eta_s}\right]^2} > C_n'(\hat{z})$$

and decreases it if the inequality is reversed.

To understand the result, recall from the analysis above that both the private and the social (gross) marginal benefit of allocating more attention to any given source come from the marginal reduction in the dispersion of individual actions around the mean action.[18] The magnitude of this reduction depends on the sensitivity of individual actions to the source, which is given by $\kappa_1 \gamma_n$ under the equilibrium strategy, and by $\kappa_1^* \gamma_n^*$ under the efficient strategy. The weight that the planner assigns to reducing the dispersion of individual actions is $|u_{kk} + u_{\sigma\sigma}|$, whereas the weight that each individual agent assigns to reducing the dispersion of her action around the mean action is $|u_{kk}|$. Staring from the equilibrium allocation of attention $\hat{z}$, forcing the agents to increase the attention they allocate to any source that receives strictly positive attention in equilibrium then increases welfare if and only if the marginal reduction in the dispersion of actions under the equilibrium strategy, weighted by the importance that each agent assigns to dispersion, falls short of the marginal reduction in dispersion under the efficient strategy, weighted by the importance that the planner assigns to dispersion. Likewise, for any source that receives no attention in equilibrium, the marginal cost exceeds the private marginal benefit of reducing dispersion. Staring from the equilibrium allocation of attention $\hat{z}$, forcing the agents to pay attention to these sources increases welfare if and only if the marginal cost also falls short of the social marginal benefit of reducing dispersion. Put differently, efficiency in the allocation of attention requires both (i) efficiency in the use of information and (ii) alignment between the private and the social value of reducing the dispersion of individual actions, which obtains when, and only when, there are no externalities from the dispersion of individual actions around the mean action, i.e., $u_{\sigma\sigma} = 0$.

To appreciate the implications of Corollary 2, it is useful to focus on a few benchmark cases. First, consider economies in which the usage of information is efficient (i.e., $\kappa = \kappa^*$ and $\alpha = \alpha^*$). When agents suffer from the dispersion of individual actions, i.e., when $u_{\sigma\sigma} < 0$, starting from the equilibrium allocation of attention, the planner can increase welfare by forcing the agents to pay more attention to any source that receives strictly positive attention in equilibrium, whereas the opposite is true when agents benefit from such a dispersion, i.e., when $u_{\sigma\sigma} > 0$. Next, consider economies in which the inefficiency in the allocation of attention is due to the sensitivity κ_1 of the complete-information equilibrium actions to the fundamentals (i.e., $\alpha = \alpha^*$, $U_{\sigma\sigma} = 0$, but $\kappa_1 \neq \kappa_1^*$). When $|\kappa_1| < |\kappa_1^*|$, the planner can increase welfare by forcing the agents to pay more attention to

[18] By usual envelope arguments, both marginal reductions are computed holding constant, respectively, the equilibrium and the efficient strategies, that is, the mappings $k(\cdot; z)$ and $k^*(\cdot; z)$.

all sources that receive strictly positive attention in equilibrium, whereas the opposite is true when $|\kappa_1| > |\kappa_1^*|$, that is, when, under complete information, agents over-respond to the fundamentals. Lastly, consider economies in which the inefficiency in the allocation of attention is due to the discrepancy $\alpha - \alpha^*$ between the equilibrium and the efficient degrees of coordination. In the Online Appendix, I show that there exists a threshold $R^* > 0$ such that, when agents over-value aligning their actions with the actions of others (i.e., when $\alpha > \alpha^*$), starting from the equilibrium allocation of attention $\hat{z}$, the planner can increase welfare by forcing the agents to pay more attention to sources of low transparency (namely, for which $\hat{z}_n > 0$ and $t_n/C_n'(\hat{z}) < R^*$) and less attention to sources of high transparency (namely, for which $\hat{z}_n > 0$ and $t_n/C_n'(\hat{z}) > R^*$). In other words, economies in which agents are over-concerned with aligning their actions with the actions of others are economies in which agents pay too much attention to sources of high transparency (which are good coordination devices) and too little attention to sources of low transparency. The opposite conclusions hold in economies in which agents undervalue aligning their actions with the actions of others, i.e., for which $\alpha < \alpha^*$. I also show that, when the cost C depends only on the total attention, too few sources receive strictly positive attention in equilibrium and the total attention $\hat{Z} \equiv \sum_{s=1}^{N} \hat{z}_s$ allocated to the various sources is inefficiently low when $\alpha > \alpha^*$, whereas the opposite conclusions hold when $\alpha < \alpha^*$ (See Claim 1 in the Online Appendix).

Results qualitatively similar to those discussed above obtain in economies in which the *planner can not control the agents' actions*. In these economies, the social benefit of changing the agents' attention to the sources must also account for the effect that such a change has on the equilibrium usage of information, i.e., the mapping from the agents' signals x^j to the agents' actions k^j (see Claim 2 in the Online Appendix). In concrete micro-founded applications, these results have direct implications for the design of policies that, by manipulating the equilibrium actions (for examples through taxes and subsidies on individual activity) also manipulate the collection of information.

3 Bounded Recall

The results in the previous section are for economies in which, at the time of committing their actions, the agents recall the information they received from each individual source. Equivalently, the agents summarize all the information received into a statistics that permits them to play the same actions as when they remember the content of individual sources (which is the case when the statistics is the equilibrium action itself). I now turn to the (perhaps more plausible) case of bounded recall in which the agents anticipate a difficulty in keeping track of the content of individual sources and do not trust that what they will recall will permit them to play the same equilibrium actions as when they perfectly recall.

3.1 Environment

To facilitate the comparison with the benchmark of full recall, I assume that the statistics that each agent recalls is linear in the signals received. Formally, let $\lambda \equiv (\lambda_n(\cdot))_{n=1}^N$ be a collection of functions, one for each source, with $\lambda_n : \mathbb{R}_+^N \to [0,1]$, $n = 1, ..., N$, such that $\sum_{s=1}^N \lambda_s(z^j) = 1$. Given the attention $z = (z_n)_{n=1}^N$ allocated to the different sources, the statistics that agent j expects to recall is given by

$$X^j \equiv \sum\nolimits_{n=1}^N \lambda_n(z) x_n^j, \tag{12}$$

where, as in the previous section, $x^j \equiv (x_n^j)_{n=1}^N$ are the signals the agent received from the various sources. The statistics provides the agent with a signal of θ whose precision is equal to

$$\pi_X(z) \equiv \left[\sum\nolimits_{s=1}^N \lambda_s(z)^2 \left(\pi_s(z)\right)^{-1} \right]^{-1}, \tag{13}$$

where

$$\pi_s(z) \equiv \frac{\eta_s z_s t_s}{z_s t_s + \eta_s}$$

continues to denote the (endogenous) precision of the information received from source n. Furthermore, when any two agents $i, j \in [0,1]$, $j \neq i$, allocate the same attention z to the sources of information, the correlation in the error in the above statistics across the two agents is given by

$$\begin{aligned} \rho_X(z) &\equiv Corr\left(\sum\nolimits_{n=1}^N \lambda_n(z)(\varepsilon_n + \xi_n^i); \sum\nolimits_{n=1}^N \lambda_n(z)(\varepsilon_n + \xi_n^j) \right) \\ &= \sum\nolimits_{n=1}^N \left(\frac{\lambda_n(z)^2 \left(\pi_n(z)\right)^{-1}}{\sum_{s=1}^N \lambda_s(z)^2 \left(\pi_s(z)\right)^{-1}} \right) \rho_n(z), \end{aligned} \tag{14}$$

where

$$\rho_n(z) \equiv Corr(\varepsilon_n + \xi_n^j, \varepsilon_n + \xi_n^i) = \frac{z_n t_n}{z_n t_n + \eta_n}$$

continues to denote the (endogenous) correlation in the error contained in source n, across the two individuals, as in the benchmark with perfect recall. In other words, $\rho_X(z)$ is a weighted average of the publicity $\rho_n(z)$ of the individual sources $n = 1, ..., N$.

Next, for any n, any $\beta \in [0,1]$, any z, let

$$\lambda_n^\beta(z) \equiv \frac{\frac{\pi_n(z)}{1-\beta\alpha\rho_n(z)}}{\sum_{s=1}^N \frac{\pi_s(z)}{1-\beta\alpha\rho_s(z)}}. \tag{15}$$

Note that the case where the statistics X^j the agents recall permits them to play the same actions as in the benchmark of perfect recall corresponds to the case in which the functions λ take the form in (15) with $\beta = 1$. On the other hand, the case in which what the agents recall is simply their posterior belief about the fundamentals θ corresponds to the case in which the functions λ are those in (15) with $\beta = 0$. As anticipated in the Introduction, the latter case also captures the possibility that, in certain environments, there are no source-specific signals x^j and, instead, given the attention

z^j allocated to the sources the agent receives a single signal X^j about θ with endogenous precision $\pi_X(z^j)$ whose error has correlation $\rho_X(z^j)$ with the error of any other agent allocating the same attention to the sources. More generally, the specification in (15) captures the idea that, in many situations of interest, the agents may expect the statistics X^j to be tilted away from the equilibrium action of the specific strategic situation under consideration towards the Bayesian projection of the fundamentals θ on the information received. This tilt in turn may reflect the expectation that the information about θ received from the various sources is relevant also for decision problems other than the one under consideration. Under the specification in (15), this possibility amounts to discounting the strategic relevance of each source n, as captured by the term $\alpha\rho_n$, in favor the source's accuracy π_n.

In the rest of this section, I consider the implications of bounded recall for the equilibrium and the efficient allocation of attention. Importantly, to isolate the effects of bounded recall, I maintain the assumption that the agents do not choose what to recall, which amounts to treating the functions $\lambda(\cdot)$ defining the summary statistics as exogenous.[19] Apart from the change described above, the environment is the same as in the previous section.

3.2 Equilibrium allocation of attention

Bounded recall amounts to imposing that agent j's actions be measurable in the sigma algebra generated by the statistics X^j. In this sense, the equilibrium under bounded recall amounts to a specific form of *analogy-based equilibrium* (as defined in Jehiel 2005) in which the coarsening of the partitions of the agents' information sets is generated by grouping together information sets corresponding to the same value of the statistics X^j. The next proposition characterizes the equilibrium allocation of attention when each agent correctly anticipates that what he will recall at the time of committing his action is the statistics X^j.

Proposition 3. *Suppose that, given any allocation of attention z, what each agent j recalls of the individual signals x^j is summarized in a statistics X^j given by (12) with arbitrary weights λ. There is a unique symmetric equilibrium. In this equilibrium, given the attention $z^\#$ allocated to the various sources of information, individual actions are given by*

$$k^i = k^\#(X^i; z^\#) = \kappa_0 + \kappa_1 \gamma^\#(z^\#) \cdot X^j \tag{16}$$

all $i \in [0,1]$, where

$$\gamma^\#(z) \equiv \frac{\frac{(1-\alpha)\pi_X(z)}{1-\alpha\rho_X(z)}}{\pi_\theta + \frac{(1-\alpha)\pi_X(z)}{1-\alpha\rho_X(z)}}. \tag{17}$$

Furthermore, for any source $n = 1, ..., N$ that receives strictly positive attention in equilibrium,

$$C_n'(z^\#) = -\frac{|u_{kk}|}{2}\frac{\partial}{\partial z_n} Var\left[k - K; z^\#, k^\#(\cdot; z^\#)\right] - \frac{|u_{kk}|}{2}(1-\alpha)\frac{\partial}{\partial z_n} Var\left[K - \kappa; z^\#, k^\#(\cdot; z^\#)\right] \tag{18}$$

[19] Clearly, if the agents could choose what to recall, in equilibrium, they would simply recall their optimal actions making the analysis identical to the one in the previous section.

where the derivatives are computed holding fixed the mapping $k^{\#}(\cdot; z^{\#})$ *given by (16).*

There are important differences relative to the case of perfect recall. First, the marginal benefit of increasing the attention allocated to each source now has two components. The first one is the marginal reduction of the dispersion of individual actions around the mean action. This component is similar to the one in the case of perfect recall and is computed holding fixed the mapping $k^{\#}(\cdot; z^{\#})$ governing the agents' actions (equivalently, the subsequent usage of information) by usual envelope arguments. Importantly, in a symmetric equilibrium, the reduction of dispersion of individual actions around the mean action is the same irrespective of whether one changes only the individual's attention or all agents' attention (this observation, which is formally proved in the proof of Proposition 3 in the Appendix, is important when comparing the equilibrium with the efficient allocation of attention).[20]

The second component reflects the fact that, with bounded recall, a change in attention also affects the dispersion of each agent's expected action around the complete-information counterpart (once again, with the change computed holding the mapping $k^{\#}(\cdot; z^{\#})$ fixed). The reason is that a change in attention changes the weights $\lambda_n(z)$ in the statistics X^j and hence affects not only the volatility of the statistics X^j but also its mean given the aggregate variables (θ, ε) The second term in the right-hand side of (18) thus represents the marginal benefit of bringing an agent's own expected action, which in a symmetric equilibrium coincides with the average action in the population, closer to the complete-information equilibrium action, κ. Importantly, while the weight the individual assigns to reducing the dispersion of his own action around the mean action continues to be given by the curvature of the individual payoffs, u_{kk}, the weight the individual assigns to reducing the volatility of his expected action around the complete-information counterpart is given by $|u_{kk}|(1-\alpha) = -(u_{kk} + u_{kK})$, and accounts also for the response of the agent's action to variations in the average action.

To facilitate the comparison with the benchmark of perfect recall, the next proposition assumes that the weights $\lambda_n(z)$ in the statistics X^j are as in (15); as explained above, the special case where what each agent recalls is simply his posterior beliefs about θ corresponds to $\beta = 0$. The following result is then true:

Proposition 4. *Suppose that, given the allocation of attention* z*, what each agent* j *recalls of the individual signals* x^j *is summarized in a statistics* X^j *given by (12) with weights given by (15), with* $\beta < 1$*. Starting from any allocation of attention* z *that is symmetric across the agents, there exist thresholds* ρ', ρ'' *with* $0 \leq \rho' \leq \rho'' \leq 1$ *such that the following is true: relative to the case of perfect recall, the benefit of locally increasing the attention allocated to source* n *is (weakly) larger if*

[20] This property is also true in the benchmark with bounded recall. There, however, the result is obvious, given that the distribution of the average action K is independent of the allocation of attention. In contrast, with bounded recall, the distribution of the average action depends on the allocation of attention, even when one holds fixed the mapping $k^{\#}(\cdot; z)$. The reason is that the allocation of attention impacts the weights λ assigned by the summary statistics to the various sources of information.

$\rho_n(z) \in [\rho', \rho'']$ and (weakly) smaller if $\rho_n(z) \notin [\rho', \rho'']$. Furthermore, when β is small (e.g., when the agent recalls only her posterior beliefs about θ, i.e., when $\beta = 0$) and α is large, $0 < \rho' < \rho'' < 1$.

To fix ideas, consider the case where what the agents recall is their posterior about θ (that is, $\beta = 0$); the discussion below, however, applies more generally to any $\beta < 1$. Observe that the endogenous publicity of any source n is given by

$$\rho_n(z) = \frac{\pi_n(z)}{\eta_n} = \frac{z_n t_n}{z_n t_n + \eta_n};$$

as explained above, the latter measures how the total error $\varphi_n^j = \varepsilon_n + \xi_n^j$ in the source (combining the error at the origin, ε_n, with the error ξ_n^j in the agent's idiosyncratic interpretation of the source's content) correlates across any two agents. Sources of low publicity are sources whose endogenous precision $\pi_n(z)$ is small relative to the source's exogenous accuracy, η_n. A low publicity in turn may be due to either a low transparency t_n of the source or little attention z_n allocated to it. The information received from such a source is thus subject to significant idiosyncratic noise in the agent's interpretation of the source's content. Relative to the case of perfect recall, the benefit of increasing the attention to such a source is smaller under bounded recall because of the impossibility for the agent to respond separately to the source's noise φ_n^j. Sources of high publicity, instead, are sources of potentially low accuracy but of high transparency or that receive significant attention. These sources serve primarily as coordination devices. With bounded recall, however, the coordination value of these sources is dampened because of the impossibility to respond separately to the noise associated with their interpretation. As a result, the benefit of expanding the attention to such sources is again smaller than under perfect recall. Finally, consider sources of intermediate publicity. These are good "compromises", in that they permit the agent to align his action well both with the fundamentals and with the other agents' actions. The benefit of expanding the attention to such sources under bounded recall is thus higher than under perfect recall.

In the case in which the attention cost depends only on total attention, the monotone relationship between the publicity of the sources in the benchmark of perfect recall and their exogenous transparency further permits me to establish the following result:

Corollary 3. *Suppose that there exists an increasing, convex, and differentiable function $c(\cdot)$ such that, for any z, $C(z) = c\left(\sum_{s=1}^{N} z_s\right)$. Let $\hat{z}$ be the allocation of attention in the unique symmetric equilibrium under perfect recall. There exist thresholds $t', t'' \in \mathbb{R}_+$ such that, starting from $\hat{z}$, any agent with bounded recall can (weakly) increases his payoff by (a) locally increasing the attention to any source for which $\hat{z}_n > 0$ and $t_n \in [t', t'']$ and (b) locally decreasing the attention to any source for which $\hat{z}_n > 0$ and $t_n \notin [t', t'']$.*

The results in Proposition 4 and Corollary 3 refer to local properties of best responses, evaluated around the equilibrium allocation of attention $\hat{z}$ in the benchmark with perfect recall. Similar conclusions hold when one compares the allocation of attention in the unique symmetric *equilibrium* with bounded recall to its counterpart under perfect recall. For simplicity, the result below is established for $\beta = 0$ (that is, for the case in which what the agents recall is simply their posterior beliefs about θ). A similar result holds for $\beta \in (0, 1)$, but with lengthier derivations in the proof.

Proposition 5. *Suppose that there exists an increasing, convex, and differentiable function $c(\cdot)$ such that, for any z, $C(z) = c\left(\sum_{s=1}^{N} z_s\right)$. Let $\hat{z}$ be the allocation of attention in the unique symmetric equilibrium with perfect recall and $z^{\#}$ the corresponding allocation of attention when, under bounded recall, the weights in the statistics X^j are given by (15) with $\beta = 0$. There exist thresholds $t', t'' \in \mathbb{R}_{++}$ such that $z_n^{\#} > \hat{z}_n$ only if $t_n \in [t', t'']$. Furthermore for any n for which $t_n \in [t', t'']$, $z_n^{\#} < \hat{z}_n$ only if $z_n^{\#} = 0$.*

The result in Proposition 5 thus establishes that it is only those sources whose transparency is intermediate that receive more attention under bounded recall than under perfect recall. In this sense, Proposition 5 extends the results in Proposition 4 and Corollary 3 from individual best responses to equilibrium allocations. The key property in the Appendix that permits me to establish the result in the proposition is that, among those sources that do receive some attention under bounded recall, those whose transparency is the highest are also those whose publicity is the highest. Recall that this property also holds under perfect recall. In that benchmark, the monotonicity extends to all sources, implying that it is only those sources whose transparency is high enough that receive some attention in equilibrium. I could not establish this stronger property under bounded recall. In other words, I could not exclude the possibility that source n with transparency t_n receives some attention whereas source n' with transparency $t_{n'} > t_n$ does not. This explains why the result in the proposition is not an "if and only if" result. However, what I could establish is that if a source of intermediate transparency receives less attention under bounded recall than under perfect recall, then it receives no attention at all.

As anticipated in the introduction, assessing the empirical support of the above predictions is quite complex because of the difficulty of measuring how agents in the field allocate their attention to a large number of information sources. However, that, in many situations of interest, sources of intermediate publicity (with an intermediate transparency/accuracy ratio) receive most of the attention seems to square well with heuristics as well as with what reported by Ragas and Tran (2015). In a survey of approximately 400 US-based journalists, it was found that the *The Wall Street Journal* and Andrew Ross Sorkin of *The New York Times* are widely perceived as the most influential financial media outlet and journalist, respectively. While many other sources appear either more sophisticated or easier to read, the influence of these sources seems to reflect a good balance between accuracy and transparency.

3.3 Efficient allocation of attention

I conclude by investigating the allocation of attention that maximizes welfare under bounded recall. The analysis is motivated by the interest in establishing whether, at least at a qualitative level, the results identified above by focusing on the benchmark of perfect recall remain valid under the (probably more plausible) case of bounded recall.

Consistently with the analysis above, I assume that the planner cannot choose the functions $\lambda_n(\cdot)$ defining the weights in the statistics X^j, but can control the agents' attention and their response to the statistics X^j (if the planner could also choose the weights λ in the statistics X^j, the efficient allocation of attention would be the same as in the previous section).

First note that, because the planner's problem is concave, it is never optimal to induce different agents to allocate different attention to the various sources of information. This in turn means that, for any symmetric allocation of attention z, efficiency in the agents' actions requires that, for any agent $j \in [0, 1]$, almost any value of the statistics X^j, the agent's action be given by

$$k^j = k^{**}(X^j; z) = \kappa_0^* + \kappa_1^* \gamma^{**}(z) X^j \tag{19}$$

with

$$\gamma^{**}(z) \equiv \frac{\frac{(1-\alpha^*)\pi_X(z)}{1-\alpha^*\rho_X(z)}}{\pi_\theta + \frac{(1-\alpha^*)\pi_X(z)}{1-\alpha^*\rho_X(z)}}, \tag{20}$$

where π_X and ρ_X are as in (13) and (14) above.[21] This implies that, for any allocation of attention z, the maximum welfare that can be achieved by having the agents follow the rule $k^{**}(\cdot; z)$ defined by (19) is given by

$$w^*(z) \equiv \mathbb{E}[u(\kappa^*, \kappa^*, 0, \theta)] - \mathcal{L}^*(z) - C(z), \tag{21}$$

where $u(\kappa^*, \kappa^*, 0, \theta)$ continues to denote welfare under the first-best allocation and where

$$\mathcal{L}^*(z) \equiv \frac{|u_{kk} + u_{\sigma\sigma}|}{2} Var[k - K \mid z, k^{**}(\cdot; z)] + \frac{|u_{kk} + 2u_{kK} + u_{KK}|}{2} Var[K - \kappa^* \mid z, k^{**}(\cdot; z)]$$

continues to denote the welfare losses due to the incompleteness of information (combining the losses from the dispersion of individual actions around the mean action with the losses stemming from the volatility of the average action around its first-best counterpart). Using the fact that $|u_{kk} + 2u_{kK} + u_{KK}| = (1 - \alpha^*) |u_{kk} + u_{\sigma\sigma}|$ and denoting by z^{**} the efficient allocation of attention with bounded recall, I then have any source of information that receives strictly positive attention under z^{**} must satisfy

$$\begin{aligned} C_n'(z^{**}) = &- \frac{|u_{kk} + u_{\sigma\sigma}|}{2} \frac{\partial}{\partial z_n} Var[k - K \mid z^{**}, k^{**}(\cdot; z^{**})] \\ &- (1 - \alpha^*) \frac{|u_{kk} + u_{\sigma\sigma}|}{2} \frac{\partial}{\partial z_n} Var[K - \kappa^* \mid z^{**}, k^{**}(\cdot; z^{**})], \end{aligned} \tag{22}$$

where all derivatives are computed holding fixed the mapping $k^{**}(\cdot; z^{**})$.

[21] The result follows from the observation that bounded recall is mathematically equivalent to a setting in which agents receive a single signal X^j about θ with precision $\pi_X(z)$ and correlation $\rho_X(z)$. The arguments that lead to the results below are then similar to those derived in the previous section.

Comparing the social to the private marginal benefit of increasing the attention allocated to any given source around the equilibrium levels then permits me to establish the following result (again, for simplicity, the result here is for $\beta = 0$, meaning that what the agents recall is their posterior belief about θ):

Proposition 6. *Suppose the weights in the statistics X^j are as in (15) with $\beta = 0$ and the planner can control the agents' response to the statistics X^j but cannot choose the weights in the statistics X^j. Let $z^{\#}$ denote the allocation of attention in the unique symmetric equilibrium with bounded recall.*

(a) Consider economies that are efficient in their use of information ($\kappa = \kappa^*$ and $\alpha = \alpha^*$). Starting from the equilibrium allocation of attention $z^{\#}$, forcing the agents to pay more attention to a source that receives strictly positive attention in equilibrium (i.e., for which $z_n^{\#} > 0$) increases welfare if $u_{\sigma\sigma} < 0$ and decreases if $u_{\sigma\sigma} > 0$.

(b) Consider economies in which there are no externalities from dispersion ($u_{\sigma\sigma} = 0$), and the equilibrium and the socially optimal degrees of coordination coincide ($\alpha = \alpha^*$). Starting from the equilibrium allocation of attention $z^{\#}$, forcing the agents to pay more attention to a source that receives strictly positive attention in equilibrium (i.e., for which $z_n^{\#} > 0$) increases welfare if $\kappa_1 < \kappa_1^*$ and decreases it if $\kappa_1 > \kappa_1^*$.

(c) Consider economies in which the sensitivity of the complete-information equilibrium actions to the fundamentals is first-best efficient (i.e., $\kappa_1 = \kappa_1^*$) and in which there are no externalities from the dispersion of individual actions (i.e., $u_{\sigma\sigma} = 0$). When $\alpha(\gamma^{\#}(z^{\#}))^2 > \alpha^*(\gamma^{**}(z^{\#}))^2$, there exists a critical threshold $\bar{\rho}(z^{\#}) \in [0, 1]$ such that, starting from the equilibrium allocation of attention $z^{\#}$, the planner would like the agents to reduce the attention they allocate to sources of high endogenous publicity (namely, for which $\rho_n(z^{\#}) > \bar{\rho}(z^{\#})$) and increase the attention to sources of low endogenous publicity (for which $\rho_n(z^{\#}) < \bar{\rho}(z^{\#})$ and $z_n^{\#} > 0$). The opposite conclusions hold for economies in which $\alpha(\gamma^{\#}(z^{\#}))^2 < \alpha^*(\gamma^{**}(z^{\#}))^2$. In the limit in which $\pi_\theta \to 0$, $\bar{\rho}(z^{\#}) \to \rho_X(z^{\#})$ and $\gamma^{\#}(z^{\#}), \gamma^{**}(z^{\#}) \to 1$, implying that the planner would like the agents to allocate less attention to sources for which $(\alpha - \alpha^*)\left[\rho_n(z^{\#}) - \rho_X(z^{\#})\right] > 0$ and more attention to sources for which the inequality is reversed and $z_n^{\#} > 0$.

The results in parts (a) and (b) parallel those in the benchmark with perfect recall. Those in part (c) are formally different but qualitatively similar to those for perfect recall: When agents overvalue aligning their actions with the actions of others, the planner would like the agents to reduce the attention they allocate to sources of high endogenous publicity and increase the attention they allocate to sources of low endogenous publicity; the opposite conclusions hold when agents undervalue aligning their actions with the actions of others. The inefficiency in the allocation of attention under bounded recall is thus qualitatively similar to the one under perfect recall.

4 Conclusions

In this paper, I compare the equilibrium to the efficient allocation of attention in a framework featuring a rich set of payoff interdependencies and a large number of information sources differing in their accuracy and transparency. I then examine how the allocation of attention is affected by bounded recall, namely by the difficulty of keeping track of the content of individual sources and by the expectation that what will be recalled is tilted away from the optimal action towards a Bayesian summary statistics of the information received.

In future work, it would be interesting to extend the analysis to dynamic settings in which agents solve a stopping problem by choosing, in each period, whether to collect further information from additional sources or irreversibly commit to an action. It would also be interesting to examine how the allocation of attention interacts with the market provision of information by endogenizing the supply of information. Such an extension would also permit one to investigate policy interventions aimed at correcting the interaction between the inefficiencies in the usage and in the provision of information.

Appendix

Proof of Proposition 1. When all agents allocate attention $\hat{z}$ to the various sources of information, the continuation game in which the agents receive information x^i and choose their actions has a unique continuation equilibrium where all agents $i \in [0,1]$ follow the linear strategy in (4). This result follows from arguments similar to those that lead to Proposition 3 in Angeletos and Pavan (2009) — the proof is thus omitted.

Next, let

$$U^j(z^j;\hat{z}) = \mathbb{E}[u(k^j, K, \sigma_k, \theta)|z^j] - C(z^j)$$

denote agent j's expected payoff when all agents $i \neq j$ pay attention $\hat{z}$ to the different sources of information and then choose their actions according to (4), whereas agent j allocates attention z^j to the various sources and then chooses his actions optimally. It is easy to show that $U^j(z^j;\hat{z})$ is continuously right-differentiable in z_n^j, any n, any $(z^j;\hat{z})$, and that, for any $z_n^j > 0$ the derivative $\partial U^j(z^j;\hat{z})/\partial z_n^j$ coincides with the partial derivative of the agent's expected payoff holding fixed the agent's optimal strategy $k_i(\cdot;z^j;\hat{z})$ by usual envelope arguments.

Next, note that when $z^j = \hat{z}$, by symmetry, the agent's optimal strategy coincides with the one of any other agent, that is, $k_i(\cdot;z^j;\hat{z}) = k(\cdot;\hat{z})$ with $k(\cdot;\hat{z})$ given by (4). Furthermore, when all agents (including agent j) follow the linear strategy in (4), for any choice of z^j, agent j's expected payoff is given by

$$\mathbb{E}[u(K,K,\sigma_k,\theta) \mid z^j, k(\cdot;\hat{z})] + \frac{u_{kk}}{2} Var[k^j - K \mid z^j, k(\cdot;\hat{z})] - C(z^j) \tag{23}$$

where the first term in the right-hand side of (23) is the payoff the agent would obtain if his action coincided with the average action in the population in every state, while the second term is the ex-ante dispersion of the agent's own action around the mean action. Note that, when all agents follow the linear strategy in (4) — more generally, when their actions are determined by any linear mapping of their signals — the distribution of K is independent of the allocation of attention. It follows that, in any symmetric equilibrium, for any source n that receives positive attention

$$\frac{\partial U^j(\hat{z};\hat{z})}{\partial z_n^j} = \frac{u_{kk}}{2}\frac{\partial}{\partial z_n^j} Var[k^j - K \mid \hat{z}, k(\cdot;\hat{z})] - C_n'(\hat{z}) \tag{24}$$

where the derivative in the right hand side of (24) is computed holding fixed the mapping $k(\cdot;\hat{z})$ and letting such mapping be the one given by (4).

Next observe that, when all agents follow the mapping in (4),[22]

$$Var[k^j - K \mid z^j, k(\cdot;\hat{z})] = \kappa_1^2 \sum\nolimits_{n=1}^{N} \frac{(\gamma_n(\hat{z}))^2}{z_n^j t_n}.$$

[22] Note that, when $z_n^j = \hat{z}_n = 0$, $(\gamma_n(\hat{z}))^2/z_n^j t_n = 0$. The contribution of source n to the dispersion of the agent's own action around the mean action can thus be written as $(\gamma_n(\hat{z}))^2/z_n^j t_n$ for any source, irrespective of whether or not such a source receives attention in equilibrium.

I conclude that, in any symmetric equilibrium, for any source of information that receives strictly positive attention, the following optimality condition must hold:

$$C_n'(\hat{z}) = \frac{|u_{kk}|}{2} \frac{(\kappa_1 \gamma_n(\hat{z}))^2}{\left(\hat{z}_n^j\right)^2 t_n}.$$

By continuity of the right-hand derivative $\partial U_+^j(z^j;\hat{z})/\partial z_n^j$, I also have that, for any source that receives no attention, the following corner condition must hold

$$C_n'(\hat{z}) \geq \frac{|u_{kk}|}{2} \frac{(\kappa_1 \gamma_n(\hat{z}))^2}{\left(\hat{z}_n^j\right)^2 t_n} = \frac{|u_{kk}|\kappa_1^2(1-\alpha)^2 t_n}{2\left[\pi_\theta + \sum_{s=1}^N \frac{(1-\alpha)\pi_s(\hat{z})}{1-\alpha\rho_s(\hat{z})}\right]^2} = \frac{|u_{kk}|}{2} \frac{(\kappa_1)^2(1-\alpha)^2 t_n}{\left[\pi_\theta + \sum_{s=1}^N \frac{(1-\alpha)\eta_s \hat{z}_s t_s}{(1-\alpha)\hat{z}_s t_s + \eta_s}\right]^2},$$

which is equivalent to the condition that $\partial U_+^j(\hat{z};\hat{z})/\partial z_n^j \leq 0$ at $\hat{z}_n = 0$.

Lastly, to see that the symmetric equilibrium is unique, let $\mathcal{U}$ denote the family of quadratic payoff functions satisfying all the conditions in the model setup. From arguments similar to those that lead to Proposition 2 in Angeletos and Pavan (2009), one can show that, given any $u \in \mathcal{U}$, there exists a unique $u' \in \mathcal{U}$ such that any symmetric equilibrium of the game where payoffs are given by u coincides with one of the efficient allocations for the economy with payoffs given by u'. Next observe that the efficient allocation for the economy with payoffs given by u' is unique – this follows from the fact that the planner's problem consisting in choosing a vector $z \in \mathbb{R}_+^N$ along with a function $k : \mathbb{R}^N \to \mathbb{R}$ so as to maximize the ex-ante expectation of u' is strictly concave. This in turn implies that the symmetric equilibrium for the economy with payoffs given by u is also unique, which establishes the result. Q.E.D.

Proof of Corollary 1. From Proposition 1, any source that receives strictly positive attention in equilibrium must satisfy (2). Substituting for

$$\gamma_n(\hat{z}) = \frac{\frac{(1-\alpha)\hat{z}_n t_n \eta_n}{(1-\alpha)\hat{z}_n t_n + \eta_n}}{\pi_\theta + \sum_{l=1}^N \frac{(1-\alpha)\hat{z}_l t_l \eta_l}{(1-\alpha)\hat{z}_l t_l + \eta_l}}$$

into Condition (2), I then have that

$$\hat{z}_n = \frac{\eta_n}{\sqrt{t_n}(1-\alpha)} \left\{ (1-\alpha)\sqrt{\frac{|u_{kk}|\kappa_1^2}{2C_n'(\hat{z})}} \frac{1}{M_1(\hat{z})} - \frac{1}{\sqrt{t_n}} \right\}, \tag{25}$$

where

$$M_1(z) \equiv \pi_\theta + \sum_{l=1}^N \frac{(1-\alpha)\eta_l z_l t_l}{(1-\alpha) z_l t_l + \eta_l} > 0. \tag{26}$$

For the right-hand-side in (25) to be positive, it must be that

$$\frac{t_n}{C_n'(\hat{z})} > R \equiv \frac{2\left(M_1(\hat{z})\right)^2}{(1-\alpha)^2 \kappa_1^2 |u_{kk}|}, \tag{27}$$

which establishes the first claim in the corollary.

Next, I prove that, for any source that receives no attention in equilibrium, condition (27) must be violated. To see this, suppose that, by contradiction, there exists a source n for which (27) holds and such that $\hat{z}_n = 0$. Suppose that the individual were to increase locally the attention allocated to this source. The continuity of the right-hand derivative of the agent's expected payoff $\partial U_+^j(\hat{z}; \hat{z})/\partial z_n^j$ implies that the net effect on the agent's expected payoff is

$$\frac{|u_{kk}|}{2} \frac{(\kappa_1 \gamma_n(\hat{z}))^2}{(\hat{z}_n)^2 t_n} - C_n'(\hat{z}) = \frac{|u_{kk}| \kappa_1^2}{2} \frac{(1-\alpha)^2 t_n}{(M_1(\hat{z}))^2} - C_n'(\hat{z}) > 0,$$

contradicting the optimality of the equilibrium allocation of attention. Q.E.D.

Proof of Example 1. Suppose that all sources receive strictly positive attention in equilibrium. The amount of attention allocated to each source n is then equal to

$$\hat{z}_n = \sqrt{\frac{|u_{kk}| \kappa_1^2}{2\bar{c}}} \frac{\gamma_n(\hat{z})}{\sqrt{t_n}}. \tag{28}$$

It follows that the influence of each source n is given by

$$\gamma_n(\hat{z}) = \sqrt{\frac{2\bar{c}}{|u_{kk}| \kappa_1^2}} \sqrt{t_n} \hat{z}_n. \tag{29}$$

Combining the above with the fact that

$$\gamma_n(\hat{z}) = \frac{\frac{(1-\alpha)\pi_n(\hat{z})}{1-\alpha\rho_n(\hat{z})}}{\pi_\theta + \sum_{s=1}^N \frac{(1-\alpha)\pi_s(\hat{z})}{1-\alpha\rho_s(\hat{z})}} \tag{30}$$

I then have that

$$\sum_{n=1}^N \gamma_n(\hat{z}) = \frac{\sum_{s=1}^N \frac{(1-\alpha)\pi_s(\hat{z})}{1-\alpha\rho_s(\hat{z})}}{\pi_\theta + \sum_{s=1}^N \frac{(1-\alpha)\pi_s(\hat{z})}{1-\alpha\rho_s(\hat{z})}} = \sqrt{\frac{2\bar{c}}{|u_{kk}| \kappa_1^2}} \sum_{n=1}^N \sqrt{t_n} \hat{z}_n.$$

This implies that

$$\pi_\theta + \sum_{s=1}^N \frac{(1-\alpha)\pi_s(\hat{z})}{1-\alpha\rho_s(\hat{z})} = \frac{\pi_\theta}{1 - \sqrt{\frac{2\bar{c}}{|u_{kk}|\kappa_1^2}} \sum_{s=1}^N \sqrt{t_s} \hat{z}_s}.$$

Replacing the latter expression into the definition of $\gamma_n(\hat{z})$ in (30) and using the fact that

$$\frac{(1-\alpha)\pi_n(\hat{z})}{1-\alpha\rho_n(\hat{z})} = \frac{(1-\alpha)\eta_n \hat{z}_n t_n}{\hat{z}_n t_n (1-\alpha) + \eta_n}$$

I then have that

$$\gamma_n(\hat{z}) = \frac{\frac{(1-\alpha)\eta_n \hat{z}_n t_n}{\hat{z}_n t_n (1-\alpha) + \eta_n}}{\frac{\pi_\theta}{1 - \sqrt{\frac{2\bar{c}}{|u_{kk}|\kappa_1^2}} \sum_{s=1}^N \sqrt{t_s} \hat{z}_s}}.$$

Combining this expression with (28) I then have that

$$\hat{z}_n = \left[\frac{1}{\pi_\theta \sqrt{\frac{2\bar{c}}{|u_{kk}|\kappa_1^2}}} - \frac{1}{\pi_\theta} \sum_{s=1}^{N} \sqrt{t_s} \hat{z}_s \right] \frac{1}{\sqrt{t_n}} \eta_n - \frac{\eta_n}{(1-\alpha)t_n}. \tag{31}$$

Multiplying both sides of (31) by $\sqrt{t_n}$, summing over n, and rearranging, I then obtain that

$$\frac{1}{\pi_\theta} \sum_{s=1}^{N} \sqrt{t_s} \hat{z}_s = \frac{\frac{\sum_{s=1}^{N} \eta_s}{\pi_\theta \sqrt{\frac{2\bar{c}}{|u_{kk}|\kappa_1^2}}} - \frac{1}{(1-\alpha)} \sum_{s=1}^{N} \frac{\eta_s}{\sqrt{t_s}}}{\pi_\theta + \sum_{s=1}^{N} \eta_s}. \tag{32}$$

Replacing (32) into (31), I conclude that

$$\hat{z}_n = \frac{\eta_n}{\sqrt{t_n}(1-\alpha)} \left[\frac{(1-\alpha)\sqrt{\frac{|u_{kk}|\kappa_1^2}{2\bar{c}}} + \sum_{s=1}^{N} \frac{\eta_s}{\sqrt{t_s}}}{\pi_\theta + \sum_{s=1}^{N} \eta_s} - \frac{1}{\sqrt{t_n}} \right]$$

as claimed. Q.E.D.

Proof of Proposition 3. First I prove that, when all agents allocate attention z to the various sources of information, the continuation game that starts when the agents, after observing their summary statistics, must choose their actions, has a unique continuation equilibrium where all agents follow the affine strategy

$$k^i = k^{\#}(X^i; z) \equiv \kappa_0 + \kappa_1 \gamma^{\#}(z) X^i. \tag{33}$$

To see this, observe that, given the attention z, observing the statistics $X^i = \sum_{n=1}^{N} \lambda_n(z) x_n^i$ is informationally equivalent to observing an additive signal

$$\theta + \sum\nolimits_{n=1}^{N} \lambda_n(z)(\varepsilon_n + \xi_n^i)$$

about the exogenous fundamentals θ, with precision $\pi_X(z)$ given by the formula in (13) and with an error whose correlation $\rho_X(z)$ across any pair of agents $i, j \in [0,1]$, $j \neq i$, is given by the formula in (14). This game is isomorphic to the one in Section 2, with the only difference that each agent receives a single signal. From Proposition 1 I then have that, in the unique continuation equilibrium, individual actions are given by (33).

Next, I characterize the allocation of attention in any symmetric equilibrium. To this purpose, suppose that all agents $i \neq j$ allocate attention $z^i = z$ to the different sources of information and then use (33) to determine their actions. Let $U^j(z^j; z)$ denote the payoff of agent j when he allocates attention z^j to the different sources and then chooses optimally the mapping from the statistics X^j into his actions. Using the envelope theorem, in any symmetric equilibrium with attention $z^{\#}$, for any source n for which $z_n^{\#} > 0$, $\partial U^j(z^{\#}; z^{\#})/\partial z_n^j$ must coincide with the partial derivative of the agent's expected payoff with respect to z_n^j, evaluated at $z_n^j = z_n^{\#}$, holding fixed the mapping

$k^{\#}(\cdot; z^{\#})$ from the agent's statistics X^j to his actions and letting this mapping be the one in (33) with $z = z^{\#}$.[23]

Observe that, when, given (z^j, z), all agents (including agent j) follow (33), then

$$U^j(z^j; z) = \mathbb{E}[u(K, K, \sigma_k, \theta) \mid z^j, z] + \mathbb{E}[u_k(K, K, \sigma_k, \theta)(k^j - K) \mid z^j, z] \\ + \frac{u_{kk}}{2} \mathbb{E}[\left(k^j - K\right)^2 \mid z^j, z] - C(z^j),$$

where the first term in the right-hand side of (23) is the expected payoff of an agent whose action coincides with the average action in the population in every state. Importantly, note that (i) because the mapping $k^{\#}(\cdot; z)$ is kept fixed, $\mathbb{E}[u(K, K, \sigma_k, \theta) \mid z^j, z]$ is independent of the agent's own information and (ii) all expectations are computed assuming that all agents' actions are determined by the linear strategy in (33).

Next observe that

$$\mathbb{E}[\left(k^j - K\right)^2 \mid z^j, z] = \mathbb{E}[\left(k^j - K^j\right)^2 + (K^j - K)^2 + 2\left(k^j - K^j\right)\left(K^j - K\right) \mid z^j, z]$$

where $K^j \equiv \mathbb{E}[k^j \mid (\theta, \varepsilon), z^j]$ denotes the agent's own average action given (θ, ε), when his attention is z^j. Using the fact that, for any (z^j, z), $k^j - K^j = \kappa_1 \gamma^{\#}(z) \left[\sum_n \lambda_n(z^j) \xi_n^j\right]$ is orthogonal to $K^j - K = \kappa_1 \gamma^{\#}(z) \left\{\sum_n \left[\lambda_n(z^j) - \lambda_n(z)\right](\theta + \varepsilon_n)\right\}$, I then have that

$$\frac{\partial}{\partial z_n^j} \mathbb{E}[\left(k^j - K\right)^2 \mid z, z] = \frac{\partial}{\partial z_n^j} \mathbb{E}[\left(k^j - K^j\right)^2 \mid z, z] + \frac{\partial}{\partial z_n^j} \mathbb{E}[\left(K^j - K\right)^2 \mid z, z] \\ = \frac{\partial}{\partial z_n^j} \mathbb{E}[\left(k^j - K^j\right)^2 \mid z, z] = \frac{\partial}{\partial z_n} Var\left[k - K \mid z, k^{\#}(\cdot; z)\right],$$

where all derivatives are computed holding fixed the agents' strategies, as given by (33). Note that the second equality follows from the fact that, when $z^j = z$,

$$\frac{\partial}{\partial z_n^j} \mathbb{E}[\left(K^j - K\right)^2 \mid z, z] = 0,$$

whereas the third equality uses the fact that, when $z^j = z$, the dispersion of each agent's action around his own average action coincides with the dispersion of each agent's action around the mean action in the cross-section of the population (in the notation for such a dispersion, I explicitly write the strategy $k^{\#}(\cdot; z)$ to make clear that the distribution of individual and aggregate actions is obtained by assuming the agents follow the mapping in (33)). Importantly, note that the derivative

$$\frac{\partial}{\partial z_n} Var\left[k - K \mid z, k^{\#}(\cdot; z)\right]$$

is computed holding fixed the agents' strategies, but accounting for the fact that a variation in z_n affects the dispersion of individual actions around the mean action both directly by changing the

[23] Furthermore, for any source for which $z_n^{\#} = 0$, the right-hand derivative $\partial U_+^j(z^{\#}; z^{\#})/\partial z_n^j$ must coincide with the limit for $z_n \to 0^+$ of the derivative $\partial U^j((z_n, z_{-n}^{\#}); (z_n, z_{-n}^{\#}))/\partial z_n^j$ by continuity of the right-hand derivative, where $z_{-n}^{\#} \equiv (z_1^{\#}, ..., z_{n-1}^{\#}, z_{n+1}^{\#}, ..., z_N^{\#})$.

distribution of the signal x_j and indirectly by changing the weights $\lambda_s(z)$ in the agents' statistics. The derivations above also establish that, in a symmetric equilibrium, the reduction of the dispersion of individual actions around the mean action is the same irrespective of whether one changes only the individual's own allocation of attention or all agents' allocation of attention, as claimed in the main text.

Finally, consider the term $\mathbb{E}[u_k(K, K, \sigma_k, \theta)(k^j - K) \mid z^j, z]$. Using the fact that

$$u_k(K, K, \sigma_k, \theta) = u_k(\kappa, \kappa, 0, \theta) + (u_{kk} + u_{kK})(K - \kappa),$$

along with the fact that $u_k(\kappa, \kappa, 0, \theta) = 0$ by definition of the complete-information equilibrium, I have that

$$\begin{aligned}\mathbb{E}[u_k(K, K, \sigma_k, \theta)(k^j - K) \mid z^j, z] &= (u_{kk} + u_{kK}) \cdot \mathbb{E}[(K - \kappa)(k^j - K) \mid z^j, z] \\ &= (u_{kk} + u_{kK}) \cdot \mathbb{E}[(K - \kappa)(K^j - K) \mid z^j, z]\end{aligned}$$

where the second equality uses the fact that $k^j - K^j$ is orthogonal to $K - \kappa$. Observe that

$$\begin{aligned}&\frac{\partial}{\partial z_n^j}\mathbb{E}[(K - \kappa)(K^j - K) \mid z, z] = \mathbb{E}\left[(K - \kappa)\frac{\partial (K^j - K)}{\partial z_n^j} \mid z, z\right] \\ &= \kappa_1 \gamma^{\#}(z)\mathbb{E}\left[(K - \kappa)\left(\sum_{s=1}^{N} \frac{\partial \lambda_s(z)}{\partial z_n}(\theta + \varepsilon_s)\right) \mid z, z\right] \\ &= \kappa_1^2 \gamma^{\#}(z) \cdot Cov\left[\left(\gamma^{\#}(z)\sum_{s=1}^{N} \lambda_s(z)(\theta + \varepsilon_s) - \theta\right); \left(\sum_{s=1}^{N} \frac{\partial \lambda_s(z)}{\partial z_n}(\theta + \varepsilon_s)\right) \mid z, z\right] \\ &= \left(\kappa_1 \gamma^{\#}(z)\right)^2 \sum_{s=1}^{N}\left(\lambda_s(z)\frac{\partial \lambda_s(z)}{\partial z_n}\right)\frac{1}{\eta_s},\end{aligned}$$

where, in the last equality, I used the fact that, for any z, $\sum_{s=1}^{N} \lambda_s(z) = 1$ and, hence, $\sum_{s=1}^{N} \frac{\partial \lambda_s(z)}{\partial z_n} = 0$, along with the fact that each ε_s is orthogonal to θ and to any ε_l, $l \neq s$.

Let $\frac{\partial}{\partial z_n} Var\left[K - \kappa \mid z, k^{\#}(\cdot; z)\right]$ denote the marginal change in the dispersion of K around κ that obtains when one changes the attention allocated to the n-th source, holding fixed the strategy in

(33). Then observe that

$$
\begin{aligned}
&\frac{1}{2}\frac{\partial}{\partial z_n}Var\left[K-\kappa \mid z, k^{\#}(\cdot;z)\right] \\
&= \frac{1}{2}\frac{\partial}{\partial z_n}Var\left[\kappa_1\left(\gamma^{\#}(z)\sum_{s=1}^{N}\lambda_s(z)\left(\theta+\varepsilon_s\right)-\theta\right) \;\middle|\; z, k^{\#}(\cdot;z)\right] \\
&= \frac{\kappa_1^2}{2}\frac{\partial}{\partial z_n}Var\left[\left(\gamma^{\#}(z)-1\right)\theta+\gamma^{\#}(z)\sum_{s=1}^{N}\lambda_s(z)\varepsilon_s \;\middle|\; z, k^{\#}(\cdot;z)\right] \\
&= \frac{\left(\kappa_1\gamma^{\#}(z)\right)^2}{2}\frac{\partial}{\partial z_n}Var\left[\sum_{s=1}^{N}\lambda_s(z)\varepsilon_s \;\middle|\; z, k^{\#}(\cdot;z)\right] \\
&= \frac{\left(\kappa_1\gamma^{\#}(z)\right)^2}{2}\frac{\partial}{\partial z_n}\left[\sum_{s=1}^{N}\lambda_s(z)^2\frac{1}{\eta_s}\right] \\
&= \left(\kappa_1\gamma^{\#}(z)\right)^2\left\{\sum_{s=1}^{N}\left(\lambda_s(z)\frac{\partial\lambda_s(z)}{\partial z_n}\right)\frac{1}{\eta_s}\right\} \\
&= \frac{\partial}{\partial z_n^j}\mathbb{E}[(K-\kappa)(K^j-K) \mid z, z].
\end{aligned}
\tag{34}
$$

Combining the different pieces and using the fact that $|u_{kk}|(1-\alpha) = -(u_{kk}+u_{kK})$, I conclude that

$$
\begin{aligned}
\frac{\partial U^j(z;z)}{\partial z_n^j} = &-\frac{|u_{kk}|}{2}\frac{\partial}{\partial z_n}Var\left[k-K \mid z, k^{\#}(\cdot;z)\right] \\
&-\frac{|u_{kk}|}{2}(1-\alpha)\frac{\partial}{\partial z_n}Var\left[K-\kappa \mid z, k^{\#}(\cdot;z)\right] - C_n'(z).
\end{aligned}
\tag{35}
$$

In any symmetric equilibrium, for any source of information $n = 1, ..., N$ that receives strictly positive attention, it must be that the above derivative vanishes, which yields (18) in the main text. Clearly, the second term in (18) disappears if the agent can choose his summary statistics X, for, in this case, by usual envelope arguments, the marginal benefit of increasing the attention to any source should be computed holding fixed both the selection of the summary statistics and the strategy that maps the latter into the selected action. When both the weights λ defining the statistics X and the mapping $k^{\#}(\cdot;z)$ are held fixed, the average action is thus invariant in z_n.

Finally, note that the uniqueness of the symmetric equilibrium follows from arguments similar to those that establish uniqueness in the model with perfect recall; the proof is thus omitted for brevity. Q.E.D.

Proof of Proposition 4. The proof is in four steps. Step 1 shows how, starting from a situation where all agents assign attention z to the various sources, the net benefit of increasing the attention allocated to any source n depends on the primitive parameters of the model. The characterization applies to any statistics X^j summarizing the agents' information. Step 2 then uses the characterization in Step 1 to compare the benefit in the benchmark with perfect recall (equivalently, when the agents can choose their summary statistics) and in the case of bounded recall when the statistics is exogenous. Step 3 shows how the comparison in Step 2 specializes in the case where the exogenous statistics X^j takes the form in (15). Finally, step 4 shows that $0 < \rho' < \rho'' < 1$ for α large enough and β small enough.

Step 1. Under bounded recall, the marginal benefit of increasing the attention to any source n is given by (35). Below, I express the various terms in (35) as a function of the key parameters. To simplify the exposition, I drop z from the arguments of the various functions, when there is no risk of confusion.

First observe that

$$\frac{\partial}{\partial z_n} Var\left[k - K \mid z, k^{\#}(\cdot)\right] = \left(\kappa_1 \gamma^{\#}\right)^2 \frac{\partial}{\partial z_n} Var\left(\sum_{s=1}^{N} \lambda_s \xi_s\right).$$

Using the fact that

$$\frac{\partial}{\partial z_n} Var\left(\sum_{s=1}^{N} \lambda_s \xi_s\right) = \frac{\partial}{\partial z_n}\left[\sum_{s=1}^{N} \frac{\lambda_s^2}{t_s z_s}\right] = \sum_{s=1}^{N} \frac{2\lambda_s}{t_s z_s} \frac{\partial \lambda_s}{\partial z_n} - \frac{\lambda_n^2}{t_n \left(z_n\right)^2},$$

I have that

$$\frac{\partial}{\partial z_n} Var\left[k - K \mid z, k^{\#}(\cdot)\right] = \left(\kappa_1 \gamma^{\#}\right)^2 \left\{\sum_{s=1}^{N} \frac{2\lambda_s}{t_s z_s} \frac{\partial \lambda_s}{\partial z_n} - \frac{\lambda_n^2}{t_n \left(z_n\right)^2}\right\}. \tag{36}$$

Next, use (34) to observe that

$$\frac{\partial}{\partial z_n} var\left[K - \kappa \mid z, k^{\#}(\cdot)\right] = \left(\kappa_1 \gamma^{\#}\right)^2 \left\{\sum_{s=1}^{N} \frac{2\lambda_s}{\eta_s} \frac{\partial \lambda_s}{\partial z_n}\right\}. \tag{37}$$

Substituting (36) and (37) into (35), I conclude that, for any source n and any allocation of attention z,

$$\begin{aligned} \frac{\partial U^j(z; z)}{\partial z_n} = & -|u_{kk}| \left(\kappa_1 \gamma^{\#}\right)^2 \left\{\sum_{s=1}^{N} \frac{\lambda_s}{t_s z_s} \frac{\partial \lambda_s}{\partial z_n} - \frac{1}{2} \frac{\lambda_n^2}{t_n \left(z_n\right)^2}\right\} \\ & - |u_{kk}|(1-\alpha) \left(\kappa_1 \gamma^{\#}\right)^2 \left\{\sum_{s=1}^{N} \frac{\lambda_s}{\eta_s} \frac{\partial \lambda_s}{\partial z_n}\right\} - C_n'(z). \end{aligned} \tag{38}$$

Next, observe that

$$t_s z_s = \frac{\pi_s}{1 - \rho_s} \quad \text{and} \quad \eta_s = \frac{\pi_s}{\rho_s}.$$

The benefit of increasing (locally) the attention to source n is thus equal to

$$\frac{\partial U^j(z; z)}{\partial z_n} = |u_{kk}| \left(\kappa_1 \gamma^{\#}\right)^2 \left\{\frac{\lambda_n^2}{2 t_n \left(z_n\right)^2} - \sum_{s=1}^{N} \frac{(1 - \alpha \rho_s) \lambda_s}{\pi_s} \frac{\partial \lambda_s}{\partial z_n}\right\} - C_n'(z). \tag{39}$$

Step 2. I now turn to the comparison between the benefit of increasing the attention to source n in the benchmark of perfect recall and in the case of bounded recall.

In the case of perfect recall (equivalently, when the sufficient statistics is chosen optimally by the agents), the value of increasing the attention to any source n is equal to (see the end of the proof of Proposition 1)

$$\frac{\partial U^j(z;z)}{\partial z_n} = |u_{kk}|\left(\kappa_1\right)^2 \left(\frac{\frac{(1-\alpha)\pi_n}{1-\alpha\rho_n}}{\pi_\theta + \sum_{s=1}^N \frac{(1-\alpha)\pi_s}{1-\alpha\rho_s}}\right)^2 \frac{1}{2t_n\left(z_n\right)^2} - C_n'(z). \tag{40}$$

Comparing (39) with (40), it is then easy to see that the benefit is larger under bounded recall if

$$\frac{\left(\lambda_n\right)^2\left(\gamma^{\#}\right)^2}{2t_n\left(z_n\right)^2} - \left(\gamma^{\#}\right)^2 \sum_{s=1}^N \frac{(1-\alpha\rho_s)\lambda_s}{\pi_s}\frac{\partial\lambda_s}{\partial z_n} > \left(\frac{\frac{(1-\alpha)\pi_n}{1-\alpha\rho_n}}{\pi_\theta + \sum_{s=1}^N \frac{(1-\alpha)\pi_s}{1-\alpha\rho_s}}\right)^2 \frac{1}{2t_n\left(z_n\right)^2}$$

and lower if the inequality is reversed. The above inequality can be rearranged as

$$\begin{aligned} &\tfrac{(\lambda_n)^2(1-\alpha\rho_n)^2}{(\pi_n)^2} - \tfrac{2t_n(z_n)^2(1-\alpha\rho_n)^2}{(\pi_n)^2}\textstyle\sum_{s=1}^N \tfrac{(1-\alpha\rho_s)\lambda_s}{\pi_s}\tfrac{\partial\lambda_s}{\partial z_n} \\ &> \tfrac{(1-\alpha)^2}{\left(\pi_\theta+\sum_{s=1}^N \frac{(1-\alpha)\pi_s}{1-\alpha\rho_s}\right)^2\left(\gamma^{\#}\right)^2}. \end{aligned} \tag{41}$$

Step 3. When the weights are as in (15), the inequality in (41) reduces to

$$\begin{aligned} &\left(\tfrac{1-\alpha\rho_n}{1-\beta\alpha\rho_n}\right)^2 - \tfrac{2t_n(z_n)^2(1-\alpha\rho_n)^2}{(\pi_n)^2}\left(\textstyle\sum_{l=1}^N \tfrac{\pi_l}{1-\beta\alpha\rho_l}\right)\textstyle\sum_{s=1}^N \tfrac{1-\alpha\rho_s}{1-\beta\alpha\rho_s}\tfrac{\partial\lambda_s}{\partial z_n} \\ &> \tfrac{(1-\alpha)^2\left(\sum_{l=1}^N \frac{\pi_l}{1-\beta\alpha\rho_l}\right)^2}{\left(\pi_\theta+\sum_{s=1}^N \frac{(1-\alpha)\pi_s}{1-\alpha\rho_s}\right)^2\left(\gamma^{\#}\right)^2}. \end{aligned} \tag{42}$$

Furthermore,

$$\frac{\partial\lambda_s}{\partial z_n} = \left[\frac{\partial\lambda_s}{\partial\pi_n} + \frac{\partial\lambda_s}{\partial\rho_n}\frac{1}{\eta_n}\right]\frac{\partial\pi_n}{\partial z_n},$$

with

$$\frac{\partial\pi_n}{\partial z_n} = \frac{\pi_n^2}{t_n\left(z_n\right)^2},$$

$$\frac{\partial\lambda_s}{\partial\pi_n} = \begin{cases} -\frac{\frac{\pi_s}{1-\beta\alpha\rho_s}\frac{1}{1-\beta\alpha\rho_n}}{\left[\sum_{l=1}^N \frac{\pi_l}{1-\beta\alpha\rho_l}\right]^2} & \text{if } s \neq n, \\ \frac{\frac{1}{1-\beta\alpha\rho_n}}{\sum_{l=1}^N \frac{\pi_l}{1-\beta\alpha\rho_l}} - \frac{\frac{\pi_n}{(1-\beta\alpha\rho_n)^2}}{\left[\sum_{l=1}^N \frac{\pi_l}{1-\beta\alpha\rho_l}\right]^2} & \text{if } s = n, \end{cases}$$

and

$$\frac{\partial\lambda_s}{\partial\rho_n} = \begin{cases} -\frac{\frac{\pi_s}{1-\beta\alpha\rho_s}\frac{\pi_n\alpha\beta}{(1-\beta\alpha\rho_n)^2}}{\left[\sum_{l=1}^N \frac{\pi_l}{1-\beta\alpha\rho_l}\right]^2} & \text{if } s \neq n, \\ \frac{\frac{\pi_n\beta\alpha}{(1-\beta\alpha\rho_n)^2}}{\sum_{l=1}^N \frac{\pi_l}{1-\beta\alpha\rho_l}} - \frac{\frac{\pi_s}{1-\beta\alpha\rho_s}\frac{\pi_n\alpha\beta}{(1-\beta\alpha\rho_n)^2}}{\left[\sum_{l=1}^N \frac{\pi_l}{1-\beta\alpha\rho_l}\right]^2} & \text{if } s = n. \end{cases}$$

Using the above expressions along with $\frac{1}{\eta_n} = \frac{\rho_n}{\pi_n}$, I can rewrite the inequality in (42) as

$$\begin{aligned} &\left(\tfrac{1-\alpha\rho_n}{1-\beta\alpha\rho_n}\right)^2 - 2\left(1-\alpha\rho_n\right)^2\left(\textstyle\sum_{l=1}^N \tfrac{\pi_l}{1-\beta\alpha\rho_l}\right)\textstyle\sum_{s=1}^N\left\{\tfrac{1-\alpha\rho_s}{1-\beta\alpha\rho_s}\left[\tfrac{\partial\lambda_s}{\partial\pi_n}+\tfrac{\partial\lambda_s}{\partial\rho_n}\tfrac{\rho_n}{\pi_n}\right]\right\} \\ &\qquad > \frac{(1-\alpha)^2\left(\sum_{l=1}^N \frac{\pi_l}{1-\beta\alpha\rho_l}\right)^2}{\left(\pi_\theta+\sum_{s=1}^N\frac{(1-\alpha)\pi_s}{1-\alpha\rho_s}\right)^2\left(\gamma^\#\right)^2}, \end{aligned} \tag{43}$$

with

$$\sum_{s=1}^N\left\{\frac{1-\alpha\rho_s}{1-\beta\alpha\rho_s}\left[\frac{\partial\lambda_s}{\partial\pi_n}+\frac{\partial\lambda_s}{\partial\rho_n}\frac{\rho_n}{\pi_n}\right]\right\} = \frac{\frac{1-\alpha\rho_n}{(1-\beta\alpha\rho_n)^3}}{\sum_{l=1}^N\frac{\pi_l}{1-\beta\alpha\rho_l}} - \frac{1}{(1-\beta\alpha\rho_n)^2}\frac{\sum_{s=1}^N\frac{\pi_s(1-\alpha\rho_s)}{(1-\beta\alpha\rho_s)^2}}{\left(\sum_{l=1}^N\frac{\pi_l}{1-\beta\alpha\rho_l}\right)^2}.$$

After some simplifications, the inequality in (43) thus reduces to

$$2\left(\tfrac{1-\alpha\rho_n}{1-\beta\alpha\rho_n}\right) + \frac{(1-\alpha)^2\left(\sum_{l=1}^N\frac{\pi_l}{1-\beta\alpha\rho_l}\right)^2}{\left(\pi_\theta+\sum_{s=1}^N\frac{(1-\alpha)\pi_s}{1-\alpha\rho_s}\right)^2\left(\gamma^\#\right)^2}\frac{1}{\left(\frac{1-\alpha\rho_n}{1-\beta\alpha\rho_n}\right)^2} < 1 + 2\frac{\sum_{s=1}^N\frac{\pi_s(1-\alpha\rho_s)}{(1-\beta\alpha\rho_s)^2}}{\sum_{l=1}^N\frac{\pi_l}{1-\beta\alpha\rho_l}}. \tag{44}$$

Now observe that the left-hand side of (44) is strictly convex in $\left(\frac{1-\alpha\rho_n}{1-\beta\alpha\rho_n}\right)$. Furthermore, $\left(\frac{1-\alpha\rho_n}{1-\beta\alpha\rho_n}\right)$ is decreasing in ρ_n. This means that there exist $\rho', \rho'' \in [0,1]$ with $0 \le \rho' \le \rho'' \le 1$ such that the inequality in (44) holds if and only if $\rho_n \in [\rho', \rho'']$ whereas the opposite inequality holds if and only if $\rho_n \notin [\rho', \rho'']$.

Step 4. Observe that, when $\beta = 0$, i.e., when the statistics X coincide with the projection of θ on x, the inequality in (44) reduces to

$$2\alpha(\rho_n - \rho_X^0) > \left(\frac{\pi_\theta + \frac{(1-\alpha)\pi_X^0}{1-\alpha\rho_X^0}}{\pi_\theta + \sum_{s=1}^N\frac{(1-\alpha)\pi_s}{1-\alpha\rho_s}}\right)^2\frac{\left(1-\alpha\rho_X^0\right)^2}{\left(1-\alpha\rho_n\right)^2} - 1, \tag{45}$$

with $\pi_X^0 \equiv \sum_{s=1}^N \pi_s$ and $\rho_X^0 \equiv \sum_{s=1}^N \frac{\pi_s}{\pi_X^0}\rho_s$. Then note that the function defined by

$$f(\rho) \equiv \frac{1}{1-\alpha\rho}$$

is convex. Hence, by Jensen inequality,

$$\sum_{s=1}^N \frac{\pi_s}{\pi_X^0}\frac{1}{1-\alpha\rho_s} > \frac{1}{1-\alpha\rho_X^0},$$

which implies that

$$\left(\frac{\pi_\theta + \frac{(1-\alpha)\pi_X^0}{1-\alpha\rho_X^0}}{\pi_\theta + \sum_{s=1}^N\frac{(1-\alpha)\pi_s}{1-\alpha\rho_s}}\right)^2 < 1.$$

If follows that the inequality in (45) always holds for sources for which $|\rho_n - \rho_X^0|$ is small, implying that $\rho' < \rho_X^0 < \rho''$. Also note that, when $\alpha = 1$, the inequality in (45) is reversed for ρ_n close to 1. By continuity, $\rho'' < 1$ for α large enough. Likewise, one can verify that, for α large enough, the inequality in (45) can be reversed when evaluated at ρ_n close to zero. This means that $0 < \rho' < \rho'' < 1$ for α large enough and β small enough. Q.E.D.

Proof of Corollary 3. The result follows from Proposition 4 along with the fact that, when the attention cost depends only on total attention, then, under perfect recall, there is an increasing relationship between the exogenous transparency of the sources and their endogenous publicity. Namely, if $t_n > t_{n'}$, then $\rho_n(\hat{z}) \geq \rho_{n'}(\hat{z})$. To see this, use the results in the proof of Corollary 1 to see that

$$\hat{z}_n = \frac{\eta_n}{\sqrt{t_n}(1-\alpha)} \max\left\{\hat{T} - \frac{1}{\sqrt{t_n}}; 0\right\} \tag{46}$$

where

$$\hat{T} \equiv (1-\alpha)\sqrt{\frac{|u_{kk}|\,\kappa_1^2}{2c'(\hat{Z})}\frac{1}{\hat{M}_1}}, \tag{47}$$

with

$$\hat{Z} \equiv \sum_{l=1}^{N} \hat{z}_l, \tag{48}$$

and

$$\hat{M}_1 \equiv \pi_\theta + \sum_{l=1}^{N} \frac{(1-\alpha)\eta_l \hat{z}_l t_l}{(1-\alpha)\hat{z}_l t_l + \eta_l}. \tag{49}$$

Clearly, for all sources for which $t_n \leq 1/\hat{T}^2$, $\hat{z}_n = 0$ and hence $\rho_n(\hat{z}) = 0$. On the other hand, for all sources for which $t_n > 1/\hat{T}^2$,

$$\rho_n(\hat{z}) = \frac{\hat{z}_n t_s}{\hat{z}_n t_n + \eta_n} = \frac{\sqrt{t_n}\hat{T} - 1}{\sqrt{t_n}\hat{T} - \alpha}$$

which is increasing in t_n. Q.E.D.

Proof of Proposition 5. Use (39) to observe that, in any equilibrium with *bounded* recall, for any source n for which $z_n^\# > 0$, the following condition must hold

$$|u_{kk}|\left(\kappa_1\gamma^\#\right)^2\left\{\frac{\left(\lambda_n^\#\right)^2}{2t_n\left(z_n^\#\right)^2} - \sum_{s=1}^{N}\frac{(1-\alpha\rho_s^\#)\lambda_s^\#}{\pi_s^\#}\frac{\partial\lambda_s^\#}{\partial z_n}\right\} = C_n'(z^\#), \tag{50}$$

where $\gamma^\#$, $\pi_s^\#$, $\rho_s^\#$, $\lambda_s^\#$ are shortcuts for $\gamma^\#(z^\#)$, $\pi_s(z^\#)$, $\rho_s(z^\#)$, and $\lambda_s(z^\#)$, respectively.

When the weights in the statistics X^j are those in (15) and the cost of attention is given by $C(z) = c(\sum_{l=1}^N z_l)$, the above condition reduces to

$$|u_{kk}|\left(\kappa_1\gamma^\#\right)^2\left\{\frac{\left(\frac{\pi_n^\#}{1-\alpha\beta\rho_n^\#}\right)^2}{\left(\sum_{l=1}^N\frac{\pi_l^\#}{1-\alpha\beta\rho_l^\#}\right)^2 2t_n\left(z_n^\#\right)^2} - \frac{1}{\left(\sum_{l=1}^N\frac{\pi_l^\#}{1-\alpha\beta\rho_l^\#}\right)}\left[\sum_{s=1}^N\frac{1-\alpha\rho_s^\#}{1-\alpha\beta\rho_s^\#}\frac{\partial\lambda_s^\#}{\partial z_n}\right]\right\} = c'(Z^\#), \tag{51}$$

where $Z^{\#} \equiv \sum_{l=1}^{N} z_l^{\#}$. Furthermore, as shown in the proof of Proposition 4, when the weights in X^j are those in (15),

$$\frac{\partial \lambda_s}{\partial z_n} = \left[\frac{\partial \lambda_s}{\partial \pi_n} + \frac{\partial \lambda_s}{\partial \rho_n}\frac{1}{\eta_n}\right]\frac{\partial \pi_n}{\partial z_n} = \left[\frac{\partial \lambda_s}{\partial \pi_n} + \frac{\partial \lambda_s}{\partial \rho_n}\frac{\rho_n}{\pi_n}\right]\frac{\pi_n^2}{t_n\left(z_n\right)^2}$$

and

$$\sum_{s=1}^{N}\left\{\frac{1-\alpha\rho_s}{1-\beta\alpha\rho_s}\left[\frac{\partial \lambda_s}{\partial \pi_n} + \frac{\partial \lambda_s}{\partial \rho_n}\frac{\rho_n}{\pi_n}\right]\right\} = \frac{\frac{1-\alpha\rho_n}{(1-\beta\alpha\rho_n)^3}}{\sum_{l=1}^{N}\frac{\pi_l}{1-\beta\alpha\rho_l}} - \frac{1}{(1-\beta\alpha\rho_n)^2}\frac{\sum_{s=1}^{N}\frac{\pi_s(1-\alpha\rho_s)}{(1-\beta\alpha\rho_s)^2}}{\left(\sum_{l=1}^{N}\frac{\pi_l}{1-\beta\alpha\rho_l}\right)^2}.$$

It follows that Condition (51) can be rewritten as

$$\frac{1}{2} - \frac{1-\alpha\rho_n^{\#}}{1-\beta\alpha\rho_n^{\#}} + \frac{\sum_{s=1}^{N}\frac{\pi_s^{\#}(1-\alpha\rho_s^{\#})}{\left(1-\beta\alpha\rho_s^{\#}\right)^2}}{\sum_{l=1}^{N}\frac{\pi_l^{\#}}{1-\beta\alpha\rho_l^{\#}}} = \frac{c'(Z^{\#})\left(\sum_{l=1}^{N}\frac{\pi_l^{\#}}{1-\alpha\beta\rho_l^{\#}}\right)^2 t_n\left(z_n^{\#}\right)^2(1-\beta\alpha\rho_n^{\#})^2}{|u_{kk}|\left(\kappa_1\gamma^{\#}\right)^2\left(\pi_n^{\#}\right)^2}.$$

In the special case in which $\beta = 0$, the above condition reduces to

$$\frac{1}{2} + \alpha\left(\rho_n^{\#} - \rho_X^{\#}\right) = \frac{c'(Z^{\#})\left(\pi_X^{\#}\right)^2 t_n\left(z_n^{\#}\right)^2}{|u_{kk}|\left(\kappa_1\gamma^{\#}\right)^2\left(\pi_n^{\#}\right)^2}, \tag{52}$$

from which I obtain that

$$z_n^{\#} = \sqrt{\frac{|u_{kk}|\left(\kappa_1\gamma^{\#}\right)^2}{c'(Z^{\#})\left(\pi_X^{\#}\right)^2}}\frac{\left(\pi_n^{\#}\right)}{\sqrt{t_n}}\sqrt{\frac{1}{2} + \alpha\left(\rho_n^{\#} - \rho_X^{\#}\right)}.$$

Using

$$\rho_n^{\#} = \frac{z_n^{\#}t_n}{z_n^{\#}t_n + \eta_n} \text{ and } \pi_n^{\#} \equiv \frac{\eta_n z_n^{\#}t_n}{z_n^{\#}t_n + \eta_n},$$

I conclude that, for any source n that receives strictly positive attention in the equilibrium with bounded recall,

$$z_n^{\#} = \frac{\eta_n}{\sqrt{t_n}}\left\{A^{\#}\sqrt{B^{\#} + \alpha\rho_n^{\#}} - \frac{1}{\sqrt{t_n}}\right\},$$

where

$$A^{\#} \equiv \sqrt{\frac{|u_{kk}|\left(\kappa_1\gamma^{\#}\right)^2}{c'(Z^{\#})\left(\pi_X^{\#}\right)^2}} \text{ and } B^{\#} \equiv \frac{1}{2} - \alpha\rho_X^{\#}.$$

Furthermore, for any source that receives no attention in the equilibrium with bounded recall, the following condition must hold

$$u_{kk}|\left(\kappa_1\gamma^{\#}\right)^2\left\{\frac{\left(\lambda_n^{\#}\right)^2}{2t_n\left(z_n^{\#}\right)^2} - \sum_{s=1}^{N}\frac{(1-\alpha\rho_s^{\#})\lambda_s^{\#}}{\pi_s^{\#}}\frac{\partial \lambda_s^{\#}}{\partial z_n}\right\} \leq C_n'(z^{\#}),$$

which, using the derivations above, reduces to

$$c'(Z^{\#}) \geq \frac{|u_{kk}| \left(\kappa_1 \gamma^{\#}\right)^2}{\left(\pi_X^{\#}\right)^2} t_n B^{\#}.$$

Next, use (46) to observe that, when $C(z) = c(\sum_{l=1}^N z_l)$, in the equilibrium with *perfect* recall, for any source n,

$$\hat{z}_n = \frac{\eta_n}{\sqrt{t_n}(1-\alpha)} \max\left\{\hat{T} - \frac{1}{\sqrt{t_n}}; 0\right\}$$

where

$$\hat{T} \equiv (1-\alpha)\sqrt{\frac{|u_{kk}| \left(\kappa_1\right)^2}{2c'(\hat{Z})} \frac{1}{M_1(\hat{z})}}$$

with $\hat{Z} \equiv \sum_{l=1}^N \hat{z}_l$ and

$$M_1(\hat{z}) \equiv \pi_\theta + \sum_{l=1}^N \frac{(1-\alpha)\eta_l \hat{z}_l t_l}{(1-\alpha)\hat{z}_l t_l + \eta_l}.$$

From the above observations, I conclude that

$$z_n^{\#} > \hat{z}_n \;\Rightarrow\; A^{\#}\sqrt{B^{\#} + \alpha\rho_n^{\#}} + \frac{\alpha}{1-\alpha}\frac{1}{\sqrt{t_n}} > \frac{\hat{T}}{1-\alpha} \text{ and} \tag{53}$$
$$0 < z_n^{\#} < \hat{z}_n \;\Rightarrow\; A^{\#}\sqrt{B^{\#} + \alpha\rho_n^{\#}} + \frac{\alpha}{1-\alpha}\frac{1}{\sqrt{t_n}} < \frac{\hat{T}}{1-\alpha}.$$

Finally, use the definition of the publicity of a source to observe that, for any source n for which $z_n^{\#} > 0$,

$$\rho_n^{\#} = 1 - \frac{1}{A^{\#}\sqrt{t_n}\sqrt{B^{\#} + \alpha\rho_n^{\#}}}. \tag{54}$$

That is, the publicity $\rho_n^{\#}$ of any source that receives some attention in the equilibrium with bounded recall must solve the following equation:

$$\left[1 - \rho_n^{\#}\right]\sqrt{B^{\#} + \alpha\rho_n^{\#}} = \frac{1}{A^{\#}\sqrt{t_n}}. \tag{55}$$

Next observe that the left-hand-side of Condition (55) is decreasing in ρ when $\alpha \leq 0$. In this case, Condition (55) implicitly defines an increasing function $\rho^{\#}(t)$ between the transparency t and the publicity $\rho^{\#}$ of those sources that receive strictly positive attention in the equilibrium with bounded recall. The same is true when $\alpha > 0$. To see this, fix $B^{\#}$, let

$$\underline{\rho}^{\#} \equiv \begin{cases} 0 \text{ if } B^{\#} \geq 0 \\ -\frac{B^{\#}}{\alpha} \text{ if } B^{\#} < 0 \end{cases}$$

and note that the function $h(\rho) \equiv (1-\rho)\sqrt{B^{\#} + \alpha\rho}$ defined by the left-hand-side of Condition (55) (a) is defined over $[\underline{\rho}^{\#}, 1]$, (b) is non-negative, (c) satisfies $h(1) = 0$ and $h(\underline{\rho}^{\#}) = 0$ when $\underline{\rho}^{\#} > 0$

and $h(\underline{\rho}^{\#}) > 0$ when $\underline{\rho}^{\#} = 0$, and (d) is concave. The above properties imply that $h(\cdot)$ is either decreasing over $[\underline{\rho}^{\#}, 1]$, or it inverted U-shaped with a stationary point $\rho^s \in (\underline{\rho}^{\#}, 1)$. In this case, Condition (55) may admit two solutions. However, when this is the case, it is always the largest one that identifies the equilibrium publicity of the source. To see this, observe that, when h is decreasing over $[\underline{\rho}^{\#}, 1]$, the unique solution to the equation defined by Condition (55) is such that $h(\rho) > \frac{1}{A^{\#}\sqrt{t_n}}$ for $\rho < \rho_n^{\#}$ and $h(\rho) < \frac{1}{A^{\#}\sqrt{t_n}}$ for $\rho > \rho_n^{\#}$. This means that, for the agent's payoff to reach at a global maximum at $z_n = z_n^{\#}$ or, equivalently, for $\rho_n^{\#}$ to be the equilibrium publicity of the source, it must be that h is locally decreasing at the equilibrium level $\rho_n^{\#}$.

I conclude that, irrespective of the sign of α, Condition (55) identifies an increasing relationship $\rho^{\#}(\cdot)$ between the endogenous publicity $\rho_n^{\#}$ and the transparency t_n of the sources that receive strictly positive attention in the equilibrium with bounded recall, with the relationship $\rho^{\#}(\cdot)$ given by the highest solution to the equation in Condition (55).

Now let $[\underline{t}, +\infty)$ denote the set of transparency levels for which the equation in Condition (55) admits a solution. Then observe that, over $[\underline{t}, +\infty)$, the highest solution to the equation in Condition (55) identifies a differentiable function with

$$\frac{\partial \rho^{\#}(t)}{\partial t} = \frac{\left[B^{\#} + \alpha \rho^{\#}(t)\right]\left[1 - \rho^{\#}(t)\right]}{t\left\{2\left[B^{\#} + \alpha \rho^{\#}(t)\right] - \left[1 - \rho^{\#}(t)\right]\alpha\right\}}.$$

Then, for ant $t \in [\underline{t}, +\infty)$, let $\Lambda(t)$ be the function defined by

$$\Lambda(t) = A^{\#}\sqrt{B^{\#} + \alpha \rho^{\#}(t)} + \frac{\alpha}{1-\alpha}\frac{1}{\sqrt{t}} = \frac{1}{\left[1 - \rho^{\#}(t)\right]\sqrt{t}} + \frac{\alpha}{1-\alpha}\frac{1}{\sqrt{t}} \tag{56}$$

with $\rho^{\#}(t)$ denoting the increasing function implicitly defined by the highest solution to the equation in Condition (55). The function $\Lambda(\cdot)$ is differentiable over $[\underline{t}, +\infty)$ with

$$\begin{aligned} \Lambda'(t) &= \frac{1}{\left[1 - \rho^{\#}(t)\right]^2 t}\left\{\frac{\partial \rho^{\#}(t)}{\partial t}\sqrt{t} - \frac{1 - \rho^{\#}(t)}{2\sqrt{t}}\right\} - \frac{\alpha}{(1-\alpha)2t\sqrt{t}} \\ &= \frac{\alpha}{2t\sqrt{t}}\left\{\frac{1}{2B^{\#} - \alpha + 3\alpha\rho^{\#}(t)} - \frac{1}{1-\alpha}\right\}. \end{aligned}$$

Note that, irrespective of the sign of α, because $\rho^{\#}(t)$ is non-decreasing, $\Lambda(t)$ is quasi-concave, meaning that either $\Lambda'(t)$ is of constant sign, or there exists $t^{\#}$ such that $\Lambda'(t) > 0$ for $t < t^{\#}$ and $\Lambda'(t) < 0$ if $t > t^{\#}$. The quasi-concavity of $\Lambda(t)$ is clearly preserved when the function $\Lambda(t)$ is restricted to the set $\left\{t_n : \; n = 1, ..., N \text{ and } z_n^{\#} > 0\right\}$. Because $\Lambda(t)$ coincides with the left-hand side of the inequalities in (53) that are responsible for whether $z_n^{\#} > \hat{z}_n$ or $0 < z_n^{\#} < \hat{z}_n$, I then conclude that, among those sources that receive attention under bounded recall, one of the following must be true: (a) $z_n^{\#} > \hat{z}_n$ for all n; (b) $z_n^{\#} < \hat{z}_n$ for all n; (c) there exists t_1 such that $z_n^{\#} > \hat{z}_n$ for those n for which $t_n < t_1$ and $z_n^{\#} < \hat{z}_n$ for those n for which $t_n > t_1$; (d) there exists t_2 such that $z_n^{\#} > \hat{z}_n$ for those n for which $t_n > t_2$ and $z_n^{\#} < \hat{z}_n$ for those n for which $t_n < t_2$; (e) there exist thresholds t_1 and t_2 such that $z_n^{\#} > \hat{z}_n$ for those n for which $t_n \in (t_1, t_2)$ and $z_n^{\#} < \hat{z}_n$ for those n for which $t_n \notin (t_1, t_2)$.

All these various cases can be summarized concisely by saying that there exist thresholds t' and t'' such that the properties in the proposition hold. Q.E.D.

Proof of Proposition 6. Fix the equilibrium allocation of attention $z^{\#}$. Using (18) and (22), I have that, starting from $z^{\#}$, the private and the social marginal benefits of increasing the attention to source n are given by, respectively,

$$PB_n(z^{\#}) = -\frac{|u_{kk}|}{2}\frac{\partial}{\partial z_n}Var\left[k-K \mid z^{\#}, k^{\#}(\cdot;z^{\#})\right] - \frac{|u_{kk}|}{2}(1-\alpha)\frac{\partial}{\partial z_n}Var\left[K-\kappa \mid z^{\#}, k^{\#}(\cdot;z^{\#})\right]$$

and

$$\begin{aligned} SB_n(z^{\#}) = &-\frac{|u_{kk}+u_{\sigma\sigma}|}{2}\frac{\partial}{\partial z_n}Var\left[k-K \mid z^{\#}, k^{**}(\cdot;z^{\#})\right] \\ &-\frac{|u_{kk}+u_{\sigma\sigma}|}{2}(1-\alpha^*)\frac{\partial}{\partial z_n}Var\left[K-\kappa \mid z^{\#}, k^{**}(\cdot;z^{\#})\right] \end{aligned}$$

where $k^{\#}(\cdot;z^{\#})$ and $k^{**}(\cdot;z^{\#})$ are, respectively, the equilibrium and the efficient strategy, with bounded recall.

Next, use the results in the Proof of Proposition 4, along with the fact that $t_s z_s = \pi_s/(1-\rho_s)$ and $\eta_s = \pi_s/\rho_s$, to observe that

$$\frac{\partial}{\partial z_n}Var\left[k-K \mid z^{\#}, k^{\#}(\cdot;z^{\#})\right] = \left(\kappa_1\gamma^{\#}(z^{\#})\right)^2\left[\sum_{s=1}^{N}\frac{2\left(1-\rho_s(z^{\#})\right)\lambda_s(z^{\#})}{\pi_s(z^{\#})}\frac{\partial\lambda_s(z^{\#})}{\partial z_n} - \frac{\left(\lambda_n(z^{\#})\right)^2}{t_n\left(z_n^{\#}\right)^2}\right]$$

and

$$\frac{\partial}{\partial z_n}var\left[K-\kappa \mid z^{\#}, k^{\#}(\cdot;z^{\#})\right] = \left(\kappa_1\gamma^{\#}(z^{\#})\right)^2\sum_{s=1}^{N}\frac{2\rho_s(z^{\#})\lambda_s(z^{\#})}{\pi_s(z^{\#})}\frac{\partial\lambda_s(z^{\#})}{\partial z_n}.$$

Similarly,

$$\frac{\partial}{\partial z_n}Var\left[k-K \mid z^{\#}, k^{**}(\cdot;z^{\#})\right] = \left(\kappa_1^*\gamma^{**}(z^{\#})\right)^2\left[\sum_{s=1}^{N}\frac{2\left(1-\rho_s(z^{\#})\right)\lambda_s(z^{\#})}{\pi_s(z^{\#})}\frac{\partial\lambda_s(z^{\#})}{\partial z_n} - \frac{\left(\lambda_n(z^{\#})\right)^2}{t_n\left(z_n^{\#}\right)^2}\right]$$

and

$$\frac{\partial}{\partial z_n}var\left[K-\kappa \mid z^{\#}, k^{**}(\cdot;z^{\#})\right] = \left(\kappa_1^*\gamma^{**}(z^{\#})\right)^2\sum_{s=1}^{N}\frac{2\rho_s(z^{\#})\lambda_s(z^{\#})}{\pi_s(z^{\#})}\frac{\partial\lambda_s(z^{\#})}{\partial z_n}.$$

Note that the last two expressions are obtained by recognizing that the efficient strategy $k^{**}(\cdot;z^{\#})$ has the same structure as the equilibrium strategy $k^{\#}(\cdot;z^{\#})$, with (κ_0^*,κ_1^*) replacing (κ_0,κ_1), and γ^{**} replacing $\gamma^{\#}$. Hence,

$$PB_n(z^{\#}) = |u_{kk}|\left(\kappa_1\gamma^{\#}(z^{\#})\right)^2\left\{\frac{\left(\lambda_n(z^{\#})\right)^2}{2t_n\left(z_n^{\#}\right)^2} - \sum_{s=1}^{N}\frac{(1-\alpha\rho_s(z^{\#}))\lambda_s(z^{\#})}{\pi_s(z^{\#})}\frac{\partial\lambda_s(z^{\#})}{\partial z_n}\right\}$$

and

$$SB_n(z^\#) = |u_{kk} + u_{\sigma\sigma}| \left(\kappa_1^* \gamma^{**}(z^\#)\right)^2 \left\{ \frac{\left(\lambda_n(z^\#)\right)^2}{2t_n\left(z_n^\#\right)^2} - \sum_{s=1}^N \frac{(1-\alpha^*\rho_s(z^\#))\lambda_s(z^\#)}{\pi_s(z^\#)} \frac{\partial \lambda_s(z^\#)}{\partial z_n} \right\}.$$

In the proof of Proposition 5, I also established that, when the weights in the statistics X^j are those in (15),

$$\frac{\left(\lambda_n(z^\#)\right)^2}{2t_n\left(z_n^\#\right)^2} - \sum_{s=1}^N \frac{(1-\alpha\rho_s(z^\#))\lambda_s(z^\#)}{\pi_s(z^\#)} \frac{\partial \lambda_s(z^\#)}{\partial z_n} = \frac{\left(\pi_n(z^\#)\right)^2 \left\{ \frac{1}{2} - \frac{1-\alpha\rho_n(z^\#)}{1-\beta\alpha\rho_n(z^\#)} + \frac{\sum_{s=1}^N \frac{\pi_s(z^\#)\left(1-\alpha\rho_s(z^\#)\right)}{\left(1-\beta\alpha\rho_s(z^\#)\right)^2}}{\sum_{l=1}^N \frac{\pi_l(z^\#)}{1-\beta\alpha\rho_l(z^\#)}} \right\}}{\left(\sum_{l=1}^N \frac{\pi_l(z^\#)}{1-\beta\alpha\rho_l(z^\#)}\right)^2 t_n\left(z_n^\#\right)^2 \left(1-\beta\alpha\rho_n(z^\#)\right)^2}.$$

When $\beta = 0$, the above expression reduces to

$$\frac{\left(\pi_n(z^\#)\right)^2 \left[\frac{1}{2} + \alpha\left(\rho_n(z^\#) - \rho_X(z^\#)\right)\right]}{\left(\pi_X(z^\#)\right)^2 t_n\left(z_n^\#\right)^2}.$$

I thus have that

$$PB_n(z^\#) = |u_{kk}| \left(\kappa_1 \gamma^\#(z^\#)\right)^2 \left\{ \frac{\left(\pi_n(z^\#)\right)^2 \left[\frac{1}{2} + \alpha\left(\rho_n(z^\#) - \rho_X(z^\#)\right)\right]}{\left(\pi_X(z^\#)\right)^2 t_n\left(z_n^\#\right)^2} \right\}.$$

Similarly,

$$SB_n(z^\#) = |u_{kk} + u_{\sigma\sigma}| \left(\kappa_1^* \gamma^{**}(z^\#)\right)^2 \left\{ \frac{\left(\pi_n(z^\#)\right)^2 \left[\frac{1}{2} + \alpha^*\left(\rho_n(z^\#) - \rho_X(z^\#)\right)\right]}{\left(\pi_X(z^\#)\right)^2 t_n\left(z_n^\#\right)^2} \right\}.$$

Note that $\frac{1}{2} + \alpha\left(\rho_n(z^\#) - \rho_X(z^\#)\right)$ is positive as shown in (52).

Part (a). Consider economies in which $\kappa_1 = \kappa_1^*$ and $\alpha = \alpha^*$. Because the term in curly bracket is positive and $\gamma^\#(z^\#) = \gamma^{**}(z^\#)$, I conclude that, starting from $z^\#$, forcing the agents to pay more attention to a source n that receives strictly positive attention in equilibrium (i.e., for which $z_n^\# > 0$) increases welfare if $u_{\sigma\sigma} < 0$ and decreases if if $u_{\sigma\sigma} > 0$.

Part (b). Next, consider economies in which $u_{\sigma\sigma} = 0$, and $\alpha = \alpha^*$. Again, because $\gamma^\#(z^\#) = \gamma^{**}(z^\#)$, I conclude that, starting from $z^\#$, forcing the agents to pay more attention to a source n that receives strictly positive attention in equilibrium (i.e., for which $z_n^\# > 0$) increases welfare if $|\kappa_1^*| > |\kappa_1|$ and decreases it if $|\kappa_1^*| < |\kappa_1|$.

Part (c). Finally, consider economies in which $\kappa_1 = \kappa_1^*$ and $u_{\sigma\sigma} = 0$. In this case,

$$SB_n(z^\#) - PB_n(z^\#) \overset{\text{sgn}}{=} Q(z^\#, \alpha, \alpha^*) + \left[\alpha^*\left(\gamma^{**}(z^\#)\right)^2 - \alpha\left(\gamma^\#(z^\#)\right)^2\right] \rho_n(z^\#),$$

where

$$Q(z^\#, \alpha, \alpha^*) \equiv \left(\gamma^{**}(z^\#)\right)^2 \left[\frac{1}{2} - \alpha^*\rho_X(z^\#)\right] - \left(\gamma^\#(z^\#)\right)^2 \left[\frac{1}{2} - \alpha\rho_X(z^\#)\right].$$

The result in part (c) in the proposition then follows by observing that (i) Q is the same across all sources of information, (ii) when $\alpha\left(\gamma^{\#}(z^{\#})\right)^2 > \alpha^*\left(\gamma^{**}(z^{\#})\right)^2$, $SB_n(z^{\#}) - PB_n(z^{\#})$ is decreasing in ρ_n, whereas the opposite is true when $\alpha\left(\gamma^{\#}(z^{\#})\right)^2 < \alpha^*\left(\gamma^{**}(z^{\#})\right)^2$. Finally note that, when $\pi_\theta \to 0$, $\gamma^{\#}(z^{\#}), \gamma^{**}(z^{\#}) \to 1$, in which case $Q \to (\alpha - \alpha^*)\,\rho_X(z^{\#})$, implying that $\bar{\rho} \to \rho_X(z^{\#})$. Q.E.D.

References

[1] Alonso, R., W. Dessein, and N. Matouschek, 2008, "When Does Coordination Require Centralization?" *American Economic Review* 98(1), 145–79.

[2] Angeletos, G.M. and A. Pavan, 2007, "Efficient Use of Information and Social Value of Information," *Econometrica*, 75, 1103-1142.

[3] Angeletos, G.M. and A. Pavan, 2009, "Policy with Dispersed Information," *Journal of the European Economic Association*, 11-60.

[4] Angeletos, G.M. and K. Sastry, 2022, "Inattentive Economies" mimeo MIT.

[5] Amir, R., and N. Lazzati, 2016, "Endogenous Information Acquisition in Bayesian Games with Strategic Complementarities," *Journal of Economic Theory*, 163: 684–698.

[6] Banerjee, S., J. Davis, and N. Gondhi, 2022, "Motivated Beliefs in Coordination Games," mimeo UNC.

[7] Benabou R., and J. Tirole, 2004, "Willpower and Personal Rules," *Journal of Political Economy*, 112(4), 848-886.

[8] Bergemann, D. and S. Morris, 2013, "Robust Predictions in Games with Incomplete Information," *Econometrica*, 81(4), 1251-1308.

[9] Calvo-Armengol, A., J. Marti', and A. Prat, 2015, "Communication and Influence," *Theoretical Economics*, 10, 649–690.

[10] Chahrour, R., 2014, "Public Communication and Information Acquisition," *American Economic Journal: Macroeconomics*, 6-3, 73-101.

[11] Colombo, L., G. Femminis, and A. Pavan, 2014, "Information Acquisition and Welfare," *Review of Economic Studies,* 81, 1438-1483.

[12] Colombo, L., G. Femminis, and A. Pavan, 2023, "Optimal Fiscal and Monetary Policy with Investment Spillovers and Endogenous Private Information," mimeo, Catholic University of Milan and Northwestern University.

[13] Denti. T., 2016, "Unrestricted Information Acquisition" forthcoming in *Theoretical Economics.*

[14] Dessein, W., A. Galeotti, and T. Santos, 2016, "Rational Inattatention and Organziational Focus," *American Economic Review*, 106, 1522–36.

[15] Dewan, T. and D. P. Myatt, 2008, "The Qualities of Leadership: Direction, Communication, and Obfuscation," *American Political Science Review* 102(3), 351-368.

[16] Galperti, S. and I. Trevino, 2016, "Competition for Attention, Information Supply, and Coordination," mimeo University of California at San Diego.

[17] Galeotti, A., B. Golub, and S. Goyal, 2020, "Targeting Interventions in Networks," *Econometrica*, 88(6), 2445-2471.

[18] Hellwig, C. and L. Veldkamp, 2009, "Knowing What Others Know: Coordination Motives in Information Acquisition," *Review of Economic Studies,* 76, 223-251.

[19] Hebert, B. and J. Jennifer La'O, 2022, "Information Acquisition, Efficiency, and Non-Fundamental Volatility" forthcoming in the *Journal Of Political Economy.*

[20] Herskovic, B. and J. Ramos, 2020, "Acquiring Information Through Peers," *American Economic Review*, 110(7), 2128-2152.

[21] Jehiel, P., 2005, "Analogy-Based Expectation Equilibrium," *Journal of Economic Theory* 123, 81–104.

[22] Jehiel, P. and D. Samet, 2007, "Valuation Equilibrium," *Theoretical Economics* 2, 163–185.

[23] Jehiel, P. and F. Koessler, 2008, "Revisiting Games of Incomplete Information with Analogy-based Expectations," *Games and Economic Behavior* 62, 533–557.

[24] Jehiel, P. and L. Samuelson, 2012, "Reputation with Analogical Reasoning," *Quarterly Journal of Economics*, 1-43.

[25] Kahneman, D., 1973, "Attention and Effort," Englewood Cliffs, NJ: Prentice-Hall.

[26] Kahneman, D., 2011, "Thinking, Fast and Slow," New York: Farrar, Strauss, Giroux.

[27] Kahneman, D., Slovic, P., and A. Tversky, 1982, "Judgment Under Uncertainty: Heuristics and Biases," New York: Cambridge University Press.

[28] Kocer, Y., 2010, "Endogenous Learning with Bounded Memory," Job Market Paper, *mimeo*, Princeton University.

[29] Lambert, N., G. Martini, and M. Ostrovsky, 2018, "Quadratic Games," mimeo USC and Stanford University.

[30] Liang, A., and X. Mu, 2020, "Complementary Information and Learning Traps," *Quarterly Journal of Economics*, 135(1): 389–448.

[31] Llosa, L.G. and V. Venkateswaran, 2022, "Efficiency Under Endogenous Information Choice: A Price-Setting Application, *Journal of the European Economic Association*, 20(4),1619-1646.

[32] Maćkowiak, B. and M. Wiederholt, 2009, "Optimal Sticky Prices under Rational Inattention," *American Economic Review*, 99, 769-803.

[33] Maćkowiak, B. and M. Wiederholt, 2012, "Inattention to Rare Events," *mimeo*, European Central Bank.

[34] Maćkowiak, B., F. Matějka, and M. Wiederholt, 2022, "Rational Inattention: A Review," *Journal of Economic Literature*, forthcoming.

[35] Mailath, G. And L. Samuelson, 2006, "Repeated Games and Reputations: Long-Run Relationships," Oxford University Press.

[36] Matejka, F. and A. McKay, 2012, "Simple Market Equilibria with Rationally Inattentive Consumers," *American Economic Review: Papers & Proceedings* 102(3), 24-29.

[37] Matejka, F. and A. McKay, 2014, "Rational Inattention to Discrete Choices: A New Foundation for the Multinomial Logit Model," *American Economic Review*, 105 (1), 272-298.

[38] Myatt, D. P. and C. Wallace, 2012, "Endogenous Information Acquisition in Coordination Games," *Review of Economic Studies*, 79, 340-374.

[39] Myatt, D. P., and C. Wallace, "Central Bank Communication Design in a Lucas-Phelps Economy," *Journal of Monetary Economics* 63, 64-79.

[40] Myatt, D. P. and C. Wallace, 2015, "Cournot Competition and the Social Value of Information," *Journal of Economic Theory*, 158(B), 466-506.

[41] Myatt, D. P. and C. Wallace, 2018, "Information Use and Acquisition in Price-Setting Oligopolies," *The Economic Journal* 128(609), 845-886.

[42] Monderer D. and L. Shapley, 1995, "Potential Games," *Games and Economic Behavior*, 14, 124–143.

[43] Morris, S. and H.S. Shin, 2002, "The Social Value of Public Information," *American Economic Review*, 92, 1521-1534.

[44] Morris, S. and M. Yang, 2022, "Coordination and Continuous Stochastic Choice," *The Review of Economic Studies*, 89(5), 2687-2722.

[45] Mullainathan S., 2002, "A Memory-Based Model of Bounded Rationality," *Quarterly Journal of Economics*, 117(3), 735-774.

[46] Pavan, A., S. Sundareasn, and X. Vives, 2023, "(In)efficiency in Information Acquisition and Aggregation through Prices," mimeo IESE, Imperial College, and Northwestern University.

[47] Ragas, M. W., and H. L. Tran, 2015, "The financial news ecosystem: Journalists' perceptions of group hierarchy," *Journalism*, 16(6), 711-729.

[48] Sims, C.A., 2003, "Implications of Rational Inattention," *Journal of Monetary Economics*, 50, 665-690.

[49] Sims, C.A., 2011, "Rational Inattention and Monetary Economics," *Handbook of Monetary Economics*, 3A, 155-182.

[50] Steiner, J., C. Stewart, and F. Matějka, 2017, "Rational Inattention Dynamics: Inertia and Delay in Decision-Making," *Econometrica*, 85(2), 521-553.

[51] Szkup, M. and I. Trevino, 2015, "Information Acquisition in Global Games of Regime Change," *Journal of Economic Theory*, 2015, 160, 387-428.

[52] Van Zandt, T., and X. Vives, 2007, "Monotone Equilibria in Bayesian Games of Strategic Complementarities, *Journal of Economic Theory*, 134, 339-360.

[53] Vives, X., 2017, "Endogenous Public Information and Welfare in Market Games," *The Review of Economic Studies*, 84(2), 935-963.

[54] Wilson, A., 2014, "Bounded Memory and Biases in Information Processing*, Econometrica,* 82(6), 2257-2294.

[55] Yang, M., 2015, "Coordination with Flexible Information Acquisition," *Journal of Economic Theory*, 158(B), 721–738.