

Article

# On Information Aggregation and Interim Efficiency in Networks

Antonio Jimenez-Martinez

División de Economía, Centro de Investigación y Docencia Económicas (CIDE), Mexico City 01210, Mexico; antonio.jimenez@cide.edu; Tel.: +52-55-5727-9800

Academic Editor: Christos Dimitrakakis

Received: 31 October 2016; Accepted: 21 February 2017; Published: 24 February 2017

**Abstract:** This paper considers a population of agents that are engaged in a listening network. The agents wish to match their actions to the true value of some uncertain (exogenous) parameter and to the actions of the other agents. Each agent begins with some initial information about the parameter and, in addition, is able to receive further information from their neighbors in the network. I derive a closed expression for the (interim) social welfare loss that depends on the initial information structure and on the possible pieces of information that can be gathered under the network. Then, I explore how changes in the network may affect social welfare for extreme levels of complementarity in the agents' actions. When the level of complementarity is very high, efficiency is achieved regardless of the network structure. For very low levels of complementarity in actions, efficiency can be either associated to more sparse or denser networks, depending on the size of the induced informative gains. The implications of this paper are relevant in security environments where agents are naturally interpreted as analysts who try to forecast the value of a parameter that describes a threat to security.

**Keywords:** networks; information aggregation; beauty-contests; strategic complementarity; social welfare

**JEL Classification:** C72; D83; D84; D85

---

## 1. Introduction

In many environments of social, economic, or political interest, decision-makers seek to match their actions both to some unknown underlying variable and to the actions chosen by others. While the first motive is typically regarded as a “fundamental motive,” the second motive is purely a “coordination motive”<sup>1</sup>. The canonical framework that captures both types of motives in strategic scenarios is that of “beauty contest” games<sup>2</sup>. Beauty contests are particularly suitable to capture environments where security issues are a main concern. A typical example is that of a group of analysts who independently try to predict the value of an exogenous variable that describes a threat to security. Each of the analysts wishes to follow an action appropriate for the true value of the uncertain variable. In addition, since coordination helps prevent (or at least mitigate) security threats, the analyst wishes

---

<sup>1</sup> For example, suppose that the profitability of some investment activity depends on an uncertain exogenous state of the world and on the aggregate investment. Here, investors would like to pick investment strategies that match both the exogenous variable and the other investors' strategies as well.

<sup>2</sup> The “beauty contest” terminology comes originally from a well-known parable by ([1], Chapter 12). Following the seminal contribution of [2], “beauty contest” games have been extensively used to explore a wide range of phenomena in a number of settings, including investment games ([3,4]), financial markets ([5]), monopolistic competition ([6]), or models of political leadership ([7]), among others.

to follow a course of action similar to the other analysts' actions as well<sup>3</sup>. In these beauty contest scenarios, decision-makers wish to collect information that helps them resolve their uncertainty about the underlying variable and the likely actions of the others.

In practice, the presence of networks is ubiquitous in contexts where agents collect information. When decision-makers can only interact locally through a network, the architecture of the network places restrictions on the ways in which they can aggregate their pieces of private information. The typical approach considers that decision-makers can only obtain information from their neighbors in the network. Often, security analysts are involved in networks where their neighbors are other forecasters. In many real-world situations, a network of security analysts encompasses formal connections between different organizations as well as more informal connections based on friendship, family, or informal online relationships. In these environments, efficiency insights provide particularly useful recommendations for the design of teams of analysts, as well as for the appropriateness of establishing collaboration links between different security organizations.

This paper considers a (relatively large) population of decision-makers (or security analysts) that have access to some private signals about the underlying state and that, in addition, can receive (locally, according to their connections in a network) the information that others obtain from their own signals. For this framework, the main contribution of this paper is to derive a closed form for the (interim) social welfare loss function (in Proposition 1) that depends, among other primitives, on the informativeness of the signals and on the collection of subsets of signals that the agents can aggregate under the network.

More specifically, each agent receives initially an informative signal about the uncertain variable and, in addition, he is able to add up to it other signals that were initially provided to his (direct) neighbors in the network. In this way, the network induces a collection of different subsets of signals that can be finally observed within the population. Intuitively, finer collections of subsets of signals are associated to more sparse networks whereas denser networks induce coarser collections. Using the welfare loss function derived in Proposition 1, I then explore how efficiency relates to changes in the collection of pieces of information that can be aggregated under the network, and how efficiency is affected by extreme values of the "coordination motive." The efficiency benchmark used in this paper considers interim utilities. Thus, the central planner is able to access the information available to the agents upon receiving their signals. This seems a reasonable approach in many environments. For security contexts, it is certainly appealing when one considers that the central planner is some central institution that coordinates the teams of analysts that constitute the network.

To investigate this normative question in the context of networks, one must understand the forces behind two different mechanisms. First, how does the informative content of the available signals change as the collection of subsets of signals induced by the network varies? Secondly, how does social welfare depend on the informative content of the collection of subsets of signals induced by the network? The first question refers to the scale properties of information aggregation within sets of signals. To answer the second question, one needs to investigate the *social value of information* when both a fundamental motive and a coordination motive are present<sup>4</sup>.

---

<sup>3</sup> For example, under a terrorist attack threat, the analyst wishes to assess which is the most likely location of the attack but also wants to come up with locations not very distant from those predicted by other analysts. In this way, counterterrorism measures could be more effective to prevent the attack.

<sup>4</sup> Ref. [4] have investigated in a very comprehensive way the social value of information in an ex-ante efficiency benchmark without restrictions in the form of local interactions and where the agents have both private and public sources of information. For that environment, they have shown that whether more informative content increases or decreases welfare depends on whether equilibrium is efficient under both complete and incomplete information or only under incomplete information. Their contribution highlights that understanding the social value of information depends crucially on the notion of efficiency used. Without a well-specified efficiency benchmark, assessing the social value of information follows the folk theorem that "everything goes" in a second-best world. Assessing the social value of information with complementarities in the presence of networks remains a question far from understood.

This paper provides a partial answer to the first question in Proposition 2, which shows that the informative content of a given subset of signals increases monotonically as we split it into smaller disjoint subsets. The paper's insights on efficiency for extreme levels the coordination motive are provided in Proposition 3. Efficiency is guaranteed, regardless of the network structure, for very high levels of complementarity in actions. For very low levels of complementarity in actions, social welfare increases as subsets of observed signals are split into smaller disjoint subsets if and only if the informative gains induced are high enough. As Example 2 illustrates, the intuitive message here is that, when agents are interested only in predicting the value of the uncertain variable, we want to have more sparse networks if and only if the informativeness of the available set of signals allows for high enough informative gains.

The rest of this paper is structured as follows. Section 2 lays out the model and Section 3 derives the social welfare loss function. Section 4 provides the results on the informative content of the information aggregated in the network and on efficiency. Section 5 concludes. While the derivation of equilibrium is a crucial step to obtain our main result, and to investigate efficiency in the proposed benchmark, it is also constructive. Therefore, the required technical details are included in the main text. Other technical details, such as the proofs of Lemma 1, and of Propositions 2 and 3, are relegated to the Appendix.

## 2. Model

There is a (measure-1) continuum of agents, indexed by  $i \in [0, 1]$ , that wish to estimate an unknown *state of the world*  $\theta \in \mathbb{R}$ . Each agent receives a utility according to a common (quadratic) *utility function*  $u(a, \theta)$  that depends on an *action profile*  $a : [0, 1] \rightarrow \mathbb{R}$ , where  $a(i)$  is the action chosen by agent  $i$ , and on the state  $\theta$ . This model can be broadly applied to many contexts of strategic interaction under uncertainty where both a fundamental and a coordination motive are present. For security environments, we can think of agents as security analysts or forecasters that wish to estimate a variable that describes a security threat.

### 2.1. Preferences and Information Structure

The main intuition is conveyed using a beauty contest game (as in [2,4,6-8]). Each agent  $i$ 's utility is given by

$$u(a, \theta) = - \left( a(i) - (1 - \lambda)\theta - \lambda \int_0^1 a(h)dh \right)^2, \quad (1)$$

where  $\lambda \in (0, 1)$  is a parameter that measures the degree of strategic complementarity in the agents' actions. Intuitively,  $\lambda$  captures the relative importance of the coordination motive in the agent's utility. Higher values of  $\lambda$  indicate higher levels of strategic complementarity in actions.

The state of the world  $\theta$  is unknown to the agents and the underlying information structure is assumed to be Gaussian with  $\theta \sim \mathcal{N}(0, \sigma^2)$ . Agents receive some (exogenous) information about  $\theta$  from a finite set of *noisy signals*  $S = \{s_1, s_2, \dots, s_n\}$  throughout two periods  $t \in \{0, 1\}$ . Although the set of agents is a continuum, the set of possible signals is assumed to be finite for tractability reasons. Specifically, signals are (exogenously) assigned to agents according to a finite partition  $\mathcal{N} = \{N_1, N_2, \dots, N_n\}$  of the set of agents  $[0, 1]$ , where  $N_1 = [0, i_1]$ ,  $N_j = (i_{j-1}, i_j]$  for each  $j \in \{2, \dots, n-1\}$ , and  $N_n = (i_{n-1}, 1]$ . Then, at  $t = 0$ , each agent  $i \in N_j$  receives a noisy signal  $s_j = \theta + \varepsilon_j$ , where  $\varepsilon_j \sim \mathcal{N}(0, \pi_j^{-1})$ . Thus,  $\pi_j$  measures the precision of the signal received by any agent  $i \in N_j$ . Each noise term  $\varepsilon_j$  is independent of the state  $\theta$  and the noise terms  $\{\varepsilon_j\}_{j=1}^n$  are independent from each other as well. Signals are assumed to be (conditional on the state) independent (i.e., the conditional random variables  $\{(s_j | \theta)\}_{j=1}^n$  are independent)<sup>5</sup>.

<sup>5</sup> Of course, signals cannot be unconditionally independent because all of them depend on the state of the world.

### 2.2. Network Structure

After receiving their initial information at  $t = 0$ , agents can listen to others at  $t = 1$  according to a (directed) network that *locally* allows them to receive other agents' signals. Agents are engaged in a pattern of listening connections or *directed links*, which determines the network. This paper considers that information flows one way between directedly linked agents and it also assumes that information cannot be transmitted to indirectly connected agents. Specifically, a *network* is a collection of neighborhoods  $\{g(i)\}_{i \in [0,1]}$  where *agent i's neighborhood*  $g(i)$  is the set of agents to whom agent  $i$  has a directed link. Thus, an agent's neighborhood includes the agents that he is connected to through a (single) directed link but not those agents that are connected through a *path* of (several) directed links. In other words, an agent  $i$  cannot hear what a neighbor  $h \in g(i)$  listens from other agents  $l \in g(h)$  such that  $l \notin g(i)$ . In this setup, a network can be conveniently specified using a *neighborhood function*

$$g : [0, 1] \rightarrow \mathcal{B}_{[0,1]},$$

where  $\mathcal{B}_{[0,1]}$  denotes the Borel  $\sigma$ -algebra on the interval  $[0, 1]$ . Each agent  $i$  can observe at  $t = 1$  the signals received at  $t = 0$  by all agents in his neighborhood,  $h \in g(i)$ . Each agent listens to himself,  $i \in g(i)$ .

For security environments, we can think of an agent's neighborhood as a group of security analysts or forecasters whom he consults for some security project. For a network described by a neighborhood function  $g$ , agent  $i$  can observe at  $t = 1$  the signals that all agents in his neighborhood (thus, including himself) received at  $t = 0$ . In other words, agent  $i$  can observe all signals  $s_j$  such that, for each neighbor  $h \in g(i)$ , we have  $h \in N_j$  for the element  $N_j$  of the partition  $\mathcal{N}$  of agents. Thus, let

$$s(i) = \{s_j \in S : g(i) \cap N_j \neq \emptyset \text{ for } N_j \in \mathcal{N}\}$$

denote a *restricted signal profile for agent i*, which consists of the subset of signals observed by agent  $i$  at  $t = 1$ , conditioned on his neighborhood  $g(i)$ . Let  $\mathcal{S}_g$  denote the collection of all the different subsets of signals that can be observed within the population under the neighborhood function  $g$  so that  $s(i) \in \mathcal{S}_g$  for any agent  $i \in [0, 1]$ . Sometimes, I will simply say that the collection  $\mathcal{S}_g$  is induced by the neighborhood function  $g$  or by the network. Since the set of available signals  $S$  is finite, both the collection of all subsets of signals  $\mathcal{S}_g$  and any restricted signal profile  $s(i) \in \mathcal{S}_g$  must be finite as well. When no reference to a particular agent that observes a subset of signals needs to be made, I will use  $s \in \mathcal{S}_g$  to denote a generic signal profile observed within the population. In addition, for technical reasons, we will consider throughout the paper any subset of signals  $s$  as an ordered string or vector that contains the elements of the subset<sup>6</sup>.

As an antidote to the abstractness of the model, and to grasp better its functioning, the goal of the following example is to illustrate how agents listen to others in a network.

**Example 1.** Consider a set of available signals  $S = \{s_1, s_2, s_3\}$  which the agents are endowed with at  $t = 0$  according to the partition  $\mathcal{N} = \{N_1, N_2, N_3\}$ , with  $N_1 = [0, 1/3]$ ,  $N_2 = (1/3, 2/3]$ , and  $N_3 = (2/3, 1]$ . Suppose that the network is specified according to the neighborhood function

$$g(i) = \begin{cases} [0, 3/10] & \text{if } i \in [0, 1/4] \\ [2/10, 5/10] & \text{if } i \in [1/4, 1/2] \\ [4/10, 8/10] & \text{if } i \in [1/2, 3/4] \\ [6/10, 1] & \text{if } i \in (3/4, 1]. \end{cases}$$

---

<sup>6</sup> For instance, suppose for some agent  $i$ , the profile  $s(i) = s \in \mathcal{S}_g$  is the subset of signals  $\{s_1, s_3, s_7, s_{100}\}$ . Then, we will consider  $s(i) = s = (s_1, s_3, s_7, s_{100})$  for technical tractability.

In this example, for each of the four neighborhoods, we can easily find another neighborhood with a non-empty intersection, so that the network is minimally connected<sup>7</sup>. As to which signals the agents can observe at  $t = 1$  under the network described by  $g$ , notice that

$$\begin{aligned} [0, 3/10] \cap N_1 &= [0, 3/10], [0, 3/10] \cap N_2 = \emptyset, [0, 3/10] \cap N_3 = \emptyset; \\ [2/10, 5/10] \cap N_1 &= [2/10, 1/3], [2/10, 5/10] \cap N_2 = (1/3, 5/10], [2/10, 5/10] \cap N_3 = \emptyset; \\ [4/10, 8/10] \cap N_1 &= \emptyset, [4/10, 8/10] \cap N_2 = [4/10, 2/3], [4/10, 8/10] \cap N_3 = (2/3, 8/10]; \\ [6/10, 1] \cap N_1 &= \emptyset, [6/10, 1] \cap N_2 = [6/10, 2/3], [6/10, 1] \cap N_3 = (2/3, 1]. \end{aligned}$$

Therefore, the restricted signal profiles that can be observed within the population at  $t = 1$  are given by

$$s(i) = \begin{cases} \{s_1\} & \text{if } i \in [0, 1/4] \\ \{s_1, s_2\} & \text{if } i \in [1/4, 1/2] \\ \{s_2, s_3\} & \text{if } i \in [1/2, 1], \end{cases}$$

so that the collection of subsets of signals observed by the agents, as induced by the neighborhood function  $g$ , is  $\mathcal{S}_g = \{\{s_1\}, \{s_1, s_2\}, \{s_2, s_3\}\}$ .

For instance, notice that all agents  $i \in [1/2, 1]$  are able to observe signals  $s_2$  and  $s_3$  at  $t = 1$ . Here, the subset of agents  $(2/3, 1] \subset [1/2, 1]$  were already able to observe  $s_3$  directly at  $t = 0$ ; as such, a signal was initially assigned to them. From this subset of agents  $(2/3, 1]$ , we observe that the agents in  $(3/4, 1]$  are linked to the agents in  $[6/10, 1]$  and, thus, through the agents in  $[6/10, 2/3] \subset [6/10, 1]$ , they are able to observe  $s_2$  as well at  $t = 1$ . In addition, from the subset of agents  $(2/3, 1]$ , we note that the agents in  $(2/3, 3/4]$  are linked to the agents in  $[4/10, 8/10]$  and, therefore, through the agents in  $[4/10, 2/3] \subset [4/10, 8/10]$ , they are able to observe  $s_2$  as well at  $t = 1$ . As to the remaining agents in  $[1/2, 2/3]$ , they were already able to observe  $s_2$  directly at  $t = 0$  as they were initially endowed with such a signal. In addition, we observe that all the agents in  $[1/2, 2/3]$  are linked to the agents in  $[4/10, 8/10]$ . Therefore, through the agents in  $(2/3, 8/10] \subset [4/10, 8/10]$ , they are able to observe  $s_3$  as well at  $t = 1$ . However, any agent  $0 \leq i < 1/2$  is unable to observe  $s_3$ . Such agents can only observe at  $t = 0$  either  $s_1$  (if  $0 \leq i \leq 1/3$ ) or  $s_2$  (if  $1/3 < i < 1/2$ ). In addition, they are linked with agents  $0 \leq h \leq 5/10$  who cannot observe  $s_3$  themselves at  $t = 0$ .

### 2.3. Optimal Actions

Given the informational constraints imposed by the network, the agents are engaged in a game where each agent  $i \in [0, 1]$  chooses at  $t = 1$  an action  $a^*(i)$  so as to maximize his conditional expected utility  $E[u(a, \theta) \mid s(i)]$ . Under the preference specification in (1), a Bayesian Nash equilibrium (BNE) is a function  $a^* : [0, 1] \rightarrow \mathbb{R}$  such that each agent  $i \in [0, 1]$  solves the problem

$$\min_{a(i) \in \mathbb{R}} E \left[ \left( a(i) - (1 - \lambda)\theta - \lambda \int_0^1 a(h) dh \right)^2 \mid s(i) \right].$$

Nonetheless, we can restrict attention to symmetric BNE where all agents that observe the same subset of signals optimally choose the same action. To see this, suppose that some agents  $h$  that observe a common subset of signals  $s'$  choose different optimal actions. Then, the expectation that other agents who receive a subset of signals  $s$  have about the average optimal action followed by the agents that observe  $s'$  is

<sup>7</sup> Formally, in the current content, the network is *minimally (directedly) connected* if for each neighborhood  $g(i) \in g([0, 1])$  there exists another neighborhood  $g(h) \in g([0, 1])$  such that  $g(i) \cap g(h) \neq \emptyset$ . I thank an anonymous referee for pointing out an error in an earlier version of this definition. This notion of minimally connectedness is equivalent to having, for any two different agents, at least a directed path in the network that connects them.

$$E \left[ \int_h a^*(h) dh \mid s \right] = \int_h E [a^*(h) \mid s] dh.$$

This expectation depends on the restricted signal profile  $s$  but not on the names of the agents. All agents that observe a common subset of signals aggregate the same information and obtain some common posteriors on  $\theta$  and on the actions chosen by other agents. In addition, since the loss function that the agents are minimizing is strictly convex, the corresponding best-reply must be unique. Therefore, we can write

$$\int_h E [a^*(h) \mid s] dh = E [a^*(s') \mid s],$$

where  $a^*(s')$  indicates the optimal action chosen by any agent  $h$  that observes the subset of signals  $s'$ . From here onwards, let us use for simplicity  $E_s[\cdot]$  and  $\text{Var}_s[\cdot]$  to indicate, respectively, the conditional expectation  $E[\cdot \mid s]$  and the conditional variance  $\text{Var}[\cdot \mid s]$  operators. Given the previous observations, for a network specified by  $g$ , an *action function*  $a^* : \mathcal{S}_g \rightarrow \mathbb{R}$  is a *symmetric BNE* if and only if each action  $a^*(s)$  satisfies

$$a^*(s) = (1 - \lambda)E_s[\theta] + \lambda \sum_{s' \in \mathcal{S}_g} E_s [a^*(s')]. \quad (2)$$

### 3. Equilibrium and Social Welfare

Consider that, after period  $t = 0$  and before period  $t = 1$ , a *central planner* can influence the structure of the network by changing the neighborhood function  $g$ . This type of intervention makes sense only if the social planner has access to the signals available to all agents in the population and thus makes uses of the information available at the *interim* stage of the game. In this sense, the current paper is addressing interim efficiency issues. The expected loss of all agents that observe a subset of signals  $s \in \mathcal{S}_g$ , under a symmetric BNE action function  $a^*$ , is

$$E_s \left[ \left( a^*(s) - (1 - \lambda)\theta - \lambda \sum_{s' \in \mathcal{S}_g} a^*(s') \right)^2 \right]. \quad (3)$$

The goal of this paper is to investigate how the network described by  $g$  influences the shape of the *social welfare loss function*

$$L(g) = \sum_{s \in \mathcal{S}_g} E_s \left[ \left( a^*(s) - (1 - \lambda)\theta - \lambda \sum_{s' \in \mathcal{S}_g} a^*(s') \right)^2 \right]. \quad (4)$$

Since the technical details required to obtain the welfare loss function in this setup are constructive, they are provided in the main text. Using such arguments, Proposition 1 then derives the relevant welfare loss function for our environment with fundamental and coordination motives.

To address this central question, we need first to characterize the class of linear symmetric BNE of the game that the agents play once they receive their signals at  $t = 1$  under the restrictions imposed by the network. As in the related literature (see, e.g., [2,4,6–8]), the existence of symmetric BNE is guaranteed under the quadratic-Gaussian structure that the model assumes. To obtain a solution to Equation (2), we must study how information is aggregated and how this influences the agents' optimal actions. For a neighborhood function  $g$  and for a subset of signals  $s \in \mathcal{S}_g$ , the pairs  $(\theta, s)$  are jointly normally distributed. Let us use  $\text{Cov}[\theta, s]$  to denote the vector of covariances between the state of the world and each of the signals in  $s$  and  $\text{Var}[s]$  to denote the variance–covariance matrix of the signals in  $s$ . It follows from some basic results on normal distributions that

$$E_s[\theta] = \text{Cov}[\theta, s]' \cdot \text{Var}[s]^{-1} \cdot s \quad (5)$$

and

$$\text{Var}_s[\theta] = \sigma^2 - \text{Cov}[\theta, \mathbf{s}]' \cdot \text{Var}[\mathbf{s}]^{-1} \cdot \text{Cov}[\theta, \mathbf{s}]. \tag{6}$$

Hence, normality ensures that the conditional expectations of the state are linear in the signals contained in  $\mathbf{s}$ . This implication allows us to focus the analysis of BNE on linear strategies. If the agents that observe a signal profile  $\mathbf{s}'$  use a linear strategy with respect to such signals, then the optimal action of the agents that observe another signal profile  $\mathbf{s}$  must be also linear in such signals. While linear strategies are in general fairly simple and intuitive to interpret, in the current context they are also robust.

Equation (2) reveals that the optimal action followed by the agents that observe a signal profile  $\mathbf{s}$  depend in a recursive way on the average posterior expectation over the true state. Hence, we need to account for arbitrarily higher-order average posterior expectations over  $\theta$ . To formalize these average posterior expectations, let  $\bar{E}[\theta] = \sum_{\mathbf{s} \in \mathcal{S}_g} E_s[\theta]$  be the average posterior expectation on the state over the collection of possible subsets of signals<sup>8</sup>. We begin with the 0-order average posterior expectation. Notice that the 0-order average posterior expectation must coincide with the true realization of the state so that we set  $\bar{E}^{(0)}[\theta] = \theta$ . Then, for the 1-order average posterior expectation, we have

$$\bar{E}^{(1)}[\theta] = \bar{E}[\bar{E}^{(0)}[\theta]] = \bar{E}[\theta] = \sum_{\mathbf{s} \in \mathcal{S}_g} E_s[\theta],$$

whereas for higher-order average posterior expectations, we use  $\bar{E}^{(m)}[\theta] = \bar{E}[\bar{E}^{(m-1)}[\theta]]$  to indicate in a recursive way the  $m$ -order average posterior expectation over  $\theta$ , for  $m \geq 2$ . With such higher-order average posterior expectations in place, recursive application of Equation (2) allows us to express the optimal action followed by the agents that observe  $\mathbf{s}$  as

$$\begin{aligned} a^*(\mathbf{s}) &= (1 - \lambda) \left[ E_s[\bar{E}^{(0)}[\theta]] + \lambda E_s[\bar{E}^{(1)}[\theta]] + \lambda^2 E_s[\bar{E}^{(2)}[\theta]] + \dots \right] \\ &= (1 - \lambda) \sum_{m=0}^{\infty} \lambda^m E_s[\bar{E}^{(m)}[\theta]]. \end{aligned} \tag{7}$$

Under the assumed information structure, we have  $\text{Cov}[\theta, \mathbf{s}] = \sigma^2 \underline{\mathbf{1}}$ , where  $\underline{\mathbf{1}}$  is a vector of ones with the same dimension as the number of signals contained in the restricted profile  $\mathbf{s}$ . Furthermore, recall that  $s_j = \theta + \varepsilon_j$ , where  $E_s[\varepsilon_j] = 0$  for all signals  $j = 1, \dots, n$ . Take a given realization of the state  $\theta$ . Then, using the expression in Equation (5), we obtain that  $E_s[\bar{E}^{(0)}[\theta]] = E_s[\theta]$ , for the 0-order average posterior expectation, and

$$\begin{aligned} E_s[\bar{E}^{(1)}[\theta]] &= E_s \left[ \left( \sigma^2 \sum_{\mathbf{s} \in \mathcal{S}_g} \underline{\mathbf{1}}' \cdot \text{Var}[\mathbf{s}]^{-1} \cdot \underline{\mathbf{1}} \right) \theta \right] \\ &= \left( \sigma^2 \sum_{\mathbf{s} \in \mathcal{S}_g} \underline{\mathbf{1}}' \cdot \text{Var}[\mathbf{s}]^{-1} \cdot \underline{\mathbf{1}} \right) E_s[\theta], \end{aligned}$$

for the 1-order average posterior expectation. Here again,  $\underline{\mathbf{1}}$  is a vector of ones whose dimension equals the number of signals in the profile  $\mathbf{s}$ . Let us use

$$\bar{\omega}_g = \sigma^2 \sum_{\mathbf{s} \in \mathcal{S}_g} \underline{\mathbf{1}}' \cdot \text{Var}[\mathbf{s}]^{-1} \cdot \underline{\mathbf{1}}$$

---

<sup>8</sup> Since this is an average over all subsets of signals,  $\bar{E}[\theta]$  equivalently indicates the average posterior expectation on  $\theta$  over all agents.



to denote the average of the inverses of the posterior variances of the state across signal profiles in the network. Given this notation for the average across (the inverse of) posterior variances, we can write  $\bar{E}[\theta] = \bar{\omega}_g \theta$  and iterate to obtain that  $\bar{E}^{(m)}[\theta] = \bar{\omega}_g^m \theta$  for each  $m \geq 0$ . Thus, we can express the equality in Equation (7) as

$$\begin{aligned} a^*(s) &= (1 - \lambda) \left[ 1 + \lambda \bar{\omega}_g + \lambda^2 \bar{\omega}_g^2 + \dots \right] E_s[\theta] \\ &= \left( \frac{1 - \lambda}{1 - \lambda \bar{\omega}_g} \right) E_s[\theta], \end{aligned} \quad (8)$$

where  $E_s[\theta]$  satisfies the equality in Equation (5). Now, if we average the expression above over all possible subsets of signals observed in the network, we obtain

$$\begin{aligned} \sum_{s \in \mathcal{S}_g} a^*(s) &= \left( \frac{1 - \lambda}{1 - \lambda \bar{\omega}_g} \right) \sum_{s \in \mathcal{S}_g} E_s[\theta] \\ &= \left( \frac{1 - \lambda}{1 - \lambda \bar{\omega}_g} \right) \bar{\omega}_g. \end{aligned}$$

Therefore, in a BNE, each agent that observes a subset of signals  $s$  wishes to match his action to the objective

$$(1 - \lambda)\theta + \lambda \sum_{s \in \mathcal{S}_g} a^*(s) = \left( \frac{1 - \lambda}{1 - \lambda \bar{\omega}_g} \right) \theta. \quad (9)$$

By plugging the expressions in Equations (8) and (9) into the expected loss function given by Equation (3), we obtain:

$$E_s \left[ \left( \frac{(1 - \lambda)E_s[\theta]}{1 - \lambda \bar{\omega}_g} - \frac{(1 - \lambda)\theta}{1 - \lambda \bar{\omega}_g} \right)^2 \right] = \left( \frac{1 - \lambda}{1 - \lambda \bar{\omega}_g} \right)^2 \text{Var}_s[\theta],$$

where the conditional variance  $\text{Var}_s[\theta]$  is given by the expression in Equation (6). By combining this with the expression in Equation (4), we obtain the the social welfare loss function is given by

$$L(g) = \left( \frac{1 - \lambda}{1 - \lambda \bar{\omega}_g} \right)^2 \sum_{s \in \mathcal{S}_g} \text{Var}_s[\theta].$$

A closing argument is needed to complete the analysis. Recall that our derivation of symmetric BNE has made use of the law of large numbers to average expectations on the state over subsets of signals. Therefore, we must consider that the number of possible subsets of signals is sufficiently large for our formal arguments to be appealing in the environment that we are studying. The analysis of higher-order beliefs used in this paper builds on the approach followed, among others, by [2,4,6,7]. As in these papers, the formal analysis used here also invokes the law of large numbers<sup>9</sup>. With this last consideration in place, the arguments provided in this section show that the welfare loss function has the form described by the following proposition.

<sup>9</sup> For applications where one considers instead a relatively small number of subsets of signals, the law of large numbers cannot be reasonably invoked to compute averages of expectations on the state. In these cases, keeping track of the higher-order beliefs that are required to characterize equilibria follows a completely different approach. In particular, under certain conditions, one can make use of the iterated application of a knowledge index matrix. The idea of using a knowledge index matrix to track individual arbitrarily higher-order beliefs in a network was originally proposed by [9]. An application of the knowledge index matrix to information acquisition problems in small populations has been recently provided by [8].



**Proposition 1** (Welfare Loss Function). *Suppose that the cardinality of  $\mathcal{S}_g$  is relatively large. Then, under the stated assumptions on preferences, the information structure, and the network of listening links, the social welfare loss function is given, as a function only of the model's primitives, by the expression*

$$L(g) = \left( \frac{1 - \lambda}{1 - \lambda \bar{\omega}_g} \right)^2 \sum_{s \in \mathcal{S}_g} \text{Var}_s[\theta], \quad (10)$$

where

$$\bar{\omega}_g = \sigma^2 \sum_{s \in \mathcal{S}_g} \mathbf{1}' \cdot \text{Var}[\mathbf{s}]^{-1} \cdot \mathbf{1}. \quad (11)$$

In this context, the social planner wishes to affect the neighborhood function  $g$  so as to minimize the welfare loss described by Equations (10) and (11) above.

#### 4. Informative Content of Signal Profiles and Efficiency

This section derives some results about the informative content of the signal profiles  $\mathbf{s} \in \mathcal{S}_g$  and about how this informative content affects the social welfare under the network described by  $g$ .

Let us use  $\gamma_g = |\mathcal{S}_g|$  to indicate the number of possible subsets of signals that can be observed within the population under the network described by  $g$ . Let us use  $k_s$  to indicate the number of signals included in the profile  $\mathbf{s}$ . In addition, in accordance with the earlier notation, I will use  $\pi_j(\mathbf{s})$  to indicate the inverse of the variance of the noise term associated to the coordinate  $s_j$  of the string  $\mathbf{s}$ , for each  $j = 1, \dots, k_s$ . Intuitively,  $\pi_j(\mathbf{s})$  indicates the precision of the  $j$ -th signal in the profile  $\mathbf{s}$ . The *informative content of the signals contained in  $\mathbf{s}$*  can be conveniently described by the number

$$w(\mathbf{s}) := \sigma^2 \mathbf{1}' \cdot \text{Var}[\mathbf{s}]^{-1} \cdot \mathbf{1}$$

which identifies the sum of all the entries of the inverse of the variance–covariance matrix of the signal profile  $\mathbf{s}$  (weighted using the variance of the state of the world). Notice that we can express the average of the inverses of the posterior variances of the state in the network as  $\bar{\omega}_g = \sum_{s \in \mathcal{S}_g} w(\mathbf{s})$ . Intuitively, each  $w(\mathbf{s})$  is simply a scalar that gives us some information about the joint precision of the signal profile  $\mathbf{s}$ . Higher values of  $w(\mathbf{s})$  are associated with lower degrees of noise in the corresponding signals and, therefore, to more informative signal profiles. The following lemma provides formally this implication.

**Lemma 1.** *The scalar  $w(\mathbf{s}) = \sigma^2 \mathbf{1}' \cdot \text{Var}[\mathbf{s}]^{-1} \cdot \mathbf{1}$  lies in the interval  $(0, 1/\sigma^2)$  and it increases (strictly) with the sum of the precision of all the signals contained in the profile  $\mathbf{s}$ ,  $\sum_{j=1}^{k_s} \pi_j(\mathbf{s})$ .*

To explore how welfare depends on the subset of collections of signals  $\mathcal{S}_g$ , we need to understand how the informative content of the available signals  $S$  changes as the collection of subsets of signals  $\mathcal{S}_g$  varies. On the one hand, the informative content of a group of signals naturally increases as its size increases or, in other words, as we add more signals to the set. This can be derived directly from Lemma 1 and it is very intuitive: enlarging a subset of signals increases its informative content. On the other hand, in principle, it is unclear how the informative content of a subset of signals changes when such a profile of signals is split into several (smaller) disjoint subsets. Proposition 2 below offers a type of “diminishing returns” result under which the final informative content increases monotonically as we split a signal profile into smaller disjoint subsets. Conversely, adding up two sets of signals decreases their informative content. Larger subsets seem to feature a type of “congestion” to aggregate information. The message conveyed here for security environments is that we want to provide agents with a higher number of (smaller) subsets of signals if we are interested in increasing the overall informative content of the sources of information available in the set of signals  $S$ .

**Proposition 2.** *The informative content of a profile of signals  $\bar{s} \in \mathcal{S}_g$  is strictly smaller than the aggregation of informative contents obtained by splitting the subset  $\bar{s}$  into two disjoint subsets, that is,  $w(\bar{s}) < w(\bar{s}_1) + w(\bar{s}_2)$  for any  $\bar{s}_1, \bar{s}_2$  such that  $\bar{s} = \bar{s}_1 \cup \bar{s}_2$  with  $\bar{s}_1 \cap \bar{s}_2 = \emptyset$ .*

The result of Proposition 2 is useful to propose ways to minimize the welfare loss function obtained in Proposition 1 as it establishes that the overall informative content described by the term  $\bar{w}_g$  always increases by separating signal profiles into smaller disjoint subsets of signals. Of course, adding up subsets of signals or splitting subsets of signals into smaller disjoint subsets are not the only ways in which changes in the neighborhood function  $g$  affect the collection of profiles  $\mathcal{S}_g$ . For instance, one can propose changes in the neighborhood function  $g$  such that a subset  $\bar{s} \in \mathcal{S}_g$  becomes split instead into two *non-disjoint* subsets  $\bar{s}_1$  and  $\bar{s}_2$ , where  $\bar{s} = \bar{s}_1 \cup \bar{s}_2$ . Unfortunately, for those complex cases, what one can say about efficiency depends very much on the various precisions  $\pi_j$  of the signals involved.

However, restricting attention to interventions that split signal profiles into disjoint sets does allow for some interesting insights on efficiency. When we move from a neighborhood function  $g$  to another  $g'$  such that a given signal profile  $\bar{s} \in \mathcal{S}_g$  is split into two disjoint subsets  $\bar{s}_1, \bar{s}_2 \in \mathcal{S}_{g'}$  (with  $\bar{s} = \bar{s}_1 \cup \bar{s}_2$  and  $\bar{s}_1 \cap \bar{s}_2 = \emptyset$ ), then we naturally have that  $g'$  specifies a more sparse network than the network specified by  $g$ . This is very intuitive as denser networks allow agents to have access to higher numbers of signals. Building up on Example 1, the following example illustrates this point.

**Example 2.** *As in Example 1, consider a set of available signals  $S = \{s_1, s_2, s_3\}$  which the agents are endowed with at  $t = 0$  according to the partition  $\mathcal{N} = \{N_1, N_2, N_3\}$ , with  $N_1 = [0, 1/3]$ ,  $N_2 = (1/3, 2/3]$ , and  $N_3 = (2/3, 1]$ . Now, suppose that the network is specified instead according to the neighborhood function*

$$g'(i) = \begin{cases} [0, 3/10] & \text{if } i \in [0, 1/4] \\ [2/10, 5/10] & \text{if } i \in [1/4, 1/2] \\ [4/10, 2/3] & \text{if } i \in [1/2, 2/3] \\ (2/3, 1] & \text{if } i \in (2/3, 1]. \end{cases}$$

Notice that there is a neighborhood,  $(2/3, 1]$ , which is disjoint to any other neighborhood so that the network is not minimally connected. In addition, we observe that  $g'$  gives us a network more sparse than the network associated to the neighborhood function  $g$  analyzed in Example 1. In particular, the agents  $i \in [1/2, 2/3]$  are now linked to a smaller set of agents,  $[4/10, 2/3]$ , compared to the one they were linked to under  $g$ ,  $[4/10, 8/10]$ . In addition, the agents  $i \in (2/3, 1]$  are now linked to a smaller set of agents,  $(2/3, 1]$ , compared to the one they were linked to under  $g$ ,  $[6/10, 1]$ .

As to which signals the agents can observe at  $t = 1$  under the network described by  $g'$ , notice that

$$\begin{aligned} [0, 3/10] \cap N_1 &= [0, 3/10], [0, 3/10] \cap N_2 = \emptyset, [0, 3/10] \cap N_3 = \emptyset; \\ [2/10, 5/10] \cap N_1 &= [2/10, 1/3], [2/10, 5/10] \cap N_2 = (1/3, 5/10], [2/10, 5/10] \cap N_3 = \emptyset; \\ [4/10, 2/3] \cap N_1 &= \emptyset, [4/10, 2/3] \cap N_2 = [4/10, 2/3], [4/10, 8/10] \cap N_3 = \emptyset; \\ (2/3, 1] \cap N_1 &= \emptyset, (2/3, 1] \cap N_2 = \emptyset, (2/3, 1] \cap N_3 = (2/3, 1]. \end{aligned}$$

Therefore, the restricted signal profiles that can be observed within the population at  $t = 1$  are given by

$$s(i) = \begin{cases} \{s_1\} & \text{if } i \in [0, 1/4] \\ \{s_1, s_2\} & \text{if } i \in [1/4, 1/2] \\ \{s_2\} & \text{if } i \in [1/2, 2/3] \\ \{s_3\} & \text{if } i \in (2/3, 1], \end{cases}$$

so that the collection of subsets of signals observed by the agents, as induced by the neighborhood function  $g'$ , is  $\mathcal{S}_{g'} = \{\{s_1\}, \{s_1, s_2\}, \{s_2\}, \{s_3\}\}$ . Thus, by moving from neighborhood function  $g$  of Example 1 to neighborhood function  $g'$ , we are able to split the profile  $\{s_2, s_3\}$  into two disjoint profiles,  $\{s_2\}$  and  $\{s_3\}$ .

To obtain some insights about efficient information aggregation networks, I turn now to derive a more tractable expression for the welfare loss function. From the expression in Equation (6), and under our normality assumptions, it follows that the average of posterior variances of the state across signal profiles is given by

$$\sum_{s \in \mathcal{S}_g} \text{Var}_s[\theta] = \sigma^2 \left( \gamma_g - \sum_{s \in \mathcal{S}_g} w(s) \right) = \sigma^2 (\gamma_g - \bar{w}_g).$$

Finally, as to the term  $w(s) = \sigma^2 \mathbf{1}' \cdot \text{Var}[s]^{-1} \cdot \mathbf{1}$ , the proof of Lemma 1 above derives

$$w(s) = \frac{\sum_{j=1}^{k_s} \pi_j(s)}{1 + \sigma^2 \sum_{j=1}^{k_s} \pi_j(s)} \in (0, \sigma^{-2}).$$

Therefore, the expressions obtained in Equations (10) and (11) for the social welfare loss can be rewritten as

$$L(g) = \sigma^2 (1 - \lambda)^2 \frac{(\gamma_g - \bar{w}_g)}{(1 - \lambda \bar{w}_g)^2}, \quad \text{where } \bar{w}_g = \sum_{s \in \mathcal{S}_g} \left( \frac{\sum_{j=1}^{k_s} \pi_j(s)}{1 + \sigma^2 \sum_{j=1}^{k_s} \pi_j(s)} \right). \quad (12)$$

The social planner would like to influence the neighborhood function  $g$  so as to induce a collection of signal profiles  $\mathcal{S}_g$  that minimizes the expression in Equation (12) above.

On the one hand, from the expression for the welfare loss derived in Proposition 1, which is rewritten in Equation (12) above, we observe that the welfare loss is always minimized as  $\lambda$  tends to 1, regardless of the collection of profiles induced by the network structure. Since the welfare loss function  $L(g)$  is continuous in  $\lambda \in (0, 1)$ , it follows that when the level of complementarity in actions is very high, interim efficiency does not depend substantially on the network structure.

On the other hand, to explore some insights when complementarity in actions is very low, let us consider two different neighborhood functions  $g$  and  $g'$  such that, moving from  $g$  to  $g'$  only implies that some profile  $\bar{s} \in \mathcal{S}_g$  is split into two disjoint profiles  $\bar{s}_1, \bar{s}_2 \in \mathcal{S}_{g'}$ . Specifically, (i) for some given  $\bar{s} \in \mathcal{S}_g$ , we have  $\bar{s}_1, \bar{s}_2 \in \mathcal{S}_{g'}$ , where  $\bar{s} = \bar{s}_1 \cup \bar{s}_2$  and  $\bar{s}_1 \cap \bar{s}_2 = \emptyset$ ; whereas (ii) for each  $s \in \mathcal{S}_g \setminus \{\bar{s}\}$ , we have  $s \in \mathcal{S}_{g'}$ . Then, following some earlier arguments and, in particular, the result of Proposition 2, I ask how splitting one signal profile into two disjoint subsets affects the welfare loss when the level of complementarity in actions is very low. Proposition 3 below shows that the answer to this question depends on the size of the informative gain derived from splitting the selected signal profile.

**Proposition 3.** Consider a collection of signal profiles  $\mathcal{S}_g$  that contains a relatively high number of profiles and split some given signal profile  $\bar{s} \in \mathcal{S}_g$  into two disjoint subsets,  $\bar{s}_1$  and  $\bar{s}_2$ . Let  $\mathcal{S}_{g'}$  denote the collection of signal profiles  $\mathcal{S}_{g'} = \{s \in \mathcal{S}_g : s \neq \bar{s}\} \cup \{\bar{s}_1, \bar{s}_2\}$  which results from splitting the chosen profile  $\bar{s}$ . Suppose that complementarity in actions is very low,  $\lambda \rightarrow 0$ . Then, under the stated assumptions on preferences, the information structure, and the network of listening links, it follows that  $L(g') < L(g)$  if and only if  $[w(\bar{s}_1) + w(\bar{s}_2)] - w(\bar{s}) > 1$ .

Proposition 3 states that, for very low levels of complementarity in actions, splitting a signal profile into two disjoint subsets is welfare-improving if and only if the induced informative gain is sufficiently high. In other words, as illustrated by Example 2, when the agents wish to follow a course of action very close to the fundamental parameter, more sparse networks increases social welfare whenever the induced informative gains are high enough. Splitting a signal profile into disjoint signal profiles always increases the informative content (as shown by Proposition 2). However, if the gain in

informative content is not sufficiently high, then denser networks are associated with higher social welfare for very low levels of complementarity in actions.

### 5. Conclusions

This paper has proposed a benchmark to explore (interim) efficiency issues in environments where decision-makers are connected through a listening network and where both a fundamental motive and a coordination motive are present in their preferences. Investigating this topic is of interest when one uses an interim efficiency benchmark. Using the ex-ante efficiency approach instead, one would obtain directly that a network is efficient if and only if it allows all agents to receive all available signals. The reason behind this implication is that, for the case of beauty contest preferences, ex-ante efficiency requires the central planner to solve the same problem that faces each agent. As a consequence, any additional source of information would be welfare-improving.

This paper has derived a closed form for the welfare loss function that depends on the informativeness of the available signals and on the way in which the network enables the agents to gather information. Efficiency is achieved for very high levels of complementarity in actions, regardless of the network structure. For very low levels of complementarity in actions, the central planner wishes to induce finer collections of possible subsets of observed signals if and only if the derived informative gains are sufficiently high. The implications of this paper are useful to provide efficiency recommendations for networks of security analysts that are coordinated by a central institution under the requirement that such an institution can access the pieces of private information available to the analysts.

A natural extension of the setting explored here would be that of considering substitutive actions as well. This possibility seems reasonable in certain contexts, but it is perhaps not very appealing in security environments where coordination helps to prevent security threats. At a more general level, understanding the social value of information when information can be aggregated only locally within neighborhoods in networks, under different efficiency benchmarks and preference specifications, remains an interesting and rather unexplored question.

**Acknowledgments:** I am indebted to the Editor in charge, Christos Dimitrakakis, and to two anonymous referees for very useful comments. I gratefully acknowledge financial support from Consejo Nacional de Ciencia y Tecnología, Sistema Nacional de Investigadores, grant 41826. I thank Alessandro Pavan for his feedback. This research project was conducted while visiting the Department of Economics at UC San Diego. I thank this institution for its generous hospitality and support. Any remaining errors are my own.

**Conflicts of Interest:** The author declares no conflict of interest.

### Appendix

**Proof of Lemma 1.** Under our normality assumptions, we have

$$\text{Var}[\mathbf{s}] = \begin{pmatrix} \sigma^2 + \pi_1^{-1}(\mathbf{s}) & \sigma^2 & \dots & \sigma^2 \\ \sigma^2 & \sigma^2 + \pi_2^{-1}(\mathbf{s}) & \dots & \sigma^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma^2 & \sigma^2 & \dots & \sigma^2 + \pi_{k_s}^{-1}(\mathbf{s}) \end{pmatrix} = \sigma^2 \mathbf{1} \cdot \mathbf{1}' + D,$$

where  $D$  is the diagonal matrix  $D = \text{diag} \left( \pi_j^{-1}(\mathbf{s}) \right)_{j=1, \dots, k_s}$  that contains the variances of the noises of the signals in the profile  $\mathbf{s}$ . Notice that

$$D^{-1} = \begin{pmatrix} \pi_1(\mathbf{s}) & 0 & \dots & 0 \\ 0 & \pi_2(\mathbf{s}) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \pi_{k_s}(\mathbf{s}) \end{pmatrix}.$$

Using a version of the Sherman–Morrison’s formula to compute the inverse of a sum of matrices, (see, e.g., [10,11]), we obtain

$$\text{Var}[\mathbf{s}]^{-1} = D^{-1} - \frac{\sigma^2}{1 + \sigma^2 \mathbf{1}' \cdot D^{-1} \cdot \mathbf{1}} \left[ D^{-1} \cdot \mathbf{1} \cdot \mathbf{1}' \cdot D^{-1} \right].$$

It follows that

$$\frac{\sigma^2}{1 + \sigma^2 \mathbf{1}' \cdot D^{-1} \cdot \mathbf{1}} = \frac{\sigma^2}{1 + \sigma^2 \sum_{j=1}^{k_s} \pi_j(\mathbf{s})}$$

and

$$\left[ D^{-1} \cdot \mathbf{1} \cdot \mathbf{1}' \cdot D^{-1} \right] = \begin{pmatrix} \pi_1^2(\mathbf{s}) & \pi_1(\mathbf{s})\pi_2(\mathbf{s}) & \cdots & \pi_1(\mathbf{s})\pi_{k_s}(\mathbf{s}) \\ \pi_2(\mathbf{s})\pi_1(\mathbf{s}) & \pi_2^2(\mathbf{s}) & \cdots & \pi_2(\mathbf{s})\pi_{k_s}(\mathbf{s}) \\ \vdots & \vdots & \ddots & \vdots \\ \pi_{k_s}(\mathbf{s})\pi_1(\mathbf{s}) & \pi_{k_s}(\mathbf{s})\pi_2(\mathbf{s}) & \cdots & \pi_{k_s}^2(\mathbf{s}) \end{pmatrix}.$$

By doing the algebra, it then follows that

$$\text{Var}[\mathbf{s}]^{-1} = \frac{1}{1 + \sigma^2 \sum_{j=1}^{k_s} \pi_j} \times \begin{pmatrix} \pi_1 + \sigma^2 \sum_{j \neq 1} \pi_1 \pi_j & -\sigma^2 \pi_1 \pi_2 & \cdots & -\sigma^2 \pi_1 \pi_{k_s} \\ -\sigma^2 \pi_2 \pi_1 & \pi_2 + \sigma^2 \sum_{j \neq 2} \pi_1 \pi_j & \cdots & -\sigma^2 \pi_2 \pi_{k_s} \\ \vdots & \vdots & \ddots & \vdots \\ -\sigma^2 \pi_{k_s} \pi_1 & -\sigma^2 \pi_{k_s} \pi_2 & \cdots & \pi_{k_s} + \sigma^2 \sum_{j \neq k_s} \pi_{k_s} \pi_j \end{pmatrix},$$

where the arguments ( $\mathbf{s}$ ) have been conveniently dropped to simplify the expression. Therefore, in order to obtain a closed expression for the scalar  $w(\mathbf{s}) = \mathbf{1}' \cdot \text{Var}[\mathbf{s}]^{-1} \mathbf{1}$ , we need to aggregate all the entries of the matrix obtain above. We obtain

$$w(\mathbf{s}) = \frac{\sum_{j=1}^{k_s} \pi_j(\mathbf{s})}{1 + \sigma^2 \sum_{j=1}^{k_s} \pi_j(\mathbf{s})} \in (0, \sigma^{-2}). \tag{A1}$$

We observe directly that higher values of the (additive) aggregation of the precision of the signals  $\sum_{j=1}^{k_s} \pi_j(\mathbf{s})$ , contained in the signal profile  $\mathbf{s}$ , determine higher values of the scalar  $w(\mathbf{s})$ .  $\square$

**Proof of Proposition 2.** Let us use  $\pi_s = \sum_{j=1}^{k_s} \pi_j(\mathbf{s})$  as a shorthand notation for the sum of the precision of the signals contained in a signal profile  $\mathbf{s} \in \mathcal{S}_g$ . Consider a signal profile  $\bar{\mathbf{s}}$  and suppose that we split it into two disjoint subsets of signals  $\bar{\mathbf{s}}_1, \bar{\mathbf{s}}_2$ . Thus, we are considering  $\bar{\mathbf{s}} = \bar{\mathbf{s}}_1 \cup \bar{\mathbf{s}}_2$  with  $\bar{\mathbf{s}}_1 \cap \bar{\mathbf{s}}_2 = \emptyset$ . First, notice that  $\pi_{\bar{\mathbf{s}}} = \pi_{\bar{\mathbf{s}}_1} + \pi_{\bar{\mathbf{s}}_2}$ . Secondly, it follows from expression (A1) obtained in the proof of Lemma 1 that

$$\begin{aligned} [w(\bar{\mathbf{s}}_1) + w(\bar{\mathbf{s}}_2)] - w(\bar{\mathbf{s}}) &= \frac{\pi_{\bar{\mathbf{s}}_1}}{1 + \sigma^2 \pi_{\bar{\mathbf{s}}_1}} + \frac{\pi_{\bar{\mathbf{s}}_2}}{1 + \sigma^2 \pi_{\bar{\mathbf{s}}_2}} - \frac{\pi_{\bar{\mathbf{s}}_1} + \pi_{\bar{\mathbf{s}}_2}}{1 + \sigma^2 (\pi_{\bar{\mathbf{s}}_1} + \pi_{\bar{\mathbf{s}}_2})} \\ &= \frac{\sigma^2 \pi_{\bar{\mathbf{s}}_1} \pi_{\bar{\mathbf{s}}_2} (2 + \sigma^2 \pi_{\bar{\mathbf{s}}_1} + \sigma^2 \pi_{\bar{\mathbf{s}}_2})}{(1 + \sigma^2 \pi_{\bar{\mathbf{s}}_1})(1 + \sigma^2 \pi_{\bar{\mathbf{s}}_2})[1 + \sigma^2 (\pi_{\bar{\mathbf{s}}_1} + \pi_{\bar{\mathbf{s}}_2})]} > 0, \end{aligned}$$

as stated.  $\square$

**Proof of Proposition 3.** For the selected signal profile  $\bar{\mathbf{s}} \in \mathcal{S}_g$ , let us use  $\alpha = \sum_{\mathbf{s} \in \mathcal{S}_g \setminus \{\bar{\mathbf{s}}\}} w(\mathbf{s})$  as shorthand notation for the informative content of all profiles different from  $\bar{\mathbf{s}}$  in the collection  $\mathcal{S}_g$ . Using

the expression for the welfare loss obtained in Equation (12), it follows that the sign of the difference  $L(g) - L(g')$  coincides with the sign of the expression

$$(1 - \lambda\alpha)^2 ([w(\bar{s}_1) + w(\bar{s}_2)] - w(\bar{s}) - 1) + \lambda^2 \left( (\gamma_g - \alpha) \left[ [w(\bar{s}_1) + w(\bar{s}_2)]^2 - w(\bar{s})^2 \right] - w(\bar{s})^2 - w(\bar{s}) [w(\bar{s}_1) + w(\bar{s}_2)] \left[ [w(\bar{s}_1) + w(\bar{s}_2)] - w(\bar{s}) \right] \right) - 2\lambda(1 - \lambda\alpha) \left( (\gamma_g - \alpha) \left[ [w(\bar{s}_1) + w(\bar{s}_2)] - w(\bar{s}) \right] - w(\bar{s}) \right).$$

Notice that the welfare loss specified in Equation (12) is continuous in  $\lambda \in (0, 1)$ . Then, for  $\lambda \rightarrow 0$ , we obtain that the the sign of the difference  $L(g) - L(g')$  coincides with the sign of  $[w(\bar{s}_1) + w(\bar{s}_2)] - w(\bar{s}) - 1$ . For very low levels of complementarity in actions ( $\lambda \rightarrow 0$ ), it then follows that  $L(g') < L(g)$  if and only if  $[w(\bar{s}_1) + w(\bar{s}_2)] - w(\bar{s}) > 1$ , as stated.  $\square$

## References

1. Keynes, J.M. *The General Theory of Employment, Interest, and Money*; MacMillan: London, UK, 1936.
2. Morris, S.; Shin, H. The Social Value of Public Information. *Am. Econ. Rev.* **2002**, *92*, 1521–1534.
3. Angeletos, G.M.; Pavan, A. Transparency of Information and Coordination in Economies with Investment Complementarities. *Am. Econ. Rev.* **2004**, *94*, 91–98.
4. Angeletos, G.M.; Pavan, A. Efficient Use of Information and Social Value of Information. *Econometrica* **2007**, *75*, 1103–1142.
5. Allen, F.; Morris, S.; Shin, H.S. Beauty Contest and Iterated Expectations in Asset Markets. *Rev. Financ. Stud.* **2006**, *19*, 720–752.
6. Hellwig, C.; Veldkamp, L. Knowing What Others Know: Coordination Motives in Information Acquisition. *Rev. Econ. Stud.* **2009**, *76*, 223–251.
7. Dewan, T.; Myatt, D.P. The Qualities of Leadership: Direction, Communication, and Obfuscation. *Am. Political Sci. Rev.* **2008**, *102*, 351–368.
8. Jimenez-Martinez, A. Information Acquisition Interactions in Two-Player Quadratic Games. *Int. J. Game Theory* **2014**, *43*, 455–485.
9. Calvó-Armengol, A.; de Martí-Beltran, J. Information Gathering in Organizations: Equilibrium, Welfare, and Optimal Network Structure. *J. Eur. Econ. Assoc.* **2009**, *7*, 116–161.
10. Sherman, J.; Morrison, W.J. Adjustment of an Inverse Matrix Corresponding to a Change in One Element of a Given Matrix. *Ann. Math. Stat.* **1950**, *21*, 124–127.
11. Henderson, H.V.; Searle, S.R. On Deriving the Inverse of a Sum of Matrices. *Siam Rev.* **1981**, *23*, 53–60.



© 2017 by the author; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).