Perspectives, Opinions, and Information Flows

Rajiv Sethi*  
Muhamet Yildiz†

January 19, 2013

Abstract

Consider a group of individuals with heterogenous prior beliefs or perspectives about a sequence of states. Individuals receive informative signals about each state, on the basis of which they form posterior beliefs or opinions. In each period, each individual can choose to observe the opinion of one other, but perspectives and signals are unobservable. The heterogeneity and unobservability of perspectives introduces a trade-off between targets who are well-informed (in the sense that their signals are precise) and those who are well-understood (in the sense that the observer’s beliefs about the target’s perspective is precise). Observing an opinion provides information about both the current period state and about the target’s perspective. Hence observed individuals become better-understood over time, although the degree to which this occurs depends on the extent to which both observer and observed are well-informed at the time of observation. This allows for both history dependence and symmetry breaking in the formation of links. We identify conditions under which history independence prevails in the long run. When this condition fails to hold, a number of network structures can arise. We focus on three of these: opinion leadership, information segregation, and static networks.

*Department of Economics, Barnard College, Columbia University and the Santa Fe Institute.
†Department of Economics, MIT.
1 Introduction

Among the many important roles played by networks in social and economic life is that of carriers of information. In fact, the formation of links among otherwise disconnected individuals is often motivated precisely by this function. Subscriptions to blog and twitter feeds have this character, as do more traditional activities such as the monitoring of radio and television broadcasts or the reading of books and newspapers.

Since an individual’s capacity to receive and process information is limited, it is necessary to make choices regarding the set of opinions that one chooses to observe at any time. These choices are further complicated by the fact that the observation of another’s opinion gives rise to an inference problem, since an opinion is based partly on one’s information and partly on one’s prior beliefs. If the prior beliefs of all individuals were mutually known, this problem would have a trivial solution, and each individual would simply choose to observe those who happen to have the most precise information.

But prior beliefs are not generally observable, and this gives rise to a trade-off between individuals who are well-informed and those who are well-understood. An individual is well-informed if the precision of her information is high. She is well-understood by an observer if the precision of the observer’s beliefs about her prior is high. That is, a person is well-understood if her beliefs reveal her information with a high degree of precision. Clearly, there is little value in observing a well-informed person who is very poorly understood in this sense, since her beliefs will provide a very noisy signals about her information to the observer. Hence individuals will not always observe those who are the best-informed, especially if informational differences across individuals are small relative to differences in the degree to which they are well-understood.

This has some interesting dynamic implications, since the observation of an opinion not only provides a signal about the information that gave rise to it, but also reveals something about the observed individual’s prior belief. In other words, the process of being observed makes one better understood. This process can give rise to unusual and interesting patterns of linkages over time, even of all individuals are identical to begin with. It is these effects with which the present paper is concerned, with particular focus on three phenomena: opinion leadership, information segregation, and static networks.

The model we explore has the following structure. There is a finite set of individuals and sequence of periods. Corresponding to each period is an unobserved state. Individuals all believe that the states are independently and identically distributed, but differ with
respect to their prior beliefs about the distribution from which these states are drawn. These beliefs, which we call perspectives, are themselves unobservable, although each individual holds beliefs about the perspectives of others. In each period, each individual receives a signal that is informative about the current state; the precision of this signal is the individual’s expertise in that period. Levels of expertise are independently and identically distributed across individuals and periods, and their realized values are public information. Individuals update their beliefs on the basis of their signals, resulting in posterior beliefs we call opinions. They then choose a target individual whose opinion may also be observed, and make this choice by selecting the target whose opinion is most informative to them.

The observation of an opinion has two effects. First, it affects the observer’s belief about the current period state and allows her to take a more informed action. Second, the observer’s belief about the target’s perspective itself becomes more precise. Hence there will be a tendency to link to previously observed targets even when they are not the best-informed in the current period. But, importantly, the level of attachment to a previously observed target depends on the expertise realizations of both observer and observed in the period in which the observation occurred. Specifically, better informed observers learn more about the perspectives of their targets since they have more precise beliefs about the signal that the target is likely to have received. But holding constant one’s own expertise, one learns more about the perspective of a poorly informed target, since the opinion of such a target will be heavily weighted to their prior rather than their signal.

This effect implies symmetry breaking over time: two observers who select the same target initially will develop different levels of attachment to that target. Hence they make different choices in a subsequent period, despite the fact that all expertise realizations are public information and a given individual’s expertise is common to all observers. Several interesting linkage patterns can arise over time as a result of this effect. Opinion leadership, where a small subset of individuals is observed with high frequency, is clearly one possible outcome. But information segregation, where the set of individuals is partitioned into groups with no observation across groups is also possible. Moreover, any static network of links can be a limit of the process of link formation. We identify conditions on the key parameters of the model—the degree of initial uncertainty about the perspectives of others, and the distribution from which expertise is drawn—under which each of these patterns can arise.

A key idea underlying our work is that there is some aspect of cognition that is variable across individuals, stable over time, and affects the manner in which information pertaining to a broad range of issues is filtered. Differences in political ideology or cultural orientation can give rise to such stable variability in the manner in which information is interpreted.
This is a feature of the cultural theory of perception (Douglas and Wildavsky, 1982) and the related notion of identity-protective cognition (Kahan et al., 2007).

A striking example of stable variability in perspectives may be found in perceptions of changes in local weather patterns. Goebbert et al. (2012) surveyed a large sample of respondents across a variety of locations, and collected information on perceived changes in local temperatures and precipitation, together with geographic and demographic information, self-declared political ideology (strong liberal to strong conservative on a seven point scale), and answers to questions that allowed for a three-way classification of respondents by worldview (egalitarian, hierarchist and individualist). They found that politically conservative and individualist respondents were significantly less likely to have perceived recent increases in local temperatures relative to their more liberal and egalitarian counterparts. Since all respondents at a given location were exposed to precisely the same local temperatures, such differences in perception cannot plausibly be attributed solely to differences in information.

Along similar lines, recent survey data on beliefs about the religion and birthplace of Barack Obama suggests considerable variability by race, faith, and political persuasion (Thrush 2009, Pew Research Center 2008). Racial identity and political ideology also correlate quite strongly with beliefs about the origins of the AIDS virus, the introduction of drugs into inner cities, and the extent of discrimination in daily life (Crocker et al. 1999, CNN/Opinion Research 2008). Beliefs about the accuracy of election polling data and even official unemployment statistics have recently shown strong political cleavages (Plambeck 2012, Voorhees 2012). Since much of the hard evidence pertaining to these issues is in the public domain, it is unlikely that such stark belief differences arise from informational differences alone.

Our analysis is connected to several stands of literature on heterogeneous priors, observational learning, and the formation of networks. Strategic communication with observable heterogeneous priors has previously been considered by Banerjee and Somanathan (2001), Che and Kartik (2009), and Van den Steen (2010) amongst others. Dixit and Weibull (2007) have shown that the beliefs of individuals with heterogeneous priors can diverge further upon observation of a public signal, and Acemoglu et al. (2009) that they can fail to converge even after an infinite sequence of signals. In our own previous work, we have considered truthful communication with unobservable priors, but with a single state and public belief announcements (Sethi and Yildiz, 2012). Communication across an endogenous network with unobserved heterogeneity in prior beliefs and a sequence of states has not previously been explored as far as we are aware.
The literature on observational learning is vast. A central concern in this body of work is the manner in which network structure affects the distribution of beliefs and actions in the long run, and whether these actions are optimal conditional on the state of the world. Bala and Goyal (1998) explore a canonical model of this kind with the following features. There is an unobserved state that affects the value of various actions, and individuals have (possibly heterogeneous) prior beliefs regarding this state. There is a sequence of periods in each of which individuals take actions, observe outcomes, and update beliefs. They may also observe the actions and payoffs of others. A directed graph represents the pattern of observability, where individuals can observe those to whom they are directly connected, and perhaps also those to whom they are indirectly connected via some path in the network. Beliefs are updated in each period based on the set of actions and outcomes observed, and the action chosen in each period is optimal conditional on the current beliefs. The authors show that convergence to optimal actions may fail to occur if there is a subset of individuals observed by all others and if prior beliefs are not sufficiently dispersed to begin with. Since updating is based on observed actions and outcomes alone, there is no communication of beliefs and no attempt to form beliefs about the beliefs of others. Furthermore, in contrast with the model developed here, the targets of observation are not themselves objects of choice, since the network is exogenously given.

The network formation literature is also substantial and growing rapidly. Two especially relevant contributions from the perspective of our work are by Galeotti and Goyal (2010) and Acemoglu et al. (2011a). Galeotti and Goyal (2010) develop a model to account for the law of the few, which refers to the empirical finding that the population share of individuals who invest in the direct acquisition of information is small relative to the share of those who acquire it indirectly via observation of others, despite minor differences in attributes across the two groups. All individuals are ex-ante identical in their model and can choose to acquire information directly, or can choose to form costly links in order to obtain information that others have paid to acquire. All strict Nash equilibria in their baseline model have a core-periphery structure, with all individuals observing those in the core and none linking to those in the periphery. Hence all equilibria are characterized by opinion leadership: those in

---

1See Goyal (2010) for a survey. Early and influential contributions include Banerjee (1992), Bikhchandani et al. (1992), and Smith and Sorensen (2000) in the context of sequential choice. For learning in networks see Bala and Goyal (1998), Gale and Kariv (2003), DeMarzo et al. (2003), Golub and Jackson (2010), Acemoglu et al. (2011b), Chatterjee and Xu (2004), and Jadbabaie et al. (2012).

2Bloch and Dutta (2010) and Jackson (2010) provide comprehensive surveys. Key early contributions include Jackson and Wolinsky (1996) and Bala and Goyal (2000); see also Watts (2001), Bramoulle and Kranton (2007), Bloch et al. (2008) and, most relevant to the present paper, Calvó-Armengol et al. (2011). We follow Bala and Goyal in focusing on the noncooperative formation of directed links.
the core acquire information directly and this is then accessed by all others in the population. Since there are no problems with the interpretation of opinions in their framework, and hence no variation in the extent to which different individuals are well-understood, information segregation cannot arise.

Acemoglu et al. (2011a) also consider communication in an endogenous network. Individuals can observe the information of anyone to whom they are linked either directly or indirectly via a path, but observing more distant individuals requires waiting longer before an action is taken. Holding constant the network, the key tradeoff in their model is between reduced delay and a more informed decision. They show that dispersed information is most effectively aggregated if the network has a hub and spoke structure with some individuals gathering information from numerous others and transmitting it either directly or via neighbors to large groups. This structure is then shown to emerge endogenously when costly links are chosen prior to communication, provided that certain conditions are satisfied. One of these conditions is that friendship cliques, defined as sets of individuals who can observe each other at zero cost, not be too large. Members of large cliques are well-informed, have a low marginal value of information, and will not form costly links to those outside the clique. Hence both opinion leadership and information segregation are possible equilibrium outcomes in their model, though the mechanisms giving rise to these are clearly distinct from those explored here.

2 The Model

Consider a population $N = \{1, \ldots, n\}$, and a sequence of periods. In each period $t$, there is a state $\theta_t \in \mathbb{R}$ that individuals cannot observe but about which they form and update beliefs. Each individual $i$ holds an idiosyncratic prior belief about the distribution from which these states are drawn, given by

$$\theta_t \sim_i N(\mu_i, 1).$$

That is, according to player $i$, the sequence of states $\theta_0, \theta_1, \ldots$ are independently, identically, and Normally distributed with mean $\mu_i$ and variance 1. We shall refer to prior mean $\mu_i$ as the perspective of individual $i$. An individual’s perspective is not directly observable by any other individual, but it is commonly known that the $n$ perspectives are independently and identically distributed according to

$$\mu_i \sim N(\mu, 1/v_0)$$
for some real numbers $\mu_1, \ldots, \mu_n$ and $\nu_0 > 0$. This describes the beliefs held by individuals about each others’ perspectives prior to the receipt of any relevant information. Note that the precision in beliefs about perspectives are symmetric in the initial period, since $\nu_0$ is common to all. This symmetry is broken as individuals learn about perspectives over time, and the revision of these beliefs plays a key role in the analysis to follow.

In each period $t$, each individual $i$ privately observes an informative signal

$$x_{it} = \theta_t + \varepsilon_{it},$$

where $\varepsilon_{it} \sim N(0, 1/\pi_{it})$. The signal precisions $\pi_{it}$ capture the degree to which any given individual $i$ is well-informed about the state in period $t$. We shall refer to $\pi_{it}$ as the expertise of individual $i$ regarding the period $t$ state, and assume that these levels of expertise are public information. Levels of expertise $\pi_{it}$ are independently and identically distributed across individuals and periods, in accordance with an absolutely continuous distribution function $F$ having support $[a, b]$, where $0 < a < b < \infty$. That is, no individual is ever perfectly informed of the state, but all signals carry at least some information.\(^3\)

**Remark 1.** Since priors are heterogenous, each individual has his own subjective beliefs. We use the subscript $i$ to denote the individual whose belief is being considered. For example, we write $\theta_t \sim_i N(\mu_i, 1)$ to indicate that $\theta_t$ is normally distributed with mean $\mu_i$ according to $i$. When all individuals share a belief, we drop the subscript. For example, $\varepsilon_{it} \sim N(0, 1/\pi_{it})$ means that all individuals agree that the noise in $x_{it}$ is normally distributed with mean 0 and variance $1/\pi_{it}$. While an individual $j$ does not infer anything about $\theta_t$ from the value $\mu_i$, $j$ does update her belief about $\theta_t$ upon receiving information about $x_{it}$. For a more extensive discussion of belief revision with incomplete information and unobservable, heterogenous priors, see Sethi and Yildiz (2012), where we study repeated communication about a single state among a group of individuals with equal levels of expertise.

Having observed the signal $x_{it}$ in period $t$, individual $i$ updates her belief about the state in conformity with Bayes’ rule.\(^4\) This results in the following posterior belief for $i$:

$$\theta_t \sim_i N\left(y_{it}, \frac{1}{1 + \pi_{it}}\right),$$

\(^3\)Since $\pi_{it}$ is observable, myopic individuals need not consider the distribution from which $\pi_{it}$ is drawn. Nevertheless, this distribution affects the pattern of linkages that emerges in the long run.

\(^4\)Specifically, given a prior $\theta \sim N(\mu, 1/v)$ and signal $s = \theta + \varepsilon$ with $\varepsilon \sim N(0, 1/r)$, the posterior is $\theta \sim N(y, 1/w)$ where

$$y = E[\theta|s] = \frac{v}{v + r} \mu + \frac{r}{v + r} s$$

and $w = v + r$.\(^7\)
where

\[ y_{it} = \frac{1}{1 + \pi_{it}} \mu_i + \frac{\pi_{it}}{1 + \pi_{it}} x_{it} \]  

(2)

is the expected value of \( \theta_t \) according to \( i \). We refer to \( y_{it} \) as individual \( i \)'s opinion at time \( t \).

A key concern in this paper is the process by which individuals choose targets whose opinions are then observed. We model this choice as follows. In each period \( t \), each individual \( i \) chooses one other individual, denoted \( j_{it} \in N \), and observes her opinion \( y_{j_{it}t} \) about the current state \( \theta_t \). This information is useful because \( i \) then chooses an action \( \hat{\theta}_{it} \in \mathbb{R} \) in order to minimize

\[ E[(\hat{\theta}_{it} - \theta_t)^2]. \]  

(3)

This implies that individuals always prefer to observe a more informative signal to a less informative one. We specify the actions and the payoffs only for the sake of concreteness; our analysis is valid so long as this desire to seek out the most informative signal is assumed. (In some applications this desire may be present even if no action is to be taken.) The timeline of events at each period \( t \) is as follows:

1. The levels \((\pi_{1t}, \ldots, \pi_{nt})\) of expertise are realized and publicly observed.
2. Each \( i \) chooses some advisor \( j_{it} \in N \setminus \{i\} \).
3. Each \( i \) observes his own noisy signal \( x_{it} \) and forms his opinion \( y_{it} \).
4. Each \( i \) observes the opinion \( y_{j_{it}t} \) of his advisor.
5. Each \( i \) takes an action \( \hat{\theta}_{it} \).

It is convenient to introduce the variable \( l_{ijt}^t \) which takes the value 1 if \( j_{it} = j \) and zero otherwise. That is, \( l_{ijt}^t \) indicates whether or not \( i \) links to \( j \) in period \( t \) and the \( n \times n \) matrix \( L^t := [l_{ijt}^t] \) defines a directed graph or network that describes who listens to whom. Note that information flows in the reverse direction of the graph. We are interested in the properties of the sequence of networks generated by this process of link formation.

We assume that individuals are myopic, do not observe the actions of others, and do not observe the realization of the state (observability of the past advisors of others will turn out to be irrelevant). Hence the payoff is the expectation (3) itself. While these are clearly restrictive assumptions, the desire to make a good decisions even when the state realization is unobserved is quite common. For instance, one might wish to vote for the least corrupt political candidate, or donate to the charity with the greatest social impact, or support
legislation regarding climate change that results in the greatest benefits per unit cost. We actively seek information in order to meet these goals, and act upon our expectations, but never know for certain whether our beliefs were accurate ex-post.

Remark 2. Even though the states, signals and expertise levels are all distributed independently across individuals and time, the inference problems at any two dates \( t \) and \( t' \) are related. This is because each individual’s ex-ante expectation of \( \theta_t \) and \( \theta_{t'} \) are the same; this expectation is what we call the individual’s perspective. As we show below, any information about the perspective \( \mu_j \) of an individual \( j \) is useful in interpreting \( j \)’s opinion \( y_{jt} \), and this opinion in turn is informative about \( j \)’s perspective. Consequently the choice of advisor at date \( t \) affects the choice of the advisor at any later date \( t' \). In particular, the initial symmetry is broken after individuals choose their first advisor, leading to potentially highly asymmetric outcomes.

3 Evolution of Beliefs and Information Networks

We now describe the criterion on the basis of which a given individual \( i \) selects a target \( j \) whose opinion \( y_{jt} \) is to be observed, and what \( i \) learns about the state \( \theta_t \) and \( j \)’s perspective \( \mu_j \) as a result of this observation. This determines the process for the evolution of beliefs and the network of information flows.

Given the hypothesis that the \( n \) perspectives are each independently drawn from a normal distribution, posterior beliefs held by one individual about the perspectives of any another will continue to be normally distributed throughout the process of belief revision. Write \( v_{ij}^t \) for the precision of the distribution of \( \mu_j \) according to \( i \) at beginning of \( t \). Initially, these precisions are identical: for all \( i \neq j \),

\[
v_{ij}^1 = v_0. \tag{4}
\]

The precisions \( v_{ij}^t \) in subsequent periods depend on the history of realized expertise levels \( (\pi^1, \ldots, \pi^{t-1}) \) and information networks \( (L^1, \ldots, L^{t-1}) \). These precisions \( v_{ij}^t \) of beliefs about the perspectives of others are central to our analyses; the expected value of an individual’s perspective is irrelevant as far as the target choice decision is concerned. What matters is how well a potential target is understood, not how far their perspective deviates from that of the observer.
3.1 Interpretation of Opinions and Selection of Advisors

Suppose that an individual $i$ has chosen to observe the opinion $y_{jt}$ of individual $j$, where

$$y_{jt} = \frac{1}{1 + \pi_{jt}} \mu_j + \frac{\pi_{jt}}{1 + \pi_{jt}} x_{jt}$$

by (2). Since $x_{jt} = \theta_t + \varepsilon_{jt}$, this observation provides the following noisy signal regarding $\theta_t$:

$$\frac{1 + \pi_{jt}}{\pi_{jt}} y_{jt} = \theta_t + \varepsilon_{jt} + \frac{1}{\pi_{jt}} \mu_j.$$

The signal is noisy in two respects. First, the information $x_{jt}$ of $j$ is itself noisy, with signal variance $\varepsilon_{jt}$. Furthermore, since the opinion $y_{jt}$ depends on $j$’s unobservable perspective $\mu_j$, the signal observed by $i$ has an additional source of noise, reflected in the term $\mu_j/\pi_{jt}$.

Taken together, the variance of the signal observed by $i$ is

$$\gamma(\pi_{jt}, v_{ij}^t) \equiv \frac{1}{\pi_{jt}} + \frac{1}{\pi_{jt}^2} v_{ij}^t.$$  \hspace{1cm} (5)

Here, the first component $1/\pi_{jt}$ comes directly from the noise in the information of $j$, and is simply the variance of $\varepsilon_{jt}$. It decreases as $j$ becomes better informed. The second component, $1/(\pi_{jt}^2 v_{ij}^t)$, comes from the uncertainty $i$ faces regarding the perspective $\mu_j$ of $j$, and corresponds to the variance of $\mu_j/\pi_{jt}$ (where $\pi_{jt}$ is public information and hence has zero variance). This component decreases as $i$ becomes better acquainted with the perspective $\mu_j$, that is, as $j$ becomes better understood by $i$.

The cost $\gamma$ reveals that in choosing an advisor $j$, an individual $i$ has to trade-off the noise $1/\pi_{jt}$ in the information of $j$ against the noise $1/(\pi_{jt}^2 v_{ij}^t)$ in $i$’s understanding of $j$’s perspective, normalized by the level of $j$’s expertise. The trade-off is between advisors who are well-informed and those who are well-understood.

Since $i$ seeks to observe the most informative opinion, she chooses to observe an individual for whom the variance $\gamma$ is lowest. Ties arise with zero probability but for completeness we assume that they are broken in favor of the individual with the lowest label. That is,

$$j_{it} = \min_{j \neq i} \left\{ \arg\min \gamma(\pi_{jt}, v_{ij}^t) \right\}.$$  \hspace{1cm} (6)

Note that $j_{it}$ and hence $L^t$ have two determinants: the current expertise levels $\pi_{jt}$ and the precision $v_{ij}^t$ of individuals’ beliefs regarding the perspectives of others. The first determinant $\pi_{jt}$ is exogenously given and stochastically independent across individuals and times. In contrast, the second component $v_{ij}^t$ is endogenous and depends on the sequence of prior advisor choices $(L^1, \ldots, L^{t-1})$, which in turn depends on previously realizes levels of expertise.
3.2 Evolution of Beliefs

We now describe the manner in which the beliefs $v_{tij}$ are revised over time. In particular we show that the belief of an observer about the perspective of her target becomes more precise once the opinion of the latter has been observed, and that the strength of this effect depends systematically on the realized expertise levels of both observer and observed.

Suppose that $j_{it} = j$, so $i$ observes $y_{jt}$. Recall that $j$ has previously observed $x_{jt}$ and updated her belief about the period $t$ state in accordance with (1-2). Hence observation of $y_{jt}$ by $i$ provides the following signal about $\mu_j$:

$$(1 + \pi_{jt})y_{jt} = \mu_j + \pi_{jt}\theta_t + \pi_{jt}\varepsilon_{jt}.$$  

The variance of the noise in this signal is

$$\pi_{jt}^2 \left( \frac{1}{1 + \pi_{it}} + \frac{1}{\pi_{jt}} \right),$$

and the precision of the signal is accordingly $\delta(\pi_{it}, \pi_{jt})$, defined as

$$\delta(\pi_{it}, \pi_{jt}) = \frac{1 + \pi_{it}}{\pi_{jt}(1 + \pi_{it} + \pi_{jt})}. \quad (7)$$

Hence, using the formula in footnote 4, we obtain

$$v_{t+1}^{ij} = \begin{cases} v_t^{ij} + \delta(\pi_{it}, \pi_{jt}) & \text{if } j_{it} = j \\ v_t^{ij} & \text{if } j_{it} \neq j, \end{cases} \quad (8)$$

where we are using the fact that if $j_{it} \neq j$, then $i$ receives no signal of $j$’s perspective, and so her belief about $\mu_j$ remains unchanged. This leads to the following closed-form solution:

$$v_{t+1}^{ij} = v_0 + \sum_{s=1}^{t} \delta(\pi_{is}, \pi_{js})l_{ij}^s. \quad (9)$$

**Remark 3.** This derivation assumes that individuals do not learn from the adviser choices of others, as described in $L^t$. If fact, under our assumptions, there is no additional information contained in these choices because $i$ can compute $L^t$ using publicly available data even before $L^t$ has been observed.\(^5\) This simplifies the analysis dramatically, and is due to the linear

---

\(^5\)One can prove this inductively as follows. At $t = 1$, $i$ can compute $L^t$ from (6) using $(\pi_{1t}, \ldots, \pi_{nt})$ and $v_0$ without observing $L^t$. Suppose now that this is indeed the case for all $t' < t$ for some $t$, i.e., $L^{t'}$ does not provide any additional information about $\mu_j$. Then all beliefs about perspectives are given by (8) up to date $t$. One can see from this formula that each $v_{kl}^{t+1}$ is a known function of past expertise levels $(\pi_{k1t'}, \ldots, \pi_{kn't'})_{t' < t}$, all of which are publicly observable. That is, $i$ knows $v_{kj}^{t+1}$ for all distinct $k, j \in N$. Using $(\pi_{1t}, \ldots, \pi_{nt})$ and these values, she can then compute $j_{kt}$ from (6) without observing $L^t$.\(^5\)
formula in Footnote 4 for normal variables. In a more general model, \( i \) may be able to obtain useful information by observing \( L \). For example, without linearity, \( v_k^{t+1} \) could depend on \( y_{jt} \) for some \( k \) with \( j_{kt} = j \). Since \( y_{jt} \) provides information about \( \mu_j \), and \( v_k^{t+1} \) affects \( j_{kt'} \) for \( t' \geq t + 1 \), one could then infer useful information about \( \mu_j \) from \( j_{kt'} \) for such \( t' \). The formula (8) would not be true for \( t' \) in that case, possibly allowing for other forms of inference at later dates.

**Remark 4.** By the argument in the previous remark, assumptions about the observability of the information network \( L \) are irrelevant for our analysis. However, assumptions about the observability of the state \( \theta_t \) and the actions \( \hat{\theta}_{kt} \) of others (including the actions of one’s advisor, which incorporate information from her own advisor) are clearly relevant.

Each time \( i \) observes \( j \), her beliefs about \( j \)’s perspective become more precise. But, by (7), the increase \( \delta(\pi_{it}, \pi_{jt}) \) in precision depends on the specific realizations of \( \pi_{it} \) and \( \pi_{jt} \) in the period of observation, in accordance with the following.

**Observation 1.** \( \delta(\pi_{it}, \pi_{jt}) \) is strictly increasing \( \pi_{it} \) and strictly decreasing \( \pi_{jt} \). Hence,

\[
\delta \leq \delta(\pi_{it}, \pi_{jt}) \leq \tilde{\delta}
\]

where \( \delta \equiv \delta(a, b) > 0 \) and \( \tilde{\delta} \equiv \delta(b, a) \)

In particular, if \( i \) happens to observe \( j \) during a period in which \( j \) is very precisely informed about the state, then \( i \) learns very little about \( j \)’s perspective. This is because \( j \)’s opinion largely reflects the signal and is therefore relatively uninformative about \( j \)’s prior. If \( i \) is very well informed when observing \( j \), the opposite effect arises and \( i \) learns a great deal about \( j \)’s perspective. Having good information about the state also means that \( i \) has good information about \( j \)’s signal, and is therefore better able to infer \( j \)’s perspective based on the observed opinion. Finally, there is a positive lower bound \( \delta \) on the amount of increase in precision, making beliefs about observed individuals more and more precise as time passes.

Given the precisions \( v_{ij}^t \) at the start of period \( t \), and the realizations of the levels of expertise \( \pi_{it} \), the links chosen by each individual in period \( t \) are given by (6). This then determines the precisions \( v_{ij}^{t+1} \) at the start of the subsequent period in accordance with (8), with initial precisions given by (4). For completeness, we set \( v_{ii}^t = 0 \) for all individuals \( i \) and all periods \( t \). This defines a Markov process, where the sample space is the set of nonnegative \( n \times n \) matrices and the period \( t \) realization is \( V^t := [v_{ij}^t] \).

For any period \( t \), let \( h_t := \{v_{ij}^1, ..., v_{ij}^t\} \) denote the history of beliefs (regarding perspectives) up to the start of period \( t \). Any such history induces a probability distribution over
networks, with the period $t$ network being determined by the realized values of $\pi_{it}$. It also induces a distribution over the next period beliefs $v_{ij}^{t+1}$. It is the long run properties of this sequence of networks and beliefs that we wish to characterize.

### 3.3 Network Dynamics

Recall from (6) that at any given date $t$, each individual $i$ chooses an advisor $j_{it}$ with the goal of minimizing the perceived variance $\gamma(\pi_{jt}, v_{ij}^t)$. At the start of this process, since the precisions $v_{ij}^1$ are all equal, the expertise levels $\pi_{jt}$ are the only determinants of this choice. Hence the criterion (6) reduces to

$$j_{i1} = \min \left\{ \arg \max_{j \neq i} \pi_{j1} \right\}.$$ 

That is, the best informed individual in the initial period is linked to by all others, and herself links to the second-best informed.

This pattern of information flows need not hold in subsequent periods. By Observation 1, individual beliefs about the perspectives of their past advisors become strictly more precise over time. Since $\gamma$ is strictly decreasing in such precision, an individual may continue to observe a past advisor even if the latter is no longer the best informed. And since better informed individuals learn more about the perspectives of their advisers, they may stick to past advisors with greater likelihood than poorly informed individuals, adding another layer of asymmetry.

For an illustration, consider the simple case of $n = 4$, and suppose (without loss of generality) that $\pi_{1t} > \pi_{2t} > \pi_{4t} > \pi_{3t}$ at $t = 1$. Then individual 1 links to 2 (i.e. $j_{1t} = 2$) and all the others link to 1 (i.e. $j_{2t} = j_{3t} = j_{4t} = 1$). Individuals 2, 3, and 4 all learn something about the perspective of individual 1. The precisions $v_{11}^2$ of their beliefs about $\mu_1$ at the start of the next period are all at least $v_0 + \delta$, while the precisions of their beliefs about the perspectives of other individuals remain at $v_0$. Moreover, they update their beliefs to different degrees, with those who are better informed about the state ending up with more precise beliefs about 1’s perspective: $v_{21}^2 > v_{41}^2 > v_{31}^2 \geq v_0 + \delta$.

Now consider the second period, and suppose that this time $\pi_{2t} > \pi_{1t} > \pi_{4t} > \pi_{3t}$. There is clearly no change in the links chosen by individuals 1 and 2, who remain the two who are best informed. On the other hand, there is an open set of expertise realizations for which 3 and 4 remain linked to 1 despite the fact that 2 is now better informed. In Figure 1, this event ($j_{32} = j_{42} = 1$) occurs for expertise realizations between the shaded region and the 45-degree
In this region, while 2 is better informed than 1 ($\pi_2t > \pi_1t$), the difference between their expertise levels is not large enough to overcome the stronger attachment of individuals 3 and 4 to their common past adviser ($\nu_{21}^2 > \nu_{22}^2$ for $i \in \{3, 4\}$). Below the shaded region, the difference in expertise levels between 1 and 2 is large enough to induce both individuals 3 and 4 to switch to the best informed target in the second period ($j_{32} = j_{42} = 2$).

Figure 1: Agents 3 and 4 choose different targets when variances lie in the shaded region.

Within the shaded region, however, symmetry is broken and individuals 3 and 4 choose different targets: 3 switches to the best informed individual ($j_{3t} = 2$) while 4 remains linked to her previous advisor ($j_{4t} = 1$). In this region, the difference between the expertise levels of 1 and 2 is large enough to overcome the preference of 3 towards 1, but not large enough to overcome the stronger preference of individual 4, who was more precisely informed of the state in the initial period, and hence learned more about the perspective of her advisor.

A particular set of realizations that generates this effect is shown in Figure 2, where a solid line indicates that links are formed in both directions and a dashed line indicates a different realization.

---

The figure has variances of $\varepsilon_1$ and $\varepsilon_2$ on the horizontal and vertical axes respectively, and is based on the specification $\nu_0 = 1, \nu_{31} = 2,$ and $\nu_{41} = 4$. Since 2 is assumed to be better informed than 1 in period 2, all expertise realizations must lie below the 45-degree line.
single link in one direction. Nodes (corresponding to individuals) are numbered in increasing order anti-clockwise, starting from the top. Nodes 1 and 2 link to each other in both periods. Nodes 3 and 4 link to node 1 (the best informed) in the first period. In the second period node 3 switches to node 2, who is now the best informed, but node 4 continues to observe node 1. This is because the perspective of 1 is better known to 4 than to 3, since 4 was better informed than 3 about the state in the initial period.

\[ \text{Figure 2: Asymmetric effects of first period observations on second period links.} \]

This example illustrates the trade-off between being well informed and being well understood. It is clear that this can prevent the formation of networks in which all individuals link to the best informed, and can give rise to history dependence. One of the key questions of interest in this paper is whether this is a temporary effect, or whether it can arise even in the long run.

In order to explore this question, we introduce some notation. We say that the link \( ij \) is active in period \( t \) if \( l_{ij}^t = 1 \). Given any history \( h_t \), we say that the link \( ij \) is broken in period \( t \) if, conditional on this history, the probability of the link being active in period \( t \) is zero. That is, the link \( ij \) is broken in period \( t \) conditional on history \( h_t \) if

\[ \Pr(l_{ij}^t = 1 \mid h_t) = 0. \]

If a link is broken in period \( t \) we write \( b_{ij}^t = 1 \). It is easily verified that if a link is broken in period \( t \) then it is broken in all subsequent periods.\(^7\) Finally, we say that a link \( ij \) is free in period \( t \) conditional on history \( h_t \) if the probability that it will be broken in this or any

\[ \text{This follows from the fact that the process } \{v_{ij}^t\} \text{ is non-decreasing, and } v_{ij} \text{ increases in period } t \text{ if and only if } l_{ij} = 1. \]
subsequent period is zero. That is, link \( ij \) is free in period \( t \) if

\[
\Pr(b_{ij}^{t+s} = 1 \mid h_t) = 0
\]

for all non-negative integers \( s \). If a link is free at time \( t \), there is a positive probability that it will be active in the current period as well as in each subsequent period.

A particular realization of this process for three individuals over the first six periods is shown in Figure 3. As before, nodes are numbered in increasing order anti-clockwise starting from the top, solid lines indicate links in both directions, and dashed lines indicate a link in one direction. Node 2 is the best informed in the initial period, while node 1 is the second best informed. Hence 2 links to 1 and the others link to 2. No links are broken, but links in both directions between 1 and 2 become free, as does the link from 3 to 2. The pattern of active links is is repeated in the second period, even though 2 need not be the best informed at this date. In the third period 1 switches to 3, who is now better informed than 2. The link from 3 to 1 is broken in the fourth period, and the link from 2 to 3 in the fifth. At this point the network is fully resolved (all links are either broken or free), and there are only two possibilities remaining feasible. In particular, links 21 and 32 are always active after this point, while 12 and 13 are sometimes active.

---

**Figure 3:** Active (top row), broken (middle) and free (bottom) links over six periods.
Remark 5. In some models of information aggregation over networks, individuals may end up selecting different actions in the long run, but their expected payoffs need to be same (see, for example, Bala and Goyal, 1998). The example in the previous paragraph shows that this need not be the case in our model. Conditional on the realized history above, the expected payoff of 1 is higher than the expected payoff of the other two in the long run. Indeed, 1 eventually observes the best informed of the two potential advisers with vanishing uncertainty about their perspectives, while the other two individuals repeatedly observe a single target.

We next identify conditions under which a link breaks or becomes free. Define

$$\varphi = \frac{a}{b(b - a)}.$$ 

and note that this threshold precision satisfies the indifference condition

$$\gamma(a, \infty) = \gamma(b, \varphi)$$

between a minimally informed individual whose perspective is known and a maximally informed individual whose perspective is uncertain with precision \(\varphi\). Define also the function \(\beta : (0, \varphi) \to \mathbb{R}^+\), by setting

$$\beta(v) = \frac{b^2}{a^2} \left( \frac{1}{v} - \frac{1}{\varphi} \right)^{-1}.$$ 

This function satisfies the indifference condition

$$\gamma(a, \beta(v)) = \gamma(b, v)$$

between a maximally informed individual whose perspective is uncertain with precision \(v\) and a minimally informed individual whose perspective is uncertain with precision \(\beta(v)\).

In our analysis, we shall ignore histories that result in ties and arise with zero probability. Accordingly, define

$$\mathcal{V} = \left\{ (v_{ij})_{i \in N, j \in N \setminus \{i\}} \mid v_{ij} \neq \beta(v_{ik}) \text{ and } v_{ij} \neq \varphi \text{ for all distinct } i, j, k \in N \right\}$$

and

$$H = \left\{ h_t \mid v^t(h_t) \in \mathcal{V} \right\}.$$ 

We shall consider only histories \(h_t \in H\).

Our first result characterizes histories after which a link is broken.

Lemma 1. For any history \(h_t \in H\), a link \(ij\) is broken at \(h_t\) if and only if \(v_{ik}(h_t) > \beta(v_{ij}(h_t))\) for some \(k \in N \setminus \{i, j\}\).
When \( v_{ik}^t > \beta(v_{ij}^t) \), individual \( i \) never links to \( j \) because the cost \( \gamma(\pi_{kt}, v_{ik}^t) \) of linking to \( k \) is always lower than the cost \( \gamma(\pi_{jt}, v_{ij}^t) \) of linking to \( j \). Since \( v_{ij}^t \) remains constant and \( v_{ik}^t \) cannot decrease, \( i \) never links to \( j \) thereafter, i.e., the link \( ij \) is broken. Conversely, if the inequality is reversed, \( i \) links to \( j \) when \( j \) is sufficiently well-informed and all others are sufficiently poorly informed.

The next result characterizes histories after which a link becomes free.

**Lemma 2.** A link \( ij \) is free after history \( h_t \in H \) if and only if

\[
v_{ij}^t(h_t) > \min \left\{ \overline{v}, \max_{k \in N \setminus \{i,j\}} \beta( v_{ik}^t(h_t) ) \right\}.
\]

When \( v_{ij}^t(h_t) > \beta(v_{ik}^t(h_t)) \) for all \( k \in N \setminus \{i,j\} \), all links \( ik \) are broken by Lemma 1, and hence \( i \) links to \( j \) in all subsequent periods, and \( ij \) is therefore free. Moreover, when \( v_{ij} > \overline{v} \), \( i \) links to \( j \) with positive probability in each period, and each such link causes \( v_{ij} \) to increase further. Hence the probability that \( i \) links to \( j \) remains positive perpetually, so \( ij \) is free. Conversely, in all remaining cases, there is a positive probability that \( i \) will link to some other node \( k \) repeatedly until \( v_{ik} \) exceeds \( \beta(v_{ij}(h_t)) \), resulting in the link \( ij \) being broken. (By Observation 1, this happens when \( i \) links to \( k \) at least \((\beta(v_{ij}(h_t)) - v_{ik}(h_t))/\delta \) times.) Note that the above lemmas imply that along every infinite history, every link eventually either breaks or becomes free.

These results may be illustrated with a simple example for \( N = \{1,2,3\} \). Figure 4 plots regions of the state space in which the links 31 and 32 are broken or free, for various values of \( v_{31} \) and \( v_{32} \) (the precisions of individual 3’s beliefs about the perspectives of 1 and 2 respectively). It is assumed that \( a = 1 \) and \( b = 2 \) so \( \overline{v} = 0.5 \). In the orthant above \((\overline{v}, \overline{v})\) links to both nodes are free by Lemma 2. Individual 3 links to each of these nodes with positive probability thereafter, eventually becoming arbitrarily close to learning both their perspectives. Hence, in the long run, she links with likelihood approaching 1 to whichever individual is better informed in any given period. This limiting behavior is therefore independent of past realizations, and illustrates our characterization of history independence in the next section.

When \( v_{32} > \beta(v_{31}) \), the region above the steeper curve in the figure, the link 31 breaks. Individual 3 links only to 2 thereafter, learning her perspective and therefore fully incorporating her information in the long run. But this comes at the expense of failing to link to individual 1 even when the latter is better informed. Along similar lines, in the region below flatter curve, 3 links only to 1 in the long run.
Now consider the region between the two curves but outside the orthant with vertex at \((\bar{v}, \bar{v})\). Here one or both of the two links remains to be resolved. If \(\overline{v} < v_{32} < \beta(v_{31})\), then although the link 32 is free, the link 31 has not been resolved. Depending on subsequent expertise realizations, either both links will become free or 31 will break. Symmetrically, when \(\overline{v} < v_{31} < \beta(v_{32})\), the link 31 is free while the other link will either break or become free in some future period.

Finally, in the region between the two curves but below the point \((\bar{v}, \bar{v})\), individual 3 may attach to either one of the two nodes or enter the orthant in which both links are free. Note that the probability of reaching the orthant in which both links are free is zero for sufficiently small values of \((v_{31}, v_{32})\). For example, when \(\beta(v_0) - v_0 < \delta\), regardless of the initial expertise levels, 3 will attach to the very first individual to whom she links. The critical value of \(v_0\) in this example is approximately 0.07, and the relevant region is shown at the bottom left of the figure.

Since the initial precisions of beliefs about perspectives lie on the 45 degree line by assumption, the size of this common precision \(v_0\) determines whether history independence is ensured, is possible but not ensured, or is not possible.\(^8\) In the first of these cases,

\(^8\)It is tempting to conclude that in the three person case, these three regimes correspond to the three
individuals almost always link to the best informed person in the long run, and the history of realizations eventually ceases to matter. In the second case, this outcome is possible but not guaranteed: there is a positive probability that some links will be broken. And in the third case, history matters perpetually and initial realizations have permanent effects. We now identify the conditions under which each of these three regimes arises and describe some of the possible structures that can emerge.

4 History Independence

If the process of network formation is history independent in the long run, then each individual will eventually observe the best informed among the rest with high probability. Specifically, this probability can be made arbitrarily close to 1 if a sufficiently large number of realizations is considered:

**Definition 1.** For any given history $h_t$, the process $\{V^t\}_{t=1}^{\infty}$ is said to be history independent at $h_t$ if, for all $\varepsilon > 0$, there exists $t^* > t$ such that

$$\Pr\left(j_{i't'} \in \arg \max_{j \neq i} \pi_{j_{i't'}} | h_t\right) > 1 - \varepsilon$$

for all $t' > t^*$ and $i \in N$. The process $\{V^t\}_{t=1}^{\infty}$ is said to be history independent if it is history independent at the initial history $h_1$.

Clearly the process cannot be history independent in this sense if there is a positive probability that one or more links will be broken at any point in time. Moreover, history independence is obtained whenever all links become free and have uniform positive bound on probability of occurrence throughout. Building on this fact and Lemma 2, the next result provides a simple characterization for history independence.

**Proposition 1.** For any $h_t \in H$, the process $\{V^t\}_{t=1}^{\infty}$ is history independent at $h_t$ if and only if $v_{ij}(h_t) > \overline{v}$ for all distinct $i, j \in N$. In particular, for $h_1 \in H$, the process $\{V^t\}_{t=1}^{\infty}$ is history independent if and only if $v_0 > \overline{v}$.

The condition for history independence may be interpreted as follows. For any given value of the support $[a, b]$ from which levels of expertise are drawn, history independence segments of the of the diagonal in Figure 4. But this is not correct, since the condition $\beta(v_0) - v_0 < \frac{\delta}{2}$ is sufficient but not necessary for at least one link to break. Specifically, there are values of $v_0$ outside the region on the bottom left of the figure such that both links can become free in the long run for any one observer, but not for all three. A fuller characterization is provided below.
arises if beliefs about the perspectives of others are sufficiently precise. That is, if each individual is sufficiently well-understood by others even before any opinions have been observed. Conversely, when there is substantial initial uncertainty about the individuals’ perspectives, the long-run behavior is history dependent with positive probability.

Depending on the extreme values $a$ and $b$ of possible expertise levels, the threshold $\overline{v}$ can take any value. When expertise is highly variable in absolute or relative terms (i.e. $b - a$ or $b/a$ are large), $\overline{v}$ is small, leading to history independence for a broad range of $v_0$ values. Conversely, when expertise is not sufficiently variable in the same sense, the threshold $\overline{v}$ becomes large, and history independence is more likely to fail. This makes intuitive sense, since it matters less to whom one links under these conditions, and hysteresis is therefore less costly in informational terms.

The logic of the argument is as follows. When $v_{ij}^0 = v_0 > \overline{v}$, there is a positive lower bound on the probability that a $i$ links to $j$ at the outset, regardless of her beliefs about others. Since $v_{ij}^t$ is nondecreasing in $t$, this lower bound is valid at all dates and histories, so $i$ links to $j$ infinitely often with probability 1. But every time $i$ links to $j$, $v_{ij}^t$ increases by at least $\delta$. Hence, after a finite number of periods, $i$ knows the perspective of $j$ with arbitrarily high precision. This of course applies to all other individuals, so $i$ comes to know all perspectives very well, and chooses advisors largely on the basis of their expertise level. Conversely, when $v_{ij}^0 = v_0 < \overline{v}$, it is possible that $i$ ends up linking to another individual $j'$ sufficiently many times, learning his perspective with such high precision that the link $ij$ breaks. After this point, $i$ no longer observes $j$ no matter how well informed the latter may be.

Proposition 1 identifies a necessary and sufficient condition for history independence at the initial history. If this condition fails to hold, then the process $\{V^t\}_{t=1}^{\infty}$ exhibits hysteresis: there exists a date $t$ by which at least one link is broken with positive probability. History independence (at the initial history) and hysteresis are complements because in our model any link either becomes free or breaks along every path, and history independence is equivalent to all links becoming eventually free with probability 1. Proposition 1 therefore can be restated as follows: $\{V^t\}_{t=1}^{\infty}$ exhibits hysteresis if and only if $v_0 < \overline{v}$.

When history independence fails, a number of interesting network structures can arise. We now consider three of these: opinion leadership, informational segregation, and static communication networks.
5 Network Structures

5.1 Opinion Leadership

One network structure that can arise is opinion leadership, with some subset of individuals being observed with high frequency even when their levels of expertise are known to be low, while others are never observed regardless of their levels of expertise. This can happen because repeated observation of a leader allows her perspective to become well understood by others, and hence her opinion can be more easily interpreted even when her information is poor.

We say that a sample path exhibits opinion leadership if there is some period $t$ and some nonempty subset $S \subset N$ such that $b_{ij} = 1$ for all $(i, j) \in N \times S$. That is, opinion leadership exists if some individuals are never observed (regardless of expertise realizations) after time $t$ along the sample path in question.

An special case of opinion leadership arises when $n$ links are free while the rest are all broken. In this case, all individuals are locked into a particular target, regardless of expertise realizations. In an extreme case, there may be a single leader to whom all others link, and a second individual to whom the leader alone links in all periods. We refer to this property of sample paths as extreme opinion leadership.

Define the cutoff $\tilde{v} \in (0, \bar{v})$ as the unique solution to the equation

$$\beta(\tilde{v}) - \tilde{v} = \hat{\delta}. \quad (10)$$

The following result establishes that unless we have history independence (in which case hysteresis is impossible) there is a positive probability of extreme opinion leadership, and such extreme leadership is inevitable when $v_0$ is sufficiently small:

**Proposition 2.** For $h_1 \in H$, $\{V^t\}_{t=1}^\infty$ exhibits extreme opinion leadership (i) with positive probability if and only if $v_0 < \bar{v}$, and (ii) with probability 1 if and only if $v_0 < \tilde{v}$.

The intuition for this result is straightforward: any network that is realized in period $t$ has a positive probability of being realized again in period $t + 1$ because the only links that can possibly break at $t$ are those that are inactive in this period. Hence there is a positive probability that the network that forms initially will also be formed in each of the first $s$ periods for any finite $s$. For large enough $s$ all links must eventually break except those that are active in all periods, resulting in extreme opinion leadership. Moreover, when $v_0 < \tilde{v}$,
we have $v_0 + \delta > \beta(v_0)$ and, by Lemma 1, each individual adheres to their very first target regardless of subsequent expertise levels. The most informed individual in the first period emerges as the unique information leader and herself links perpetually to the individual who was initially the second best informed.

More generally, two or more information leaders may emerge, who might themselves have different sets of targets. An example is shown in Figure 5, where nodes 1 and 4 emerge as leaders, and themselves link to 5 and 3 respectively. By the sixth period all links that target a member of the set $\{2, 6\}$ are broken, and these two individuals are never subsequently observed. Furthermore, the two information leaders are each locked in to a single target, while the remaining individuals observe both information leaders with positive probability in all periods.

Figure 5: Emergence of Information Leadership

5.2 Information Segregation

Despite the ex ante symmetry of the model, it is possible for clusters to emerge in which individuals within a cluster link only to others within the same cluster in the long run. In this case there may even be a limited form of history independence within clusters, so that
individuals tend to link to the best informed in their own group, but avoid linkages that cross group boundaries.

We say that a sample path exhibits segregation over a partition \( \{ S_1, S_2, \ldots, S_m \} \) of \( N \) if there is a period \( t \) such that \( b_{ij}^t = 1 \) for all \((i, j) \in S_k \times S_l \) with \( k \neq l \). That is, segregation over a partition \( \{ S_1, S_2, \ldots, S_m \} \) is said to arise if no link involving elements of different clusters can form after some period is reached, and members of each cluster \( S_k \) communicate only with fellow members of their own cluster. We say that a sample path exhibits segregation if it exhibits segregation over some partition with at least two disjoint clusters.

The first few periods of a sample path that exhibits segregation is illustrated in Figure 6. In this case the disjoint clusters \( \{ 1, 2, 3 \} \) and \( \{ 4, 5, 6 \} \) emerge with positive probability. Although this network is not resolved by the end of the last period depicted, it is easily seen that there as a positive probability of segregation after this history since no link that connects individuals in two different clusters is free.

![Figure 6: Emergence of Segregated Clusters](image)

In order for a segregation to arise over a partition \( \{ S_1, S_2, \ldots, S_m \} \), each \( S_k \) must have at least two elements. Excluding the trivial partition \( \{ N \} \), write \( \mathcal{P} \) for the set of all partitions \( \{ S_1, S_2, \ldots, S_m \} \) with \( m \geq 2 \) and \( |S_k| \geq 2 \) for all \( k \). This is the set of all partitions over which segregation could conceivably arise.
Segregation can arise only if initial precision level $v_0$ are small enough to rule out history independence. Furthermore, if $v_0 > \bar{v} - \delta$, all links to the best informed individual in the first period become free. This is because all such links are active in the first period, and the precision of all beliefs about this particular target’s perspective rise above $v_0 + \delta > \bar{v}$. These links are then free by Proposition 1, which clearly rules out segregation. So $v_0$ cannot be too large if segregation is to arise. And it cannot be too small either, otherwise individuals get locked into common early targets. For example, extreme opinion leadership, in which a single information leader is observed repeatedly by all others, is inconsistent with segregation and arises with certainty when $v_0 < \bar{v}$ (Proposition 2). The following result establishes that in all the other cases, segregation arises with positive probability over any partition in $\mathcal{P}$:

**Proposition 3.** Suppose $n \geq 4$. For any $h_1 \in H$ and any partition $\{S_1, S_2, \ldots, S_m\} \in \mathcal{P}$, the process $\{V^t\}_{t=1}^{\infty}$ exhibits segregation over $\{S_1, S_2, \ldots, S_m\}$ with positive probability if and only if $v_0 \in (\bar{v}, \bar{v} - \delta)$.

The forces that give rise to segregation can be understood by reconsidering the example depicted in Figure 6, where two segregated clusters of equal size emerge in a population of size 6. Nodes 1, 2 and 3 are the best informed, respectively, in the first three periods. After period 4, all links from this cluster to the nodes 4–6 are broken. Following this nodes 4–6 are best informed and link to each other, but receive no incoming links. Although the network is not yet resolved by the ends of the sixth period, it is clear that segregation can arise with positive probability because any finite repetition of the period 6 network has positive probability, and all links across the two clusters must break after a finite number of such repetitions. Hence a very particular pattern of expertise realizations is required to generate segregation, but any partition of the population into segregated clusters can arise with positive probability.

### 5.3 Static Networks

When $v_0 > \bar{v}$, all links are free to begin with. At the other extreme, when $v_0 < \bar{v}$, the long run outcome is necessarily extreme opinion leadership, resulting in the lowest possible level of information aggregation. For intermediate values of $v_0$, while extreme opinion leadership remains possible, other structures can also arise. As shown above, individuals can be partitioned into any arbitrary set of clusters of at least two individuals, with no cross-cluster communication at all.

This indeterminacy of network structures extends further. We shown next that each
individual may be locked into a single, arbitrarily given target in the long run. This implies that every worst case scenario (with respect to information aggregation) can arise with positive probability.

Let $\mathcal{G}$ denote the set of functions $g : N \to N$ that satisfy $g(i) \neq i$. Each element of $\mathcal{G}$ thus corresponds to a directed graph in which each node is linked to one (not necessarily unique) target. We say that a sample path converges to $g \in \mathcal{G}$ if there exists a period $t^*$ such that, for all $i \in N$ and all $t > t^*$, $j_{it} = g(i)$. The process $\{V^t\}_{t=1}^{\infty}$ converges to $g$ with positive probability if the probability that a sample path will converge to $g$ is positive. In this case there is a positive probability that each individual eventually links only to the target prescribed for her by $g$.

In order to identify the range of parameter values for which any given network $g \in \mathcal{G}$ can emerge with positive probability as an outcome of the process, we make the following assumption.

**Assumption 1.** There exists $\pi \in (a,b)$ such that $\gamma(\pi, v_0) < \gamma(a, v_0 + \delta(\pi, b))$ and $\gamma(b, v_0) < \gamma(\pi, v_0 + \delta(\pi, b))$.

Note that this assumption is satisfied whenever $v_0 > v^*$ where $v^*$ is defined by

$$\beta(v^*) - v^* = 2\delta(b, b).$$

In addition to Assumption 1, convergence to an arbitrary network $g \in \mathcal{G}$ requires that $v_0$ be sufficiently small:

**Proposition 4.** Assume that $v_0 < \bar{v} - \delta(b,b)$ and satisfies Assumption 1. Then, for any graph $g \in \mathcal{G}$, the process $\{V^t\}_{t=1}^{\infty}$ converges to $g$ with positive probability.

A sufficient condition for such convergence to occur is $v_0 \in (v^*, \bar{v} - \delta(b,b))$, and it is easily verified that this set is nonempty. For instance if $(a, b) = (1, 2)$, then $(v^*, \bar{v} - \delta(b,b)) = (0.13, 0.20)$.

While the emergence of opinion leadership is intuitive, the possibility of convergence to an arbitrary graph is much less so. Since all observers face the same distribution of expertise in the population, and almost all link to the same target in the initial period, the possibility that they may all choose different targets in the long run is counter-intuitive. Nevertheless, there exist sequences of expertise realizations that result in such strong asymmetries.
6 Strong Hysteresis

The three classes of networks discussed in the previous section are not by any means exhaustive, and a variety of other outcomes are possible when the condition for history independence does not hold at the initial history. Recall that the process \( \{V^t\}_{t=1}^\infty \) exhibits hysteresis if there exists a date \( t \) by which at least one link is broken with positive probability. Note that this is consistent with the possibility that all links become free with positive probability. Hysteresis rules out history independence at the initial history, but allows for history independence to arise after some histories with positive probability.

We now introduce a stronger notion of hysteresis, which rules out the possibility that all links will eventually be free. For any given history \( h_t \), the process \( \{V^t\}_{t=1}^\infty \) is said to exhibit strong hysteresis at \( h_t \) if the probability that no links will break in period \( t + 1 \) is zero. It is said to exhibit strong hysteresis if it exhibits strong hysteresis at the initial history \( h_0 \).

An immediate implication of Proposition 2 is that the process exhibits strong hysteresis if \( v_0 < \tilde{v} \), since this is sufficient for opinion leadership to arise with probability 1. In this case each individual links perpetually to the first person they observe. However, \( v_0 < \tilde{v} \) is not necessary for strong hysteresis. To see why, consider the three agent example described in Section 3.3. Here \( v_0 < \tilde{v} \) corresponds to the segment of the 45 degree line in the bottom left section of Figure 4. If \( v_0 \) lies within this range, one of the two links originating at 3 will break after the first observation is made. If \( v_0 \) lies outside this range, then there is a positive probability that both links 31 and 32 will eventually be free. But this does not mean that there is a positive probability that all links in the network will be free: sample paths that result in both 31 and 32 being free might require that some other link be broken. This is in fact the case.

We next identify a necessary and sufficient condition for strong hysteresis. To this end, define \( \hat{v} \) as the unique solution to

\[
\beta(\hat{v}) - \hat{v} = \delta(b,b).
\]  

(11)

We then have:

**Proposition 5.** For any \( h_1 \in H \), the process \( \{V^t\}_{t=1}^\infty \) exhibits strong hysteresis if and only if \( v_0 < \hat{v} \).

It is easily verified that \( \hat{v} > \tilde{v} \), as expected. The condition \( v_0 < \hat{v} \) is necessary and sufficient for all links to break in the initial period except for the ones that are active,
resulting in opinion leadership. The weaker condition $v_0 < \hat{v}$ is necessary and sufficient for at least one link to break. This rules out history independence at any future period, but allows for a broad range of network structures to emerge in the long run, including segregation and static networks.

7 Conclusions

Interpreting the opinions of others is challenging because such opinions are based in part on private information and in part on prior beliefs that are not directly observable. Individuals seeking informative opinions may therefore choose to observe those whose priors are well-understood, even if their private information is noisy. This problem is compounded by the fact that observing opinions is informative not only about private signals but also about prior perspectives, so preferential attachment to particular persons can develop endogenously over time. And since the extent of such attachment depends on the degree to which the observer is well-informed, there is a natural process of symmetry breaking. This allows for a broad range of networks to emerge over time, including opinion leadership and informations segregation.

Our analysis has been based on a number of simplifying assumptions. We have assumed that just one target can be observed in each period rather than several, and this could be relaxed by allowing for costs of observation that increase with the number of targets selected. Observation of the actions of others, and observation of the state itself could also be informative and affect beliefs about perspectives. It would also be worth relaxing the assumption of myopic choice, which would allow for some experimentation. We suspect that perfectly patient players will choose targets in a manner that implies history independence, but that our qualitative results will survive as long as players are sufficiently impatient. But these and other extensions are left for future research.
Appendix

Proof of Lemma 1. To prove sufficiency, take \( v_{ik}^t(h_t) > \beta \left( v_{ij}^t(h_t) \right) \). By definition of \( \beta \),
\[
\gamma \left( a, v_{ik}^t(h_t) \right) < \gamma \left( a, \beta \left( v_{ij}^t(h_t) \right) \right) = \gamma \left( b, v_{ij}^t(h_t) \right)
\]
where the inequality is by monomotonicity of \( \gamma \) and the equality is by definition of \( \beta \). Hence, \( \Pr \left( l_{ij}^t = 1 | h_t \right) = 0 \). Moreover, by (9), at any \( h_{t+1} \) that follows \( h_t \), \( v_{ij}^{t+1} \left( h_{t+1} \right) = v_{ij}^t \left( h_t \right) \) and \( v_{ik}^{t+1} \left( h_{t+1} \right) \geq v_{ik}^t \left( h_t \right) \), and hence the previous argument yields \( \Pr \left( l_{ij}^{t+1} = 1 | h_t \right) = 0 \). Inductive application of the same argument shows that \( \Pr \left( l_{ij}^s = 1 | h_t \right) = 0 \) for every \( s \geq 0 \), showing that the link \( ij \) is broken at \( h_t \). Conversely, suppose that \( v_{ik}^t \left( h_t \right) < \beta \left( v_{ij}^t \left( h_t \right) \right) \) for every \( k \in N \setminus \{i,j\} \). Then, by definition of \( \beta \), for all \( k \not\in \{i,j\} \),
\[
\gamma \left( b, v_{ij}^t \left( h_t \right) \right) = \gamma \left( a, \beta \left( v_{ij}^t \left( h_t \right) \right) \right) < \gamma \left( a, v_{ik}^t \left( h_t \right) \right),
\]
where the inequality is by \( \gamma \) being decreasing in \( v \). Hence, by continuity of \( \gamma \), there exists \( \eta > 0 \) such that for all \( k \not\in \{i,j\} \),
\[
\gamma \left( b - \eta, v_{ij}^t \left( h_t \right) \right) < \gamma \left( a + \eta, v_{ik}^t \left( h_t \right) \right).
\]
Consider the event \( \pi_{jt} \in [b - \eta, b] \) and \( \pi_{kt} \in [a, a + \eta] \) for all \( k \neq j \). This has positive probability, and on this event \( l_{ij}^t = 1 \), showing that link \( ij \) is not broken at \( h_t \).

Proof of Lemma 2. To prove sufficiency, first take any \( i, j \) with \( v_{ij}^t \left( h_t \right) > \overline{v} \). Then, by definition of \( \overline{v} \), for any \( k \not\in \{i,j\} \),
\[
\gamma \left( b, v_{ij}^t \left( h_t \right) \right) < \gamma \left( b, \overline{v} \right) \leq \gamma \left( a, v_{ik}^t \left( h_t \right) \right),
\]
where the first inequality is because \( \gamma \) is decreasing in \( v \) and the second inequality is by definition of \( \overline{v} \). Hence, by continuity of \( \gamma \), there exists \( \eta > 0 \) such that for all \( k \not\in \{i,j\} \),
\[
\gamma \left( b - \eta, v_{ij}^t \left( h_t \right) \right) < \gamma \left( a + \eta, v_{ik}^t \left( h_t \right) \right).
\]
Consider the event \( \pi_{jt} \in [b - \eta, b] \) and \( \pi_{kt} \in [a, a + \eta] \) for all \( k \neq j \). This has positive probability, and on this event \( l_{ij}^t = 1 \). Hence \( \Pr(b_{ij}^t = 1) = 0 \). For any \( s \geq t \), since \( v_{ij}^s \geq v_{ij}^t \geq \overline{v} \), we have \( \Pr(l_{ij}^s = 1) > 0 \), showing that the link \( ij \) is free. On the other hand, if \( v_{ij}^t \left( h_t \right) \geq \max_{k \in N \setminus \{i,j\}} \beta \left( v_{ik}^t \left( h_t \right) \right) \), then, by Lemma 1, all the links \( ik \) with \( k \in N \setminus \{i,j\} \) are broken at \( h_t \), and hence \( i \) links to \( j \) with probability one thereafter. Therefore, the link \( ij \) is free. This proves sufficiency.

For the converse, take \( v_{ij}^t \left( h_t \right) < \min \left\{ \overline{v}, \max_{k \in N \setminus \{i,j\}} \beta \left( v_{ik}^t \left( h_t \right) \right) \right\} \). We will show that the link \( ij \) will break with positive probability by some \( t^* > t \). Since \( v_{ij}^{t^*} \left( h_t \right) < \overline{v}, \beta \left( v_{ij}^{t^*} \left( h_t \right) \right) \)
is finite. Moreover, since \( v_{ij}^t (h_t) \leq \max_{k \in N \setminus \{i,j\}} \beta (v_{ik}^t (h_t)) \), there exists \( k \neq j \) such that 
\[
\gamma (b, v_{ik}^t (h_t)) > \gamma (a, v_{i,k'}^t (h_t)) \quad \text{for every } k'.
\]
If \( v_{ik}^t (h_t) > \beta (v_{ij}^t (h_t)) \), by Lemma 1, the link \( ij \) is broken at \( h_t \), as desired. Assume that \( v_{ik}^t (h_t) < \beta (v_{ij}^t (h_t)) \). By continuity of \( \gamma \), there exists \( \eta > 0 \) such that \( \gamma (\pi_{kt}, v_{ik}^t (h_t)) > \gamma (\pi_{k't}, v_{ik'}^t (h_t)) \) on the positive probability event that \( \pi_{kt} \in [b - \eta, b] \) and \( \pi_{k't} \in [a, a + \eta] \) for all \( k' \neq k \). In that case, \( i \) links to \( k \), increasing \( v_{ik}^t \) and \( v_{ik'}^t \). Hence, \( i \) keeps linking to \( k \) on the positive probability event that \( \pi_{ks} \in [b - \eta, b] \) and \( \pi_{k's} \in [a, a + \eta] \) for all \( k' \neq k \) and \( s \in \{t, t + 1, \ldots, t^* \} \) where \( t^* = t + \lceil (\beta (v_{ij}^t (h_t)) - v_{ik}^t (h_t)) / \delta \rceil \). \(^{9}\) Then, on that event, by (9),
\[
v_{ik}^t = v_{ij}^t (h_t) + \sum_{s = t}^{t^*} \delta (\pi_{is}, \pi_{ks}) \geq v_{ik}^t (h_t) + \lceil (\beta (v_{ij}^t (h_t)) - v_{ik}^t (h_t)) / \delta \rceil \delta > \beta (v_{ij}^t (h_t)),
\]
where the inequality is by Observation 1. Therefore, the link \( ij \) breaks by \( t^* \) on this event. \( \square \)

**Proof of Proposition 1.** First take \( v_{ij}^t (h_t) < \overline{\nu} \) for some distinct \( i, j \in N \). If \( v_{ij}^t (h_t) \geq \max_{k \in N \setminus \{i,j\}} \beta (v_{ik}^t (h_t)) \), then all the links \( ik \) with \( k \neq j \) are broken at \( h_t \). Otherwise, as shown in the proof of Lemma 2, the link \( ij \) is broken with positive probability by some \( t^* > t \). In either case, \( \Pr (j_{is} \in \arg \max_k \pi_{ks} | h_t) \) is bounded away from 1, showing that \( \{V^t\}_{t=1}^{\infty} \) is not history independent at \( h_t \).

Assume now \( v_{ij}^t (h_t) > \overline{\nu} \) for all distinct \( i, j \in N \). Of course, \( v_{ij}^s (h_s) \geq v_{ij}^t (h_t) > \overline{\nu} \) for all distinct \( i, j \in N \) and for every history after \( h_t \). Now, since \( \gamma (\pi, \nu) \) is continuous in \( \pi \) and \( 1 / \nu \) and \( F \) is continuous over \( [a, b] \), for every \( \varepsilon > 0 \), there exists \( \tilde{\nu} < \infty \) such that 
\[
\Pr (j_{is} \in \arg \max_{j \neq i} \pi_{js}) > 1 - \varepsilon \quad \text{whenever } v_{ij}^s > \tilde{\nu} \quad \text{for all distinct } i \text{ and } j.
\]
Hence, it suffices to show that, conditional on \( h_t \), \( v_{ij}^s \to \infty \) as \( s \to \infty \) for all distinct \( i \) and \( j \) almost surely. To this end, observe that
\[
\gamma (b, v_{ij}^t (h_t)) < \gamma (b, \overline{\nu}) \leq \gamma (a, v) \quad (\forall v, i, j),
\]
where the first equality is because \( \gamma \) is decreasing in \( v_{ij}^t \) and the second inequality is by definition of \( \overline{\nu} \). Hence, by continuity of \( \gamma \), there exists \( \eta > 0 \) such that
\[
\gamma (b - \eta, v_{ij}^t (h_t)) < \gamma (a + \eta, v) \quad (\forall v, i, j).
\]
Since \( v_{ij}^s (h_s) \geq v_{ij}^t (h_t) > \overline{\nu} \), this further implies that
\[
\gamma (b - \eta, v_{ij}^s) < \gamma (a + \eta, v_{ik}^s)
\]
\(^{9}\)Here, \( \lceil x \rceil \) denotes the smallest integer larger than \( x \).

30
for every history that follows \( h_t \), for every distinct \( i, j, k \), and for every \( s \). Consequently, \( l_{ij}^{s+1} = 1 \) whenever \( \pi_{js} > b - \eta \) and \( \pi_{ks} \leq a + \eta \) for all other \( k \). Thus,

\[
\Pr(l_{ij}^{s+1} = 1) \geq \lambda
\]

after any history that follows \( h_t \) and any date \( s \geq t \) where

\[
\lambda = F(a + \eta)^{n-2} (1 - F(b - \eta)) > 0.
\]

Therefore, \( l_{ij}^{s+1} = 1 \) occurs infinitely often for all distinct \( i, j \in N \) almost surely conditional on \( h_t \). But whenever \( l_{ij}^{s+1} = 1 \), \( v_{ij}^{s+1} \geq v_{ij}^{s} + \delta \), where \( \delta = \delta(a, b) > 0 \), showing that \( v_{ij}^{s} \to \infty \) as \( s \to \infty \) for all distinct \( i, j \in N \) almost surely conditional on \( h_t \). This completes the proof.

Proof of Proposition 2. Clearly, when \( v_0 > \overline{v} \), the long-run outcome is history independent by Proposition 1, and hence opinion leadership is not possible. Accordingly, suppose that \( v_0 < \overline{v} \). Consider the positive probability event \( A \) that for every \( t \leq t^* \), \( \pi_{1t} > \pi_{2t} > \max_{k>2} \pi_{kt} \) for some \( t^* > (\beta(v_0) - v_0)/\delta \). Clearly, on event \( A \), for any \( t \leq t^* \) and \( k > 1 \), \( j_{kt} = 1 \) and \( j_{1t} = 2 \), as the targets are best informed and best known individuals among others. Then, on event \( A \), for \( ij \in S \equiv \{12, 21, 31, \ldots, n1\} \),

\[
v_{ij}^{t^*+1} = v_0 + t^* \delta > \beta(v_0)
\]

while \( v_{ik}^{t^*+1} = v_0 \) for any \( ik \not\in S \). (Here, the equalities are by (9); the weak inequality is by Observation 1, and the strict inequality is by definition of \( t^* \).) Therefore, by Lemma 1, all the links \( ik \not\in S \) are broken by \( t^* \), resulting in the extreme opinion leadership as desired.

To prove the second part, note that for any \( v_0 \leq \bar{v} \) and \( i \in N \),

\[
v_{ij,1}^2 = v_0 + \delta(\pi_{i1}, \pi_{j1}) \geq v_0 + \delta > \beta(v_0)
\]

while \( v_{ik}^2 = v_0 \) for all \( k \neq j_{1i} \), showing by Lemma 1 that all such links \( ik \) are broken after the first period. Since \( j_{1i} = \min \arg \max, \pi_{i1} \) for every \( i \neq \min \arg \max, \pi_{i1} \), this shows that extreme leadership emerges at the end of first period with probability 1. The claim that extreme opinion leadership arises with probability less than 1 if \( v_0 > \bar{v} \) follows from Proposition 3, which is proved below.

Proof of Proposition 3. Take any \( v_0 \in (\bar{v}, \overline{v} - \delta) \) and any partition \( \{S_1, \ldots, S_m\} \) where each cluster \( S_k \) has at least two elements \( i_k \) and \( j_k \). We will now construct a postive probability
event on which the process exhibits segregation over partition \( \{S_1, \ldots, S_m\} \). Since \( v_0 \in (\tilde{v}, \eta - \delta) \), there exists a small \( \varepsilon > 0 \) such that
\[
v_0 + \delta (a + \varepsilon, b - \varepsilon) < \min \{ \beta (v_0), \tilde{v} \}
\]
and
\[
\delta (b - \varepsilon, b) > \delta (a + \varepsilon, b - \varepsilon).
\]
By (13) and by continuity and monotonicity properties of \( \gamma \), there also exist \( \pi^* \in (a, b) \) and \( \varepsilon' > 0 \) such that
\[
\gamma (\pi^* - \varepsilon', v_0 + \delta (b - \varepsilon, b)) < \gamma (b, v_0)
\]
\[
\gamma (\pi^* + \varepsilon', v_0 + \delta (a + \varepsilon, b - \varepsilon)) > \gamma (b - \varepsilon, v_0).
\]
For every \( t \in \{2, \ldots, m\} \), the realized expertise levels are as follows:
\[
\pi_{it} > \pi_{jt} > \pi_{st} > b - \varepsilon \quad (\forall i \in S_t)
\]
\[
\pi^* + \varepsilon' > \pi_{ipt} > \pi_{ijt} > \pi_{sit} > \pi^* - \varepsilon' \quad (\forall i \in S_k, k < t)
\]
\[
\pi_{it} < a + \varepsilon \quad (\forall i \in S_k, k > t).
\]
Fixing
\[
t^* > (\beta (v_0 + \delta (a + \varepsilon, b - \varepsilon)) - v_0) / \delta,
\]
the realized expertise levels for \( t \in \{m + 1, \ldots, m + t^*\} \) are as follows:
\[
\pi^* + \varepsilon' > \pi_{ipt} > \pi_{ijt} > \pi_{sit} > \pi^* - \varepsilon' \quad (\forall i \in S_k, \forall k)
\]
The above event has clearly positive probability. We will next show that the links \( ij \) from distinct clusters are all broken by \( m + t^* + 1 \).

Note that at \( t = 1 \), \( j_{i_1} = j_1 \) and \( j_{i_1} = i_1 \) for all \( i \neq i_1 \). Hence,
\[
v_{i_{i_1}}^2 \geq v_0 + \delta (b - \varepsilon, b) > v_0 + \delta (a + \varepsilon, b - \varepsilon) \geq v_{j_{i_1}}^2 \quad (\forall i \in S_1, \forall j \not\in S_1),
\]
where the strict inequality is by (13). Therefore, by (14), at \( t = 2 \), each \( i \in S_1 \) sticks to his previous link
\[
j_{i_1} = j_1 \text{ and } j_{i_1} = i_1 \forall i \in S_1 \setminus \{i_1\},
\]
while each \( i \not\in S_1 \) switches to a new link
\[
j_{i_2} = j_2 \text{ and } j_{i_2} = i_2 \forall i \in N \setminus (S_1 \cup \{i_2\}).
\]
Using the same argument inductively, observe that for any \( t \in \{2, \ldots, m\} \), for any \( i \in S_k \) and \( i' \in S_t \) with \( k < t \leq l \), and for any \( s < t \),
\[
v_{i_{j_{i(t-1)}}}^t \geq v_0 + \delta (b - \varepsilon, b) > v_0 + \delta (a + \varepsilon, b - \varepsilon) \geq v_{i'_{j_{i's}}}^2.
\]
Hence, by (14),

\[
    j_{it} = \begin{cases} 
        j_{i(t-1)} & \text{if } i \in S_k \text{ for some } k < t \\
        j_t & \text{if } i = i_t \\
        i_t & \text{otherwise.} 
    \end{cases}
\]

In particular, at \( t = m \), for any \( i \in S_k \), \( j_{im} = i_k \) if \( i \neq i_k \) and \( j_{ikm} = j_k \). Once again,

\[
    v^{t}_{ijm} \geq v_0 + \delta \left( b - \varepsilon, b \right).
\]

Moreover, \( i \) could have observed any other \( j \) at most once, when \( \pi_{it} < a^* + \varepsilon \) and \( \pi_{jt} > b - \varepsilon \), yielding

\[
    v^{t}_{ij} \leq v_0 + \delta \left( a + \varepsilon, b - \varepsilon \right).
\]

Hence, by (14), \( i \) sticks to \( j_{im} \) by date \( m + t^* \), yielding

\[
    v^{m+t^*+1}_{ijm} \geq v_0 + \delta \left( b - \varepsilon, b \right) + t^* \delta > \left( v_0 + \delta \left( a + \varepsilon, b - \varepsilon \right) \right) \geq \beta \left( v^{m+t^*+1}_{ij} \right)
\]

for each \( j \neq j_{im} \). By Lemma 1, this shows that the link \( ij \) is broken. Since \( j_{im} \in S_k \), this proves the result.

\[
    \Box
\]

Proof of Proposition 4. Take \( v_0 \) as in the hypothesis, and take any \( g : N \rightarrow N \). We will construct some \( t^* \) and a positive probability event on which

\[
    j_{it} = g(i) \quad \forall i \in N, t > n + t^*.
\]

Now, let \( \pi \) be as in Assumption 1. By continuity of \( \delta \) and \( \gamma \), there exists a small but positive \( \varepsilon \) such that

\[
    \begin{align*}
    \gamma (\pi, v_0) & < \gamma (a, v_0 + \delta (b - \varepsilon, \pi + \varepsilon)) \quad (15) \\
    \gamma (b - \varepsilon, v_0) & < \gamma (\pi + \varepsilon, v_0 + \delta (\pi + \varepsilon, b - \varepsilon)) \quad (16) \\
    \delta (b - \varepsilon, \pi + \varepsilon) & > \delta (\pi + \varepsilon, b - \varepsilon). \quad (17)
    \end{align*}
\]

Fix some

\[
    t^* > \left( \beta \left( v_0 + \delta \left( \pi + \varepsilon, b - \varepsilon \right) \right) - v_0 \right) / \delta,
\]

and consider the following positive probability event:

\[
    \begin{align*}
    \pi_{it} & \geq b - \varepsilon > \pi + \varepsilon \geq \pi_{g(t)t} \geq \pi > a + \varepsilon \geq \pi_{jt} \quad (\forall j \in N \setminus \{t, g(t)\}, \forall t \in N), \\
    (\pi_{1t}, \ldots, \pi_{mt}) & \in A \quad (\forall t \in \{n+1, \ldots, n+t^*\})
    \end{align*}
\]

where

\[
    A \equiv \left\{ (\pi_1, \ldots, \pi_n) \mid \gamma (\pi_i, v_0 + \delta (\pi + \varepsilon, b - \varepsilon)) > \gamma (\pi_j, v_0 + \delta (b - \varepsilon, \pi + \varepsilon)) \forall i, j \in N \right\}.
\]
Note that $A$ is open and non-empty (as it contains the diagonal set). Note that at every date $t \in N$, the individual $t$ becomes an ultimate expert (with precision nearly $b$), and his target $g(t)$ is the second best expert.

We will next show that the links $ij$ with $j \neq g(i)$ are all broken by $n + t^* + 1$. Towards this goal, we will first make the following observation:

At every date $t \in N$, $t$ observes $g(t)$; every $i < t$ observes either $t$ or $g(i)$, and every $i > t$ observes $t$.

At $t = 1$, the above observation is clearly true: $1$ observes $g(1)$, while everybody else observes $1$. Suppose that the above observation is true up to $t - 1$ for some $t$. Then, by date $t$, for any $i \geq t$, $i$ has observed each $j \in \{1, \ldots, t - 1\}$ once, when his own precision was in $[a, \pi + \varepsilon]$ and the precision of $j$ was in $[b - b, b]$. Hence, by Observation 1, $v_{ij}^t \leq v_0 + \delta (\pi + \varepsilon, b - \varepsilon)$. He has not observed any other individual, and hence $v_{ij}^t = v_0$ for all $j \geq t$. Thus, by (16), for any $i > t$, $\gamma (\pi_{tt}, v_{ii}^t) < \gamma (\pi_{jt}, v_{ij}^t)$ for every $j \in N \setminus \{i, t\}$, showing that $i$ observes $t$, i.e., $j_{it} = t$. Likewise, by (15), for $i = t$, $\gamma (\pi_{gt}, v_{tt}^t) < \gamma (\pi_{jt}, v_{ij}^t)$ for every $j \in N \setminus \{t, g(t)\}$, showing that $t$ observes $g(t)$, i.e., $j_{tt} = g(t)$. Finally, for any $i < t$, by the inductive hypothesis, $i$ has observed any $j \neq g(i)$ at most once, yielding $v_{ij}^t \leq v_0 + \delta (\pi + \varepsilon, b - \varepsilon)$. Hence, as above, for any $j \in N \setminus \{i, t, g(i)\}$, $\gamma (\pi_{tt}, v_{ii}^t) < \gamma (\pi_{jt}, v_{ij}^t)$, showing that $i$ does not observe $j$, i.e., $j_{it} \in \{g(i), t\}$.

By the above observations, after the first $n$ period, each $i$ has observed any other $j \neq g(i)$ at most once, so that

$$v_{ij}^{n+1} \leq v_0 + \delta (\pi + \varepsilon, b - \varepsilon) \quad (\forall j \neq g(i)).$$

He has observed $g(i)$ at least once, and in one of these occations (i.e. at date $i$), his own precision was in $[b - \varepsilon, b]$ and the precision of $g(i)$ was in $[\pi, \pi + \varepsilon]$, yielding

$$v_{ig(i)}^{n+1} \geq v_0 + \delta (b - \varepsilon, \pi + \varepsilon).$$

By definition of $A$, inequalities (18) and (19) imply that each $i$ observes $g(i)$ at $n + 1$. Consequently, the inequalities (18) and (19) also hold at date $n + 2$, leading each $i$ again to observe $g(i)$ at $n + 2$, and so on. Hence, at dates $t \in \{n + 1, \ldots, t^* + n\}$, each $i$ observes $g(i)$, yielding

$$v_{ig(i)}^{n+t^*+1} \geq v_{ig(i)}^{n+1} + t^* \delta \geq v_0 + \delta (b - \varepsilon, \pi + \varepsilon) + \beta (v_0 + \delta (\pi + \varepsilon, b - \varepsilon)) - v_0$$

$$> \beta (v_0 + \delta (\pi + \varepsilon, b - \varepsilon)).$$

34
For any $j \neq g(i)$, since $v_{ij}^{n+t^*+1} = v_{ij}^{n+1}$, together with (18), this implies that

$$v_{ig(i)}^{n+t^*+1} > \beta \left(v_{ij}^{n+t^*+1}\right).$$

Therefore, by Lemma 1, the link $ij$ is broken at date $t^* + n + 1$. \hfill \Box

**Proof of Proposition 5.** Take $v_0 \leq \hat{v}$, so that $v_0 + \delta (b, b) \geq \beta (v_0)$. Write $i^* = \arg \max_i \pi_{i1}$ and $j^* = \arg \max_{i \neq i^*} \pi_{i1}$. With probability 1, $\pi_{i^*1} > \pi_{j^*1}$. Hence,

$$v_{i^*j^*}^{2} = v_0 + \delta (\pi_{i^*1}, \pi_{j^*1}) > v_0 + \delta (\pi_{i^*1}, \pi_{i^*1}) \geq v_0 + \delta (b, b) \geq \beta (v_0),$$

showing that the link $i^*j^*$ is broken by Lemma 1. To see the penultimate equality, note that $\delta (\pi, \pi)$ is decreasing in $\pi$. Conversely, when $v_0 > \hat{v}$, there exists $\varepsilon > 0$ such that $v_0 + \delta (b, b - \varepsilon) < \beta (v_0)$.

Then, no link is broken in the first period when $(\pi_{11}, \ldots, \pi_{n1}) \in [b - \varepsilon, b]^N$. \hfill \Box
References


