Persuasion in Networks*

Francesco Squintani
University of Warwick†

July, 2024

Abstract

This paper brings together two major research streams in economic theory: information transmission in networks and strategic communication. The model embeds "persuasion" games of strategic disclosure by Milgrom (1981) into the communication network framework by Jackson and Wolinsky (1996). I find that the unique optimal network is a line in which players are ordered according to their bliss points. This ordered line is also pairwise-stable. This finding stands in sharp contrast to previous results in network studies, that identify stars as the optimal and pairwise-stable networks when communication is non-strategic and subject to technological constraints. While stars are the most centralized minimally-connected networks, the line is the most decentralized one. These results may be especially relevant to political economy applications, such as networks of policymakers, interest groups, or judges.

*For helpful comments and discussions I thank Torun Dewan, Andrea Galeotti, Alex Frankel, Christian Ghiglino, Ilan Kremer, Pablo Montagnes, Phil Reny, Erik Snowberg, Pierre Yared, as well as audiences at seminars/conferences, including Brown University, Columbia University, Duke University, EIEF Bank of Italy, Erasmus University Rotterdam, Harvard University, Northwestern University, Paris School of Economics, Royal Holloway University, Stanford University, UCSD, UCLA, Universitá di Verona, University of Bonn, University of Chicago, University of Nottingham, University of Queensland, University of St Andrews, University of Utah, University of Urbana-Champaign, and Yale University.

†Department of Economics, Gibbett Hill Road, Coventry CV47AL, UK; f.squintani@warwick.ac.uk

1 Introduction

This paper proposes the marriage of two major research streams in economic theory: information transmission in networks and strategic communication. The model embeds "persuasion" games of strategic disclosure by Milgrom (1981) into the "social communication" network framework introduced by Jackson and Wolinsky (1996). My analysis yields results that overturn established insights in both the strategic communication literature, which usually assumes that experts communicate directly to decision-makers, and the information transmission in networks literature, which typically does not consider how strategic misrepresentation of information influences its transmission in a network.[1]

Specifically, I find that the insight from direct strategic communication, that verifiable information is fully disclosed in equilibrium, extends to strategic communication in networks if and only if all players on the path that connects an expert to a decision-maker are biased in the same direction relative to the decision-maker. As a consequence, the optimal network for strategic information turns out to be the "ordered line" in which each player forms a link only with the players with the most closely aligned preferences. This is the most economical network that ensures no verifiable information is strategically withheld in communication. This result stands in sharp contrast to established findings that optimal communication networks are highly centralized when abstracting from the possibility of strategic misrepresentation of information. Further, I show that such an ordered line endogenously forms as a Nash Equilibrium of a game of bilateral link sponsorship based on Myerson (1991), and is pairwise stable.

In order to describe my model and results in more detail, let us start by briefly reviewing social communication in networks. Consider a symmetric network, where each link represents who may communicate with whom within a group of players. Each player enjoys a benefit from being connected with any other player, possibly through a path of intermediaries in the network. Such a benefit represents, in reduced form, the value of the other players' information for the player when called to make decisions. The value of an indirect connection may decay exponentially with the length of the connecting path. Each player bears a cost when forming a link with any other player.

The optimality and stability of networks depend on the link costs. If they are very high,

[1] In both cases, there are only a few exceptions, which I discuss in the literature review (Section 2).

no links form, and the optimal network is empty. If costs are very low and there is decay in communication, a fully connected network is optimal and stable. For intermediate costs, if there is no information decay along the network, then any minimally optimal network (tree) is optimal. But if there is any decay, then only star networks are optimal: one player (center) is connected to all others (periphery), and no other links form. In fact, stars are the trees that minimize the sum of the lengths of the paths, and hence aggregate communication decay.

Turning to describe persuasion games, consider an uninformed agent who makes a decision based on the information provided by a perfectly informed expert. The former would like to match her decision with the state of the world. The expert would like the decision to match the state plus a constant "bias." The information transmitted by the expert is verifiable, hence the expert cannot lie, but he can withhold information. Despite the expert's bias, in equilibrium, he fully discloses all his information, and the decision-maker makes a perfectly informed decision. This outcome is supported by "worst-case" beliefs off the equilibrium path. If the expert were to withhold any information, the decision-maker would believe that the state of the world is the one most contrary to the expert's bias, compatibly with the information he released.

My model embeds a persuasion game into a network, and is presented in Section 3. An expert and a decision-maker are randomly selected from the players in the network, with each pair of players having a positive probability. Once chosen, their identities become common knowledge, and a persuasion game is played in the network. The expert's information is sent to the decision-maker through any path of players that connects them. Players have misaligned preferences, meaning that each would like to persuade the decision-maker to choose a different action, given the same state of the world.

For each network and pair of realized expert and decision-maker, I determine the value of the most informative equilibrium in the resulting persuasion game. The average of such values, weighted by the ex-ante probability of each pair of expert and decision-maker, determines the value that each individual player assigns to the network. The optimality and stability of networks are assessed using such values. This simple, micro-founded construction captures the main features of both social communication in networks—namely, that each player benefits from the information she may receive from any other player—and strategic (verifiable) information transmission, where each player may withhold information to influence the decision-maker's choice.

To keep the exposition simple and to highlight the departure from the existing networks literature, I abstract from non-strategic communication constraints and focus on strategic persuasion incentives. Also, I frame the analysis within the canonical setup of strategic communication where the players' payoffs follow a quadratic loss function with misaligned bliss points. I relax these assumptions in Section 5, which shows that my results are qualitatively the same as long as the players' payoffs satisfy mild assumptions, as well as when multiple experts and decision-makers may be called to play, and experts do not know the state and only observe noisy signals. Finally, I show that my results are robust when the information transmitted across the network decays independently of the players' choices, as long as the decay probability is small relative to the players' bliss point misalignments.

My first result is that the equilibrium analysis of persuasion games, for a given network and realized expert and decision-maker, depends on the characteristics of all the players along the paths that connect the expert with the decision-maker. I introduce the concept of a "bias reversal" path, which denotes a path where not all players are biased in the same direction relative to the decision-maker. For example, this occurs if the expert is biased to the right relative to the decision-maker, but can only communicate with her through an intermediary that is biased to the left.

In the case where the expert and the decision-maker are linked through only one path, I show that the state is transmitted precisely in equilibrium if and only if the path has no bias reversals. Only then, in fact, do all players pass on the expert's information precisely along the path, and Milgrom's (1981) logic extends to communication through intermediaries. The reason is that there cannot exist off-path beliefs that simultaneously punish the expert and all intermediaries for withholding information, if some are biased to the right and others are biased to the left. To punish the former, the decision-maker's choice should be as leftist as possible, compatibly with the information received, and simultaneously as rightist as possible to punish the latter. But of course, this is impossible. Hence, the state is transmitted precisely to the decision-maker in equilibrium if and only if the path from the expert has no bias reversals. If it does, and the biases are large enough, the decision-maker receives no information in equilibrium. In this case, the results of persuasion games without intermediaries are completely overturned.

This simple result bears important implications for the study of communication in networks. Existing analyses based on diffusion models assume that players learn information from the players they are linked with (i.e., their neighbors), possibly randomly, but abstract from the possibility that information is strategically misrepresented.[2] My results show that, when explicitly considering strategic communication, whether or not a player learns another's information does not only depend on the shape of the network, nor only on the characteristics of the pair of players considered. Whether one player's information reaches another crucially depends on the biases and characteristics of *all the players* on the connecting paths. As I explain in Section 2, this implies that existing diffusion models cannot appropriately represent strategic communication in networks, where players may actively withhold or manipulate information based on their biases.

These implications are highlighted by the contrast in terms of the results on network optimality and stability. I find that the unique optimal network is a line in which each player forms a link only with those with the closest bliss points. This ordered line is also implemented as an equilibrium of a game of bilateral link sponsorship based on Myerson (1991), and it is pairwise-stable. These results stand in sharp contrast with fundamental findings in network theory, such as those by Jackson and Wolinsky (1996). If communication is non-strategic and subject to technological constraints such as information decay, they find that the optimal and stable networks are stars. In terms of network centrality, the contrast could not be more extreme. The star is the most centralized minimally connected network, whereas the line is the most decentralized one.

Further, my results hold despite the fact that, unlike in Galeotti, Goyal, and Kamphorst (2006) and Calvó, de Martí, and Prat (2015), for example, the cost to link agents is the same across pairs, and (while this assumption is not needed) the value of every player's information is the same for each other player. The reason why it is optimal that the players with the closest bliss points form links is not based on homophily (McPherson, Miller, and Cook, 2001). Rather, it is that this induces the least costly network where Milgrom's (1981) logic kicks in. All verifiable information is disclosed through the ordered line network in equilibrium because each possible decision-maker would know who to blame if any information were withheld. The signals of every expert biased to the left (right)

[2] As I detail in the literature review, this is so even for models of information diffusion in networks based on Bayesian learning, because they assume that there are no externalities across agents' decisions or that their preferences are aligned. A fortiori, this is the case for models of non-Bayesian learning in networks.

are transmitted through a path where all players are biased to the left (right). If any information about such signals does not reach the decision-maker, she would infer that it must be evidence that would move her decision to the right (left). She reacts by taking the most leftist (rightist) decision possible compatibly with the information received, and that deters all the players to her right (left) from ever withholding information.

The contrast with non-strategic information diffusion models in networks is also reflected in different applications. Diffusion models naturally apply to economic problems where there are no action externalities, or the agents share similar goals. For example, they have been applied to the study of firms and other organizations in business economics, where the various groups in the organization share the same objectives (e.g., profit maximization). As well as the optimality of stars to deal with information decay, hierarchical, centralized networks are regarded as the optimal structure in this context, for reducing the costs of information processing (Radner, 1993; Bolton and Dewatripont, 1994), and for preventing conflicts between subordinates and their superiors (Friebel and Raith, 2004).

By contrast, my paper studies networks where agents have a strategic incentive to mislead each other because each is affected by everyone's decision, and their preferences are not aligned. The main motivation and application of my work are networks of political decision-makers.[3] Social connections are understood to be an important, if not fundamental, feature of politics.[4] While there is now an empirical literature that investigates political networks,[5] systematic theoretical modeling, as provided by network economics, is still in its infancy.

Possibly quintessential in politics are preference divergences across agents, driven by ideological differentiation, which often lead agents to try and mislead each other. It is also common that political agents form alliances with those who are ideologically closest to them, as predicted by my analysis. Indeed, especially career political agents usually have a good sense of each other's ideological views. In this context, the implication of my paper's main finding is that not only do links form among the ideologically closest agents, but this is also optimal as it ensures the transmission of verifiable information across all political agents.

[3] A different application is to organizations such as companies where distinct divisions have divergent strategic objectives, risk attitudes, or discount factors.

[4] This has been recognized as early as the work by Routt (1938).

[5] An exhaustive review is provided by the handbook edited by Victor, Montgomery, and Lubell (2017).

The insights provided here about political networks are fully developed in my companion paper, Squintani (2018). Unlike here, I consider a simple binary signal model. Each agent receives an i.i.d. signal that is Bernoulli distributed given the state of the world. Of special interest is the case in which agents are partitioned into ideologically diverse groups, each composed of agents with similar views.[6] My analysis predicts that these groups optimally organize as stars, and that each center forms a link with the centers of the ideologically adjacent stars. This gives rise to familiar organizational structures: factions of like-minded politicians report to a leader (star center), and faction leaders communicate with each other along ideological lines. While factionalization is often perceived negatively, my analysis shows that it is the least demanding network architecture that ensures the free flow of verifiable information.

As I explain in the companion paper in detail, this analysis can be applied to several different political networks. For example, think about policymakers in different jurisdictions. They face similar policy challenges and would benefit from the experience gathered by their colleagues. Often, this information takes the form of policy experiments, whose results cannot be falsified, but can be hidden in the abundance of bureaucratic documents. To make better decisions, each policymaker consults her network of peers, but these connections are costly to maintain.[7] Likewise, members of legislative assemblies greatly value communication within informal networks.[8] Other examples include networks of interest groups, lobbyists, political experts, consultants, investigative journalists, bloggers, or academics.[9] When one of these agents is called to campaign or provide advice on different matters, she will reach out through her peer network to collect information. But again,

[6] As posited by Hume (1741), and documented since at least Rose (1964), these distributions of ideologies arise naturally in several political environments.

[7] The empirical/descriptive literature on networks of policymakers includes work by Zafonte and Sabatier (1998) on the cooperative links among San Francisco Bay area government agencies, and the study by Gerber, Henry, and Lubell (2013) on the connections among planning and management agencies in five main California regions.

[8] In a classical piece, Fiellin (1962) surveys 1961 Democrat representatives about their network of friends in the New York State House, and documents that "probably the most important functions of informal groups and relationships result from their use as communication networks." This work led to many studies, e.g., Zhang et al. (2008), and Bratton and Rouse (2011).

[9] Among studies on interest group networks, Laumann and Knoke (1987) and Carpenter, Esterling, and Lazer (2004) trace the connections among lobbyists, government agencies, and congressional staff in the 1970s health and energy policy domains, König and Bräuninger (1998) investigate the ties among interest groups, trade unions, governmental agencies, and legislators in the events that shaped 1980s German labor policies, Koger, Masket, and Noel (2009) study the network of US party candidates, activists, interest groups, and media outlets using donor and subscriber name data, and Box-Steffensmeier and Christenson (2014) build a network of interest group coalitions using *amicus curiae* briefs to the US Supreme Court.

these connections are costly to maintain. Furthermore, my theoretical analysis can also be related to the study of informal networks of judges, due to the available network datasets built by linking judges that cite other judges' rulings as precedents in their decisions.[10]

2 Related Literature

Within the large body of literature developed over the years in network theoretical economics,[11] an important focus is information transmission in networks.[12] The diffusion approach usually adopted applies to communication among agents who do not have any incentive to strategically mislead each other. This is also true for so-called "Bayesian learning" models, where each player learns through equilibrium inference based on their neighbors' choices, either in the form of explicit information transmission or just by observing their actions. This is because such models assume that there are no externalities across agents' play and that their preferences are aligned.[13]

Specifically, Bala and Goyal (1998) studied a pioneering model in which there is no explicit information transmission; players observe the choices of their neighbors and thus indirectly learn about information (see also Acemoglu, Dahleh, Lobel, and Ozdaglar, 2011). Acemoglu, Bimpikis, and Ozdaglar (2014) provided an explicit model of information transmission. While in these papers, information originates only among the players in the network, Egorov and Sonin (2020) study a model where a biased outside sender may also send a signal to the players. In line with the "Bayesian persuasion" model by Kamenica and Gentzkow (2011), the sender controls the informativeness of her signal but cannot falsify it, and information flows through the network with decay.[14] Instead, Galperti and Perego

[10] For example, Caldera (1985) constructs the network of citations across US State supreme courts, and Fowler et al. (2007) the network of citations across all three levels of US federal courts.

[11] Such literature is so vast that it is impossible to survey adequately here. A detailed review of network economics literature can be found in the handbook edited by Bramoullé, Galeotti, and Rogers (2017), for example.

[12] Jackson and Yariv (2011) review the literature from both economics and other fields on the role of externalities and diffusion in networks.

[13] A fortiori, strategic information transmission in networks is covered by so-called "naive" models where players' learning is not based on equilibrium beliefs in a game. For example, Golub and Jackson (2010) study a model where players' beliefs are assumed to be weighted averages of their neighbors'.

[14] Although there is no strategic information transmission within the network due to aligned preferences, the outside sender's communication is strategic, and so is the decision within the network on whether to listen. As a result, the outside sender cannot reach high-centrality agents, who prefer to rely on the information from their network neighbors.

(2024) consider the case where information does not decay through the network, and only a subset of networked players can receive the sender's information.

In all these papers, there are no externalities across players' choices in the network, and their preferences are aligned. Hence, the players in the network have no incentive to strategically manipulate information and mislead each other, which is the research subject of this paper. Further, a maintained assumption of existing network models is that the value of connecting a pair of agents i and j is independent of the characteristics of the other players on the connecting path. This assumption is shown not to hold in this paper's full-fledged analysis of strategic communication. And while this is shown here for the case of verifiable information transmission, Ambrus, Azevedo, and Kamada (2013) prove an analogous result for the case of cheap talk. These results clarify why existing network models do not cover strategic information transmission.[15]

Building on the seminal papers by Milgrom (1981), Grossman (1981), and Crawford and Sobel (1982), strategic information transmission has developed into one of the central topics in the economics of information.[16] One of the main insights of this literature is that, while communication of unverifiable information (cheap talk) generally leads to imprecise decisions, withholding verifiable information is incompatible with equilibrium.[17] This insight is overturned in my analysis of indirect communication. When an expert communicates with a decision-maker through intermediaries, verifiable information is fully disclosed in equilibrium if and only if all the players on the communication path are biased in the same direction relative to the decision-maker.

While the literature on strategic information transmission has branched out theoretically in several directions, the study of indirect communication through intermediaries is still underdeveloped.[18] Ambrus, Azevedo, and Kamada (2013) study the case of cheap

[15]Sadler (2020) and Lipnowski and Sadler (2019) do not consider information transmission in networks, but each player's action may signal information to her neighbors in the network. Action externalities may take the form of strategic complementarity or substitution, but they are limited to the choice of neighbors, and so is the incentive to mislead neighbors through action choice.

[16]Such literature is so vast that it is impossible to survey adequately here. Strategic communication models have been applied to several different fields of research, including political economy (e.g., Gilligan and Krehbiel, 1987; Morris, 2001), accounting (e.g., Verrecchia, 1983; Dye, 1985; Skinner, 1994), organization design (e.g., Dessein, 2002; Alonso, Dessein, and Matouschek, 2008), contract theory and industrial organization (e.g., Vives 1984; Lizzeri 1999; Dewatripont and Tirole, 1999).

[17]This result is extended to a model in which the sender has a cost for lying by Kartik, Ottaviani, and Squintani (2007) and Kartik (2009).

[18]In fact, most studies of communication are staged in 2-player models with one expert and one decision-

talk, demonstrating that the analysis is significantly more complex than for the case of verifiable information disclosure considered here. The result that intermediation cannot improve information transmission holds only for pure strategy equilibria. They provide a partial characterization of mixed strategy equilibria, show instances in which intermediation improves upon direct communication, and provide necessary conditions. For the case of verifiable information disclosure I consider here, however, intermediation cannot ever improve upon direct communication, in equilibrium.

Further, Ivanov (2010) considers the intermediation of cheap talk through a strategic mediator, whereas Patty and Penn (2014) explore sequential decision-making in networks of three privately informed agents with misaligned preferences. There is no explicit information transmission in their model, but each agent's action may signal information to subsequent decision-makers in the network. While I consider the transmission of multi-agent (verifiable) information in networks allowing for any profile of players' biases, Bloch, Demange, and Kranton (2018) study the spreading of possibly false information in a network where agents may either wish correct decisions are made or have a private agenda in favor of one alternative. Unlike my paper, they do not consider network optimality. Instead, Migrow (2019) studies optimal hierarchies of possibly biased agents who strategically report unverifiable information to a single decision-maker.

Closer to my work, Gieczewski (2022) studies a model of learning in networks with the transmission of verifiable information among agents with misaligned preferences. Unlike my paper, he does not consider network formation, stability, or optimality. Experts learn the state with probability less than one, and this prevents full information disclosure. He finds that full learning requires sufficiently dense networks in his framework. Signals closer to the mean are more likely to propagate because agents tend to block signals contrary to their bias, as is the case in my paper. When agents are forward-looking, concerns about learning cascades cause the players to divide into like-minded, non-communicating groups.

Unlike these papers, Onuchic and Ramos (2023) do not consider indirect communication, nor networks. They study team production where output is verifiable, and its disclosure may impact team members differently. They consider all disclosure protocols,

maker. Among exceptions are the studies by Battaglini (2002) on many-to-one communication, Farrell and Gibbons (1989) on one-to-many communication, and Galeotti, Ghiglino, and Squintani (2013) on many-to-many communication. None of these or other papers on multi-player communication consider indirect communication through intermediaries.

from unilateral to consensual disclosure, and find that full disclosure is the unique outcome only if team members can unilaterally disclose information on output. This result is similar in spirit to my findings on disclosure through intermediaries. They characterize productive environments where different protocols maximize effort incentives in teams, finding that partial disclosure can boost incentives.

3 The model and preliminaries

This Section provides a simple model of persuasion in networks by putting together the early model of persuasion in games by Milgrom (1981) with the model of "social communication" networks by Jackson and Wolinsky (1996). I begin the exposition by briefly reviewing these models and their most important results, and by setting up a common notation.

Jackson and Wolinsky (1996) I now turn to introduce the model of "social communication" in networks by Jackson and Wolinsky (1996). Such model is a reduced-form model in which there is no explicit information transmission. Rather, each player i's payoffs for a possibly indirect connection with another player j are interpreted as the value of j's information for i. Specifically, suppose that a set $\mathcal{N}$ of n players is connected in a symmetric network N, a symmetric $n \times n$ matrix, where $N_{ij} \in \{0, 1\}$, $N_{ii} = 1$ for all $i, j \in N$. The utility of each player i from graph N is:

$$u_i(N) = \sum_{j \neq i} (\delta^{\ell(i,j)} - cN_{ij}),$$

where $\ell(i, j)$ is the length of shortest path between i and j in N.[19] Each link describes who transmits information to whom. Each link costs $c \geq 0$ to player i, and the parameter $\delta \in (0, 1]$ represents "communication decay." The idea is that player j may hold information valuable to player i. When this is the case, i's information travels to j along one of the shortest paths that connects them. At each step on the path, the information decays with probability $1 - \delta$. So, j obtains i's information with probability $\delta^{\ell(i,j)}$.

Network optimality trades off how well information is communicated through the network with the cost the players pay to form links. A network N is optimal if it maximizes

[19] Two agents i and j are linked by the path $p = (i, h_1, ..., h_{\ell-1}, j)$ of length ℓ in network N, if i is linked to h_1, h_k is linked to h_{k+1} for every $k = 1, ..., \ell - 2$, and $h_{\ell-1}$ is linked to j.

$w(N) = \sum_i u_i(N)$. Network optimality is characterized through the following definitions. The empty network is such that there does not exist a link between any player i and j. The complete network is such that every player i is linked with every other player j. A minimally connected network or tree N is a network such that every pair of players i, j is connected via a unique path. A star is a minimally connected network in which one player, called the center, is connected to all other players.

Proposition 1 (Jackson and Wolinsky 1996) *There exist thresholds $\underline{c}$ and $\overline{c}$ functions of δ, and n such that: for large link cost c, i.e. $c > \overline{c}$, the optimal network is empty; for small c and $\delta < 1$, i.e. $c < \underline{c}$, the optimal network is complete; for intermediate cost c and $\delta < 1$, i.e. $\underline{c} < c < \overline{c}$, every star is optimal; for $\delta = 1$ and $0 = \underline{c} < c < \overline{c}$, every tree is optimal.*

The logic behind these results is easy to grasp. When forming links is too costly, it is optimal not to connect players at all, and the optimal network is empty. If instead forming links is very cheap, then it is optimal to link all players in a complete graph, as long as there is decay in communication $(\delta < 1)$. In the case links are neither too costly nor too cheap, it is optimal to connect all players with a network without redundant paths, i.e., with a tree. In the presence of decay, $\delta < 1$, the optimal graph is the star as it is the tree that minimizes the sum $\sum_{i,j} \ell(i,j)$ of the lengths $\ell(i,j)$ of the paths across every pair of players i and j.

Jackson and Wolinsky (1996) further prove that for intermediate link costs c and $\delta < 1$, every star is "pairwise stable." Informally, stars are the networks with the property that no link forms that would not be worth the cost to one of the connected player, and every link forms that would be beneficial to both the connected players.[20] The main message of their analysis is that optimality and equilibrium stability select highly centralized network such as the stars, when avoiding information decay is the most important concern in the design of communication networks. Later, I will show how strategic communication incentives overturn these insights.

Milgrom (1981) The model of persuasion by Milgrom (1981) considers a biased, informed expert e who may or may not disclose his information to a decision maker d. The

[20] To avoid the burder of too many preliminaries at this stage, I postpone the formal definition of pairwise stability to the next section.

information disclosed is verifiable, hence the expert cannot lie, but can withhold information. Specifically, there is a state of the world $x \in X = [\underline{x}, \overline{x}] \subset \mathbb{R}$.[21] The decision maker does not know x, she only knows the distribution of x, which I assume to have a continuous density f strictly positive on X. The expert knows x, and may disclose information about x, in the form of a closed set $\hat{m} \subseteq X$, which may be a singleton set. The information $\hat{m}$ is verifiable, i.e., the expert is restricted to send a message $\hat{m}$ such that $x \in \hat{m}$. After receiving $\hat{m}$, the decision maker chooses an action $\hat{y} \in \mathbb{R}$.

The decision maker's payoff is the loss function $L_d(x, \hat{y}) = -(\hat{y} - x)^2$: she would like to match her action $\hat{y}$ with the state x. Relative to the decision maker, the expert is biased: he would like that $\hat{y} > x$. Milgrom (1981) assumes that the expert's loss function $L_e(x, \hat{y})$ is strictly increasing in $\hat{y}$ for all x. For reasons that will become evident later, I make the less extreme assumption that the expert would like that action $\hat{y}$ is matched with $x + b$, where $b > 0$ measures the expert's bias. Specifically, his payoff is $L_e(x, \hat{y}) = -(\hat{y} - x - b)^2$. The extreme case considered by Milgrom is captured by $b > \overline{x} - \underline{x}$.

The main result of the analysis is:

Proposition 2 (Milgrom 1981) *There is a unique (pure and mixed strategy) perfect Bayesian equilibrium, in which the biased expert e reveals x precisely,* $\hat{m} = \{x\}$, *and the decision maker d chooses* $\hat{y} = \min \hat{m}$ *for every set* $\hat{m}$ *disclosed by the expert. Hence, on the equilibrium path,* $\hat{y} = x$.

Despite her bias, the expert always discloses the state x precisely in equilibrium, and the decision maker acts fully informed. The equilibrium is supported by "worst-case" beliefs off the equilibrium path. If the expert did not disclose x precisely, i.e., she sent a closed set $\hat{m}$ that does not coincide with x, the decision maker would believe that the state equals $\min \hat{m}$, and pick action $\hat{y} = \min \hat{m}$. But because information is verifiable, $x \in \hat{m}$, it follows that $\min \hat{m} \leq x$ whenever $\hat{m} \neq \{x\}$. The expert does not gain by deviating from the equilibrium strategy of sending $\hat{m} = \{x\}$.[22]

[21] In Milgrom (1981), the state space X consists of the positive reals. For reasons that will become clear later, it is useful to let X be a closed interval, here.

[22] The result that there is no other equilibrium is based on the following so-called "unraveling argument." Suppose to the contrary that the decision maker chose $\hat{y} = E[x|\hat{X}]$ whenever $x \in \hat{X}$ for some non-degenerate set $\hat{X} \subseteq X$. Because the distribution of x is full support, it must be that $E[x|\hat{X}] < \max \hat{X}$. But then, for every state x such that $E[x|\hat{X}] < x \leq \max \hat{X}$, the expert would prefer to send message $\hat{m} = \{x\}$ and disclose x precisely, so that $\hat{y} = x > E[x|\hat{X}]$, rather than following the equilibrium message strategy and obtain $\hat{y} = E[x|\hat{X}]$.

In sum, the received wisdom of persuasion games where an informed expert directly transmits verifiable information to a decision maker is that all information is disclosed in equilibrium. I will later show how this insight is overturned when considering transmission of verifiable information through intermediaries.

Persuasion in networks In order to study strategic information transmission in networks I embed the persuasion game by Milgrom (1981) into the social communication construction by Jackson and Wolinsky (1996). As a by-product, I also provide a microfoundation of social communication networks with an explicit model of information transmission.

Again, suppose that a set $\mathcal{N}$ of n players is connected in a symmetric network N, where $N_{ij} = 1$ means that i is linked to j, i.e., i can transmit verifiable information to j. After the network N is formed, a pair of players d and e is randomly selected from $\mathcal{N}$, according to a full support probability P. Player d takes the role of the uninformed decision maker and e of the informed expert, she is the unique player who knows the state of the world $x \in X$. Their identities becomes common knowledge among all players in $\mathcal{N}$ after they are selected.

As in Jackson and Wolinsky (1996), player e's information travels to d through the network N. But I generalize the construction by allowing information to travel along any path $p(e, d)$ that connects e and d. As in Milgrom (1981), information is verifiable and takes the form of closed sets. Specifically, letting $\bar{\ell}(e, d)$ be the length of the longest path p from e to d, there are $T = \bar{\ell}(e, d)$ periods of information transmission. At time $t = 0$, player e transmits a message $\tilde{m}^0_{ej}$ to any one of her neighbors j who belong to a path p from e to d.[23] The expert's message $\tilde{m}^0_{ej}$ is a closed subset of X and it is verifiable, in the sense that $\{x\} \subseteq \tilde{m}^0_{ej}$.

Then, for $t > 1$, communication proceeds as follows. For any path p from e to d and any player i on p, denote by the 'distance' $\ell(p(e, i))$ between i and e on p, the length of the subpath $p(e, i)$ from e to i contained in p. At any time $t = 1, ..., T - 1$, each player i on a path p from e to d at distance t from e sends a message $\tilde{m}^t_{ij} \subseteq X$ to any one of her neighbors j who are on a path p' from e to d and are at distance $t + 1$ from e. Again, for each such i and j, the message $\tilde{m}^t_{ij}$ is closed and verifiable. Specifically, for every history h^t, letting

[23] The set N_i of neighbors of i in N is the set of players j with whom i is linked, i.e., such that $N_{ij} = 1$.

$\omega_i(h^t)$ be the information held by player i about the state x, the verifiability requirement is that $\omega_i(h^t) \subseteq \tilde{m}^t_{ij}$. In the first period, given the null history h^0, the information of the expert is $\omega_e(h^0) = \{x\}$ and every other player i's information is $\omega_i(h^0) = X$.

With probability $1-\delta$ independently across periods and pairs of players i, j, the message $\hat{m}_{ij}$ is lost in transmission, and player j learns nothing about X, i.e., she observes $\hat{m}^t_{ij} = X$. Else, j observes $\hat{m}^t_{ij} = \tilde{m}^t_{ij}$. I restrict attention to small or no decay, δ close or equal to 1. Verifiable information $\omega_i(h^t)$ is updated in the standard manner: if a player j observes the vector of messages $\hat{\mathbf{m}}_j$ at history h^t, she updates her information at history h^{t+1} according to the rule: $\omega_j(h^{t+1}) = \omega_j(h^t) \cap_i \hat{m}^t_{ij}$. At time T, given any history h^T, player d chooses $\hat{y}_d \in \mathbb{R}$ on the basis of her verifiable information $\omega_d(h^T)$ and on equilibrium beliefs.

In line with my specification of persuasion games, each player i would like that any decision maker d's choice $\hat{y}_d$ matches her realized bliss point $x+b_i$. The "relative bliss point" $b_i \in \mathbb{R}$ identifies player i's idiosyncratic preference component, relative to the common state x. For the sake of realism, I allow for each player i to care about some decisions more than others. Player i's loss function for decision $\hat{y}_d$ is

$$L_i(\hat{y}_d, x) = -\alpha_{id}(\hat{y}_d - x_d - b_i)^2,$$

where the utility weights α_{id} are such that $\alpha_{id} > 0$ for all i, d and $\sum_{d \in \mathcal{N}} \alpha_{id} P(d) = 1$ for all d. I assume that the ex-ante bliss points are ordered, $b_1 < ... < b_n$, and that they are common knowledge, in line with the motivating applications presented in the Introduction.[24]

Given N, e and d, I let μ_{edN} denote a possibly mixed equilibrium strategy profile of the players $i \neq d$ on any path p from e to d, and y_d the associated equilibrium strategy of player d. Hence, each player i's expected value of equilibrium (μ_{edN}, y_d) in the persuasion game given by network N, expert e, and decision maker d is:

$$u_i(\mu_{edN}, y_d; e, d, N) = -\alpha_{id} E[(y_d(h^T; \mu_{edN}) - x - b_i)^2],$$

where the expectation is taken over h^T and x. In case multiple equilibria μ_{edN} exists in a persuasion game defined by a triple N, e and d, I select the equilibrium μ_{edN} that yields the highest expected value $u_i(\mu_{edN}, y_d; e, d, N)$ to all players i (I will later show that the

[24] While these assumptions capture multi-player persuasion games, the "opposite" case where $b_1 = ... = b_n \equiv b$ provides a simple micro-foundation of the social communication games by Jackson and Wolinsky (1996).

equilibria μ_{edN} are Pareto ranked ex-ante for all e, d, N). I define such equilibrium μ^*_{edN} as the 'most informative equilibrium.'

Aggregating across the possible realizations of expert-decision maker pairs (e, d), we obtain that each player i's ex-ante value of network N, including the costs of the links in N, is:

$$U_i(N) = -\sum_{(d,e)\in\mathcal{N}^2:d\neq e} \alpha_{id} E[(y_d(h^T; \mu^*_{edN}) - x - b_i)^2] P(e, d) - c\sum_{j\neq i} N_{ij}.$$

As in social communication game presented above, I let the welfare of each network N be simply the sum of the players' ex-ante values: $W(N) = \sum_i U_i(N)$. Later, I will determine the relationship between such utilitarian welfare concept and the balance between a network N's information transmission efficiency and aggregate link costs.

4 Results

First, I solve for perfect Bayesian equilibrium in the persuasion game with fixed network N. Then, I will determine optimality and stability. To highlight my point of departure from models of information diffusion in networks, I here abstract from any technological communication constraint and focus on the case of no decay, $\delta = 1$. Also, to simplify the exposition, I restrict the players' messages space to rule out partial information disclosure. Formally, for any N, e and d, at any time t in which a player $i \neq d$ on a path $p(e, d)$ is called to play, she can only send a message $\hat{m}^t_{ij} \in \{\{x\}, X\}$ to her immediate successor j on $p(e, d)$. That is, player i can either disclose x precisely or withhold all information (and obviously, the former is feasible only if she knows x). To avoid ambiguity, I refer to the resulting communication game from e to d as a "disclosure game," reserving the term "persuasion game" for when partial disclosure is also allowed.

Equilibrium in the Disclosure Game I begin with two simple results that are immediate consequences of the players' quadratic loss specifications. For any d, e, x, in every equilibrium μ of the disclosure game, in all histories h^T, the decision maker d plays

$$y_d(h^T) = E[x|\omega_d(h^T), \mu] + b_d,$$

and hence each player i's ex-ante expected equilibrium payoff is:

$$-\alpha_{id}E(y_d(h^T) - x - b_i)^2 = -\alpha_{id}E[Var(x|\omega_d(h^T), \mu)] + (b_i - b_d)^2. \quad (1)$$

The decision maker d chooses a decision equal to the expected value $E[x|\omega_d(h^T), \mu] + b_d$ of her bliss point $x + b_d$ given her information $\omega_d(h^T)$ at history h^T, together with knowledge of equilibrium strategies μ. As a result, the expected loss $E(y_d(h^T) - x - b_i)^2$ of each player i can be decomposed into the expected residual variance $E[Var(x|\omega_d(h^T), \mu)]$ of x given the information $\omega_d(h^T)$ and the equilibrium strategies μ, and the square of the bias $b_i - b_d$ of player i relative to the decision maker d. It follows that the players' ex-ante payoffs are aligned in any disclosure game determined by any N, e, d. They would all prefer to minimize d's expected residual variance $E[Var(x|\omega_d(h^T), \mu)]$, i.e., that d makes her decision with as precise information $\omega_d(h^T)$ as possible.

Momentarily restricting attention to minimally connected networks N, I now characterize the equilibrium of the disclosure game on N given realized expert e and decision maker d. I make use of the following concept. Say that *a path $p(e, d)$ from e to d has a bias reversal if there exists i, j on $p(e, d)$ such that $b_i < b_d < b_j$.* The next result shows that verifiable information flows along N if and only if the path that connects e to d does not have bias reversals. And if such bias reversals are sufficiently large, then no information reaches the decision maker. In other terms, the result of Milgrom (1981) is fully overturned when verifiable information is transmitted through intermediaries i with biases $b_i - b_d$ that are opposed to the expert's bias $b_e - b_d$.

Proposition 3 *In any equilibrium of a disclosure game on a tree N, for any expert e and decision maker d, letting $T = \ell(e, d)$ be the length of the unique path $p(e, d)$ from e to d,*

1. if the path $p(e, d)$ has no bias reversals, then the decision maker d always learns x and plays $\hat{y} = x$, so that $E[Var(x|\omega_d(h^T), \mu)] = 0$,

2. if $p(e, d)$ has bias reversals, then d does not always learn x: $E[Var(x|\omega_d(h^T), \mu)] > 0$;

3. if $b_d - b_i > 0$ and $b_j - b_d > 0$ are large enough, then d acts with no information: $y_d(h^T) = E[x] + b_d$ for all h^T.

These result are quite intuitive. Say $p(e, d)$ has no bias reversals, e.g. $b_i > b_d$ for all $i \neq d$. Then, the logic of Milgrom (1981) generalizes. There exists a perfect Bayesian equilibrium such that every i relays all information she has about x precisely along the path $p(e, d)$, formally $\hat{m}_{ij} = \omega_i(h^t)$ for all $i \neq d$ and history h^t at which i is called to play. As a result, the decision maker d learns x precisely and plays $\hat{y}_d = x$, so that $E[Var(x|\omega_d(h^T), \mu)] = 0$. Such equilibrium is supported by "worse-case" off-path beliefs that interpret vague information as evidence contrary to agents biases, and hence "punish" players' for withholding information. Formally, such beliefs assign probability one to $x = \min \omega_j(h^t)$ for every agent $j \neq e$ and every history h^t at which j is called to play. Due to these off-path beliefs, each player $i \neq d$ can only shift the decision $\hat{y}_d$ to the left, contrary to her bias $b_i - b_d > 0$, by withholding information at any history h^t where she is called to play.[25] Therefore, withholding information only reduces her payoff, and she will choose to fully disclose in equilibrium.

Further, it cannot be that $E[Var(x|\omega_d(h^T), \mu)] > 0$ in equilibrium, by the following generalization of the standard "unravelling" argument. If such an equilibrium existed, it would need to be the case that d does not know x for some history h^T. In the set of states $X(h^T)$ that d considers possible at h^T, there would exist states x sufficiently close to the upper bound of $X(h^T)$ such that all players $i \neq d$ would like to reveal the state x precisely, as this moves the decision $\hat{y}_d$ to the right, relative to the decision based only on the information that the state is in $X(h^T)$, and on equilibrium beliefs.

Instead, both the "worse-case beliefs" and the "unravelling" arguments break down if the (unique) path $p(e, d)$ from the expert e to the decision maker d has bias reversals. Plainly, there cannot exist off-path beliefs, and consequent decisions $\hat{y}_d$, that simultaneously punish players i biased rightward, $b_i > b_d$ and players j biased leftward, $b_j < b_d$, for withholding information. There is no equilibrium such that player d learns x precisely, and hence the loss $E[Var(x|\omega_d(h^T), \mu)]$ is strictly bounded above zero (the precise bound is derived in the proof of Proposition 3, in Appendix).

To elaborate, first note that every player $i \neq d$ on the path $p(e, d)$ can withhold all information about x from all successors by sending the message $\hat{m}_{ij}^t = X$ to her immediate successor j at the time t when she is called to play. When this happens, the decision maker will not learn anything about x; that is, her information set will be $\omega_d(h^T) = X$, and she will play $\hat{y}_d(X) = E[x|X, \mu] + b_d$. Consider any equilibrium belief $E[x|X, \mu] > \underline{x}$, and suppose $x \geq \underline{x}$ is below $E[x|X, \mu]$ but less than $b_i - b_d$ away from $E[x|X, \mu]$. Every player i with $b_i > b_d$ who knows x strictly prefers that d play $\hat{y}_d(X) = E[x|X, \mu] + b_d$ rather than $\hat{y}_d(\{x\}) = x + b_d$. Hence, she prefers to send $\hat{m}_{ij}^t = X$ rather than reveal x precisely. Similarly, for any equilibrium belief $E[x|X, \mu] < \overline{x}$, every player i with $b_i < b_d$

[25] This is because player i cannot lie about her information, because all that d knows about x she learnt it from the communication path $p(e, d)$, and because all successors of i on $p(e, d)$ will relay their information precisely along the path, formally $\omega_i(h^t) \subseteq \hat{m}_{ij} = \omega_d(h^T)$ for all $i \neq d$, and hence $\min \hat{m}_{ij} \leq \min \omega_i(h^t)$.

strictly prefers to withhold x because she is better off if d plays $\hat{y}_d(X)$ instead of $\hat{y}_d(\{x\})$, for any $x \in (E[x|X, \mu], \min\{\overline{x}, E[x|X, \mu] - b_i + b_d\})$. Thus, regardless of the value assigned to $E[x|X, \mu]$, there always exists an interval of states x that is not precisely revealed to the decision maker in equilibrium.

Further, if the player's biases $b_i - b_d > 0$ and $b_j - b_d < 0$ are large enough, then, player i just wants to maximize $\hat{y}$, and player j to minimize $\hat{y}$, regardless of the value of the state x. Indeed, such a case of "transparent motives" as defined in Lipnowski and Ravid (2020) is the case originally studied by Milgrom (1981). Further, every player i on the path $p(e, d)$ can force the decision $\hat{y}_d = E[x] + b_d$ simply by withholding all information (i.e., by sending $\hat{m}_{ij} = X$) from his neighbor j and hence from all her successors including player d. So, if the players' biases on path $p(e, d)$ are sufficiently large, then the only possible equilibrium outcome is $\hat{y}_d = E[x] + b_d$. The decision maker acts as if she receives no information, and the expected loss is as if there were no information transmission: $E[Var(x|\omega_d(h^T), \mu)] = Var(x)$.

Proposition 3 bears some important implications for the study of information transmission in networks. With few exceptions (see Section 2), the study of information diffusion in networks abstracts from strategic communication incentives. This is so even in 'Bayesian' learning models. Players learn from each other with some probability as long as they are connected. So, the value of a path that connects two players i, j depends only on the (possibly weighted) length, and possibly on characteristics of i, j (as in Galeotti, Goyal and Kamphorst, 2006, for example). My full fledged model of strategic communication shows that the value of a path connecting e and d depends also on characteristics of all the players $i \neq e, d$ on the path $p(e, d)$. Hence, network analysis based on information diffusion models cannot cover strategic communication.

The next part of the Section shows how Proposition 3 leads to my characterization of network optimality.

Optimal Network I now deliver the most important result of the paper: The unique optimal network N of my model is the line in which the players are ordered according to their bliss points, which I define as the 'ordered line.' The statement is provided for meaningful links costs, i.e., costs c that are (i) strictly positive, and (ii) not so large that the optimal network would not be connected even in the hypothetical case that information flowed unconstrained in the network, and strategic disclosure did not matter. Formally, I define the cost threshold $\bar{c}$ as the largest cost c, function of the probability P and weights α, such that the optimal N would be connected if every d received x along every path $p(e, d)$ from every e.

Proposition 4 *For any cost $c \in (0, \bar{c})$, utility weights $\boldsymbol{\alpha} > \mathbf{0}$, and full support recognition probability P, the unique optimal network N is the ordered line.*

The proof of Proposition 4 is relatively simple and for this reason, I present it here in the main body. It consists in showing that the ordered line is the unique minimally connected network such that the signal s of every possible expert e is relayed precisely to every possible decision maker d, through the unique path $p(e, d)$ that connects them. In every other tree N, there exists at least one pair e, d that is connected through a path $p(e, d)$ with bias reversals.[26] As a consequence, the ordered line is the unique optimal tree. Because the link cost $c < \bar{c}$, it is suboptimal to consider networks that are not connected, and because $c > 0$, non-minimally connected networks are also suboptimal.

Proof of Proposition 4. Suppose momentarily the optimal network is a tree. Evidently, the ordered line has no bias reversal paths. For any pair of realized decision maker d and expert e with $b_e < b_d$, it is the case that $b_i < b_d$, for each player i on the path $p(e, d)$ from d to e, and vice versa when $b_e > b_d$. By Proposition 3, the state x is always relayed precisely along the path $p(d, e)$ regardless of the realized identities of players d and e.

Consider every other tree N, focus momentarily on $n = 4$ players. Interchanging the players' identities, there are only two classes of trees: the lines and the stars. Of course, every non-ordered line contains bias reversal paths. Every 4-player star has at least a bias reversal path. When the star's centre is player $i = 1, 2$, that is the path from $e = i + 2$ to $d = i + 1$, see Figure 1. Symmetrically, when the center is $i = 3, 4$, the path from $e = i - 2$ to $d = i - 1$ has a bias reversal.For any number of players $n \geq 4$, the only tree that does not contain a 4-player star or a non-ordered line is the n-player ordered line. (The same is true, trivially, for $n = 2$, and also for $n = 3$, where stars and lines coincide.) So, for every n, because $\boldsymbol{\alpha} > \mathbf{0}$, the unique tree N in which x is transmitted precisely for every realized

[26] Note that this is not an immediate consequence of the concept of bias reversal, as it is easy to find paths such as $p = (2, 1, 3)$ in which the players are not ordered and yet there are no bias reversals.

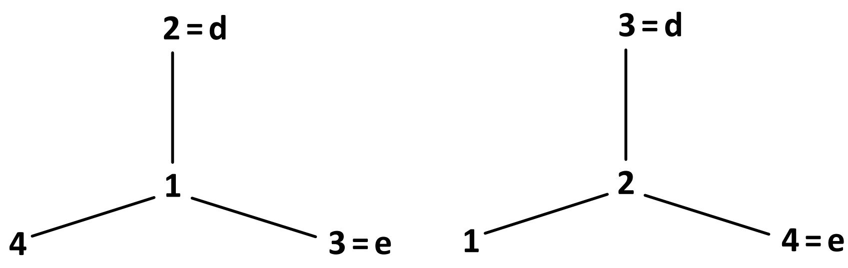

Figure 1: 4-player stars

d, e is the ordered line. By Proposition 3, the ordered line is the unique tree N such that for all e, d, the expected variance of d's decision is $E[Var(x|\omega_d(h^T), \mu)] = 0$.

Using the mean-variance decomposition (1), write welfare

$$W(N) = -\sum_{(e,d):e\neq d} nE[Var(x|\omega_d(h^T), \mu)]P(e,d) - \sum_{i\in\mathcal{N}} \sum_{(e,d):e\neq d} (b_i - b_d)^2 P(e,d) - \sum_{i\in\mathcal{N}} \sum_{j\neq i} N_{ij}c$$

The aggregate link cost is $\sum_{i\in\mathcal{N}} \sum_{j\neq i} N_{ij}c = (n-1)c$ in every tree, hence the tree N that maximizes the welfare $W(N)$ is the tree that minimizes the sum of expected losses $\sum_{(e,d):e\neq d} nE[Var(x|\omega_d(h^T), \mu)]P(e,d)$.Therefore, the ordered line is the unique optimal tree for any full support probability distribution P.

Now, let's consider networks that are not minimally connected. Because $c > 0$, adding links to the ordered line is wasteful. By definition of the cost threshold $\bar{c}$, unless $c > \bar{c}$, deleting links from the ordered line is detrimental: each realized expert e's information is useful to every realized decision maker d.

This concludes that the ordered line is the unique optimal network. ■

Endogenous Network Formation I now turn to considering what networks would form endogenously as the equilibrium of a game in which individual players pay the cost of their links. As is the case for the welfare analysis, also the network formation game is formulated ex-ante, i.e., before the identities of the expert e and decision maker d are drawn, and before the state x is realized.

Specifically, I model network formation as a 'bilateral sponsorship' game a-la Myerson (1991), in which both linked players i and j need to pay their cost c for any link N_{ij} to form. For any fixed recognition probability P and utility weights $\boldsymbol{\alpha}$, the network N is formed as

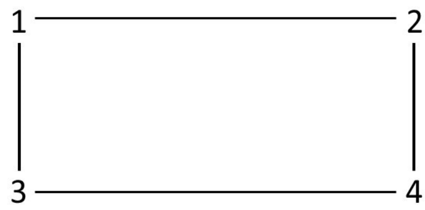

Figure 2: pairwise stable circle

follows. The players $i \in \mathcal{N}$ simultaneously submit a list $\ell_i \in \{0,1\}^{\mathcal{N}\backslash\{i\}}$ of players $j \neq i$ they wish to link to at cost c, where $\ell_{ij} = 1$ means that i commits to pay. For every pair of players i and j, it is the case that $N_{ij} = N_{ji} = 1$ if and only if both players commit to pay: $\ell_{ij} = 1 = \ell_{ji}$. Once the network N is formed, an expert e and decision maker d are drawn, the state x is realized and revealed to e, and the disclosure game is played on the network N. As is fairly obvious, the 'bilateral sponsorship' game is such that players can mis-coordinate in Nash Equilibrium: a player i can play $\ell_{ij} = 0$ only because she knows that j also plays $\ell_{ji} = 0$ although the link between i and j would benefit both in the disclosure game. I focus the analysis on Nash equilibria networks N with the extra requirement that if $N_{ij} = N_{ji} = 0$, then at least one among players i and j would gain less than c if forming a link, i.e., $U_i(N + \tilde{N}_{ij}) < U_i(N)$ or $U_j(N + \tilde{N}_{ji}) < U_j(N)$, where $\tilde{N}_{ij}$ is a $n \times n$ matrix in which all entries $\tilde{n}$ are equal to zero, with the exception of $\tilde{n}_{ij} = \tilde{n}_{ji} = 1$. I call such Nash equilibria 'pairwise stable' in line with a terminology introduced by Jackson and Wolinsky (1996).

With these definitions, I can now deliver my main result on endogenous network formation. Unless the link cost c is too high, the ordered line is the unique pairwise stable minimally connected network. Other pairwise stable networks exist, but they must have loops. As demonstrated in the Appendix, a minimal example of a pairwise stable network different from the ordered line is a 4-player is the circle portrayed in figure 2, in which both 1 and 4 are linked with 2 and 3.

Proposition 5 *For every utility weights $\boldsymbol{\alpha} > \mathbf{0}$, and full support recognition probabilities P, there exists a cost threshold $\hat{c}$ such that for all $c \in (0, \hat{c})$, the ordered line is a pairwise stable Nash network, there are other pairwise stable networks but they all entail a higher aggregate link cost.*

Qualitatively, this result parallels the welfare outcomes described in Proposition 4, and the proof is largely analogous. Because the ordered line ensures that every player d receives signal x precisely from every possible expert e, no player is willing to pay for extra links, as $c > 0$, nor anyone wishes to delete any link, for c small enough. But there are two important caveats. First, the ordered line is uniquely optimal but not the unique pairwise stable network. Players can mis-coordinate and form excessively costly non-minimally connected pairwise stable networks to ensure that every player d receives precise information from every expert e. Undo such networks and replace them with the more efficient ordered line would entail coordination in deleting and reforming links by more than 2 players.

Second, the range of link costs $(0, \hat{c})$ for which the ordered line is pairwise stable can be significantly narrower than the range $(0, \bar{c})$ for which it is uniquely optimal. This discrepancy arises because $\hat{c}$ is the maximal cost c such that every pair of players i and j are willing to pay to ensure that every player d receives precise information from any possible expert e, whereas $\bar{c}$ is the maximum cost at which the total link cost of a minimally connected network is justified by ensuring that every d receives accurate information from any e. Clearly, when players do not care equally about different decision makers d's decisions (i.e., α_{id} is not the same across i), some pairs of players may be willing to pay less than others to ensure that a specific d receives information from a particular e, and this results in $\hat{c} < \bar{c}$.

As well as studying endogenous network formation for small link cost c and fixed utility weights $\boldsymbol{\alpha} > \mathbf{0}$, it is also interesting to consider the case in which the players' payoffs u_i is primarily influenced by their own decision $\hat{y}_i$, i.e., fixing (small) cost c, I take α_{ij} arbitrarily small for all players i and $j \neq i$.

Remark 1 For every full support recognitions probability P, there exists a cost threshold $\tilde{c}$ such that for all $c \in (0, \tilde{c})$, there exists a threshold $\hat{\alpha} > 0$ such that for all $\boldsymbol{\alpha}$ such that $\alpha_{ij} \in (0, \hat{\alpha})$ for all players i and $j \neq i$, the ordered line and every star are pairwise stable. Indeed, the ordered line is pairwise stable for the same reasons as in Proposition 5: every player d cares about receiving the signal x precisely from every possible expert e and so none of them is willing to delete any link for c small enough, nor anyone is willing to add links as $c > 0$. But now, also all stars are pairwise stable. The reason is that, although for any star network N, there exist pairs of players e and d such that the path $p(e, d)$ from e to d has bias reversals, there does not exist any pair i, j such that both paths $p(i, j)$ from

i to j and $p(j,i)$ from j to i have bias reversals.[27] Hence, for any pair i, j either i receives x precisely from j, or vice versa. For α_{ij} sufficiently small for all i and $j \neq i$, they will be both willing to upset N by forming a direct link (if they do not already have one).

The next part of the Section shows how my results can be generalized, beginning by introducing small decay, to then consider general functional forms for the players' utility functions and information, and finally by allowing that all players may be experts and decision makers.

5 Generalizations

Small Decay and Partial Disclosure I consider the case of small decay, $\delta < 1$ but close to one, and return to fully developed persuasion games, where partial disclosure is allowed, through messages $\tilde{m}_{ij}^t : \omega_i(h^t) \subseteq \tilde{m}_{ij}^t \neq X$. The welfare and stability results, Propositions 4 and 5, generalize. The proofs of these generalizations are based on the following extension of Proposition 3, which characterizes equilibrium in a persuasion game with small decay when there is a unique path from the expert e to the decision maker d. This generalizes to persuasion on networks the analysis by Dye (1985), who studied strategic disclosure by an expert who may or may not informed of the state x, or equivalently whose chosen message may not reach the decision maker with positive probability.

Proposition 6 *In every equilibrium of the persuasion game on a tree N, for any realized pair d, e, letting $T = \ell(e, d)$,*

1. if the path $p(e,d)$ has no bias reversals, then there exists an equilibrium μ such that $E[Var(x|\omega_d(h^T), \mu)] \to 0$ for $\delta \to 1$; specifically, if $b_i > (<) b_d$ for all $i \neq d$, then there exists a threshold $\hat{x}$, function of δ such that every $i \neq d$ on $p(e,d)$ discloses every $x > (<) \hat{x}$, so that player d learns x precisely, and further, $\lim_{\delta \to 1} \hat{x} = \underline{x}$ (respectively, $\lim_{\delta \to 1} \hat{x} = \overline{x}$);

2. if $p(e,d)$ has bias reversals, then d cannot learn x precisely in equilibrium, and $E[Var(x|\omega_d(h^T), \mu)] \to \Delta > 0$ for $\delta \to 1$.

[27] Suppose that $p(j,i)$ has a reversal of bias. Hence it must be that neither j nor i are the centre of the star N. Letting k be the centre, say without loss of generality that $b_k > b_j$. Then, for $p(j,i)$ to have a bias reversal, it must be that $b_i < b_j$. Hence, $b_i < b_j < b_k$, so that the path $p(j,i)$ from j to i does not have bias reversals.

The intuition behind this result mirrors the intuition behind Proposition 3. While the latter showed that the logic and results of Milgrom (1981) generalize if and only if the path $p(e,d)$ from the expert e to the decision maker d has no bias reversals, Proposition 6 shows that the logic and results of Dye (1985) generalize under the same conditions on $p(e,d)$. Specifically, suppose that $b_i > b_d$ for all $i \neq d$ on $p(e,d)$, and let $\hat{x}(\delta)$ be the state that solves

$$\left(1 - \delta^T\right) E[x] + \delta^T E[x|x \leq \hat{x}] = \hat{x}. \tag{2}$$

Because $\delta < 1$, there exist terminal histories h^T on path such that $\omega_d(h^T) = X$. Suppose that all players $i \neq d$ on $p(e,d)$ disclose x if and only if $x > \hat{x}$, and otherwise do not disclose anything, i.e., send $m_{ij} = X$. Then, player d's equilibrium expectation of x upon observing $\omega_d(h^T) = X$ is equal to the left-hand side of (2). Because this is equal to $\hat{x}$, the decision maker plays $y_d(X;\mu) = \hat{x} + b_d$. Since $b_i > b_d$ for all $i \neq d$ on $p(e,d)$, all such players i strictly prefer to disclose x than to send $m_{ij} = X$ if and only if $x > \hat{x}$, thus supporting the conjectured equilibrium strategy. (As in Milgrom (1981), partial disclosure messages are ruled out by "worst-case" off-equilibrium-path beliefs that $x = \min \omega_j(h^t)$ for all histories h^t and $\omega_j(h^t) \neq X$.) Further, it is immediate that $\lim_{\delta \to 1} \hat{x}(\delta) = \underline{x}$, and hence that d always learn x in the limit.

In practice, when all players $i \neq d$ on $p(e,d)$ are biased to the right relative to d, the decision maker's strategy ensures that they disclose x unless the state is very close to its left bound $\underline{x}$ by taking an action $\hat{x}$ close to $\underline{x}$ when not receiving any information. The action $\hat{x}$ leads to an equilibrium in which indeed x is disclosed unless $x < \hat{x}$, whenever $\hat{x}$ satisfies equation (2). Conversely, when $b_i < b_d$ for all $i \neq d$ on $p(e,d)$, then all such players i disclose x if and only if $x > \hat{x}$ such that $\left(1 - \delta^T\right) E[x] + \delta^T E[x|x \geq \hat{x}] = \hat{x}$. And since $\lim_{\delta \to 1} \hat{x} = \overline{x}$, again d always learn x in the limit for $\delta \to 1$.

Armed with Proposition 6, I now generalize my welfare and stability results, Propositions 4 and 5. For small decay, there exist a range of 'intermediate' costs such that the unique optimal network N is the ordered line. Further, as decay vanishes, the optimality of the ordered line covers the whole meaningful range of costs $(0, \bar{c})$. Likewise, for small decay the ordered line is pairwise stable in a different intermediate costs range, that converges to $(0, \hat{c})$ for vanishing decay.

Proposition 7 *For all $\boldsymbol{\alpha} > \mathbf{0}$, and full support P, there exist a decay threshold $\bar{\delta} < 1$, as* *well as intermediate cost ranges* (c_-, c^+) *and* $(\hat{c}_-, \hat{c}^+)$ *with* $\lim_{\delta\to 1} c_-(\delta) = \lim_{\delta\to 1} \hat{c}_-(\delta) = 0$, $\lim_{\delta\to 1} c^+(\delta) = \bar{c}$ *and* $\lim_{\delta\to 1} \hat{c}^+(\delta) = \hat{c}$, *such that for all* $\delta \in (\bar{\delta}, 1]$ *and* $c \in (c_-, c^+)$, *the unique optimal network* N *is the ordered line, and for all* $c \in (\hat{c}_-, \hat{c}^+)$, *the ordered line is pairwise stable.*

The proof of Proposition 7 is essentially the same as the proof Propositions 4 and 5, once determined in Proposition 6 that the expected loss $E[Var(x|\omega_d(h^T), \mu)] \to 0$ for $\delta \to 1$ in (the best) equilibrium of the persuasion game given any e, d, N if and only if the path between e and d has no bias reversals. By continuity, exactly the same results holds in the limit for $\delta \to 1$. But because of decay, $\delta < 1$, networks that are not minimally connected may outperform the ordered line when the link cost c is very small, i.e., $0 < c < c_-$, because they allow to transmit information from extreme experts, e.g., $e = 1$, to oppositely extreme decision makers such as $d = n$ through shorter paths than the ordered line. Likewise, when the link cost c is very close to the upper bound $\bar{c}$ that makes connected networks optimal with frictionless communication, it may that optimal networks are not connected because of decay.

I conclude this Section by taking a detour to consider the model of 'unilateral sponsorship' endogenous network formation by Bala and Goyal (2000). In such a framework, every link can be fully paid by one of the linked players. The game form is similar to the bilateral sponsorship game presented in Section 4. For any fixed recognition probability P and utility weights $\boldsymbol{\alpha}$, the players $i \in \mathcal{N}$ simultaneously submit a list $\ell_i \in \{0,1\}^{\mathcal{N}\setminus\{i\}}$ of players $j \neq i$ they wish to link. Here, $\ell_{ij} = 1$ means that i commits to pay the whole cost $2c$ of link N_{ij}. For every pair of players i and j, it is thus the case that $N_{ij} = N_{ji} = 1$ if and only if $\ell_{ij} + \ell_{ji} > 0$, i.e., at least one among i and j commits to pay. Once the network N is formed, an expert e and decision maker d are drawn, the state x is realized and revealed to e, and the persuasion game is played on the network N.

Unlike the case of pairwise stability, I now show that the unilateral sponsorship game's results differ dramatically when moving from no decay to considering small decay. Suppose that the recognition probability $P(e, d)$ is the same for all pairs e, d, and each player i cares mostly about her own decision, i.e., α_{ij} is small for all $j \neq i$. If there is any information decay, the ordered line fails to be a Nash equilibrium (when there are more than 5 players).

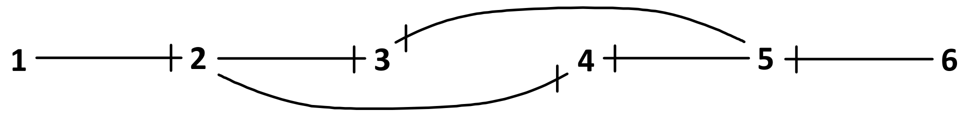

Figure 3: unilateral sponsorship network

Proposition 8 *For all* $\boldsymbol{\alpha} > \mathbf{0}$, *and full support* P, *there exists cost threshold* $\tilde{c}$ *such that for link costs* $c \in (0, \tilde{c})$, *the ordered line is a Nash equilibrium of the unilateral sponsorship game when there is no decay* ($\delta = 1$). *For any decay,* $\delta < 1$, *and any link cost* c, *the ordered line is not a Nash equilibrium when* $n \geq 6$, *the recognition probability* P *is uniform, and* α_{ij} *is sufficiently small for all* i *and* $j \neq i$

Some intuition for the result can be gleaned through Figure 3, where, as customary, the sponsor of each link is denoted by vertical dash. Of course, when there is no decay, either among 3 and 4 would sponsor a direct link. Instead, for any $\delta < 1$, neither player 3 nor player 4 are willing to sponsor that link. Player 3 can access the same information with less decay by forming a link with 5, and so does player 4 by connecting to 2.

General Functional Forms The analysis in Section 4 was undertaken under the assumption that the expert e knew precisely the state x, and that the players' loss functions L_i followed a simple quadratic form $L_i(y, x) = -(y - x - b_i)^2$ with state-independent bias b_i. I now show how to dispense of these assumptions and generalize my findings.

Suppose that, instead of knowing x, player e observes a signal $s \in S \subset \mathbb{R}$, where $S = [\underline{s}, \overline{s}]$ is a closed interval. The distribution of s given x is determined by the density $g(s|x)$ which I assume to be strictly positive on S. The signal s is informative of x in the sense that it satisfies the monotone likelihood ratio property: If $s' > s$ and $x' > x$, then $g(s'|x')/g(s|x') > g(s'|x)/g(s|x)$. Further, I consider any loss function $L_i(y, x)$ that is twice continuously differentiable and that satisfies concavity, $\partial^2 L_i/\partial y^2 < 0$, and single-crossingness: $\partial^2 L_i/\partial y \partial x > 0$ and $\partial L_{i+1}/\partial y > \partial L_i/\partial y$. Hence, every player i's expected value of equilibrium (μ_{edN}, y_d) in the disclosure game given network N, expert e, decision maker d is:

$$u_i(\mu_{edN}, y_d; e, d, N) = \alpha_{id} E[L_i(y_d(h^T; \mu, e), x)].$$

Concavity of L_i and the monotone likelihood ratio property guarantee that, for any signal $s \in S$, there exists a unique decision $y_i(s)$ that maximizes player i's expected payoff

$E[L_i(y,x)|s]$ given s in the disclosure game. Together with single-crossingness, they guarantee that the expected payoff $E[L_i(y,x)|s]$ of each player $i > (<)d$ is strictly increasing (decreasing) in y for all $y \leq (\geq) y_d(s)$. So, for any signal s, player $i > d$ would always like to bias d's decision to the right, and vice versa. I denote by $EL_i(x,s)$ the expected loss $L_i(x,y)$ when $y = y_i(s)$. I say that a path $p(e,d)$ from e to d has bias reversals if there exist i, j on $p(e,d)$ such that $i < d < j$.

Unlike with quadratic loss and constant bias, a mean-variance decomposition analogous to (1) is not available. The players' ex-ante equilibrium payoffs $u_i(\mu_{edN}, y_d; e, d, N)$ cannot be decomposed into a common part such as the residual variance $E[Var(x|\omega_d(h^T), \mu)]$, and an idiosyncratic part such as $(b_i - b_d)^2$ independent of the equilibrium. My concept of network optimality trades off how well information is aggregated with the aggregate link costs. I earlier represented a network N information efficiency as the sum of the players' expected equilibrium payoffs $u_i(\mu_{edN}, y_d; e, d, N)$ in disclosure games given realized expert e, and decision maker d, weighted by the probability P of pairs (e,d). Because expected equilibrium payoffs in the disclosure game need now not be aligned across players, this representation is no longer appropriate, and hence neither is expressing welfare simply as $W(N) = \sum_i U_i(N)$.

Equilibrium informativeness is usually represented in strategic communication games as the ex-ante payoff of the decision maker, the rationale being that she is the one who minimizes loss using the information transmitted in equilibrium. In line with this convention, I now let the information efficiency of a network N be the sum of the players' expected equilibrium payoffs in disclosure games when they take the role of decision makers. Again the best equilibrium given N is selected in case of multiplicity. Hence, I write the welfare function as follows:[28]

$$\hat{W}(N) = \max_{(\mu, y_d)} \sum_{d \in \mathcal{N}} \left[\sum_{e \neq d} E[L_d(y_d(h^T; \mu, e), x)] P(e|d) - c \sum_{j \neq d} N_{dj} \right].$$

I say that a network N is optimal if it maximizes $\hat{W}(N)$.

The following result generalizes Proposition 3, the equilibrium characterization of disclosure games played on a (minimally-connected) network N given expert e and decision

[28] This welfare function also approximates the sum of players utilities in the network N-optimal equilibrium (μ, y_d) when the utility weights $\boldsymbol{\alpha}$ are such that α_{ij} is sufficiently small for all i and $j \neq i$, i.e., each player i cares mostly about her own decision.

maker d. If the path $p(e,d)$ from e to d has no bias reversals, then d always learns s on the equilibrium path and plays $y_d(s)$. As a result, the expected equilibrium payoff $E[L_d(y,x|\omega_d(h^T),\mu)]$ achieves its maximum $E_s[EL_d(x|s)]$ based on the information contained in the signals s. Instead, if the path $p(e,d)$ has bias reversals, then d does not learn s precisely on a positive probability set $\hat{S}$, and the expected payoff $E[L_d(y,x|\omega_d(h^T),\mu)]$ stays smaller than the maximum $E_s[EL_d(x|s)]$.

Proposition 9 *Suppose that a randomly drawn expert e observes a signal s that satisfies the monotone likelihood ratio property, and that each player's loss function L_i is concave and single crossing. In every equilibrium of the disclosure game on a tree N, for any realized pair d,e, letting $T=\ell(e,d)$,*

1. if the path $p(e,d)$ has no bias reversals, then the decision maker d learns s precisely in equilibrium and plays $\hat{y}=y_d(s)$, so that $E[L_d(y,x|\omega_d(h^T),\mu)]=E_s[EL_d(x|s)]$;

2. if $p(e,d)$ has bias reversals, then d cannot learn s precisely in equilibrium, and $E[L_d(y,x|\omega_d(h^T),\mu)]<E_s[EL_d(x|s)]$.

The logic behind this result is analogous to Proposition 3, once realized that concavity and single crossing of the payoff functions $L_i(y,x)$ for each i, together with the monotone likelihood ratio property of $g(s|x)$, imply that for each signal s, player i would like to persuade a decision maker $d<i$ to make a decision to the right of her bliss point $y_d(s)$, and vice versa for $d>i$. When the path $p(e,d)$ has no bias reversals, e.g., $i>d$ for all $i\neq d$ on $p(e,d)$, the off-equilibrium beliefs that $s=\min\omega_j(h^t)$ for all i's successors j ensures that i discloses all her information $\omega_j(h^t)$. Instead when the path $p(e,d)$ has bias reversals, there does not exist off-equilibrium beliefs that simultaneously induce the players $i>d$ and $j<d$ on the path $p(e,d)$ to fully disclose their information.

With the use of Proposition 9 in lieu of Proposition 3, I prove the following result, which generalizes my earlier optimality and stability results, Propositions 4 and 5.

Proposition 10 *Suppose that a randomly drawn expert e observes a signal s that satisfies the monotone likelihood ratio property, and that each player's loss function L_i is concave and single crossing. Then, for every utility weights $\boldsymbol{\alpha}>\mathbf{0}$, and full support recognitions probability P, the unique optimal network N is the ordered line for any link cost $c\in(0,\bar{c})$, and there exists a threshold $\hat{c}$ such that the ordered line is a pairwise stable Nash network for any $c\in(0,\hat{c})$.*

Many Experts and Decision Makers I finally consider the case where multiple players may have information on x and may be called to make decisions. A non-empty set $E \subseteq \mathcal{N}$ of experts and a non-empty set $D \subseteq \mathcal{N}$ of decision makers are randomly drawn from a full support distribution P over the set $\wp^2_+(\mathcal{N}) \equiv \{(E, D) : \varnothing \neq E \subseteq \mathcal{N}, \varnothing \neq D \subseteq \mathcal{N}, D \cap E = \varnothing\}$.

Every player $e \in E$ holds a signal $s_e \in S$ and every $d \in D$ is called to make a decision $\hat{y}_d \in \mathbb{R}$ after communication takes place in network N. I maintain that each player i's loss function L_i take the quadratic form of Section 3: $L_i(\hat{\mathbf{y}}_D, x) = -\sum_{d \in D} \alpha_{id} (\hat{y}_d - x - b_d)^2$. In line with the ideas of Jackson and Wolinsky (1996), the signal s_e of each expert e is valuable to every decision maker d and cannot be replicated by the information of others in the network. Specifically, for any set of experts E, I assume that the distribution g_E of the signal profile $\mathbf{s}_E$ satisfies the monotone likelihood ratio property, and that for any signal profile $\mathbf{s}_E$, any player $e \in E$ and signal s_e, and any set $S_e : \{e\} \subsetneq S_e \subseteq S$, knowing s_e is more informative that knowing only S_e in the sense that it induces a lower expected quadratic loss. This is the case, for example, when the signals s_e are i.i.d. conditional on x.

Assumption 1 *For any set of experts $E \subseteq \mathcal{N}$, the distribution g_E is such that, if $\mathbf{s}'_E \geq \mathbf{s}_E$ and $x' > x$, then $g_E(\mathbf{s}'_E|x') / g_E(\mathbf{s}_E|x') > g_E(\mathbf{s}'_E|x) / g_E(\mathbf{s}_E|x)$, and, for any signal profile $\mathbf{s}_E$, any player $e \in E$ and signal s_e and any set $S_e : \{e\} \subsetneq S_e \subseteq S$,*

$$E[Var(x|\{\mathbf{s}_{E \setminus \{e\}}\} \times S_e)] > E[Var(x|\{\mathbf{s}_E\})]. \tag{3}$$

Given a network N, and realized sets E and D, information transmission of the signals s_e of the experts $e \in E$ to decision makers $d \in D$ is defined as in Section 3, with the qualification that each player i's information set $\omega_i(h^t)$ at any history h^t is a possibly proper subset of S^E. Let $\bar{\ell}(E, D) = \max_{e \in E, d \in D} \bar{\ell}(e, d)$ be the length of the longest path p from some $e \in E$ to some $d \in D$, and say there are $T = \bar{\ell}(E, D)$ periods of information transmission. For every pair (e, d) such that $e \in E$, $d \in D$, and any path p from e to d, at any time $t = 0, ..., \ell(p) - 1$, the player i on the path p at distance $\ell(p(e, i)) = t$ from e sends a verifiable message $\hat{m}_{ij} \subseteq S^E$, i.e., such that $\omega_i(h^t) \subseteq \hat{m}_{ij}$, to any one of her neighbors j who are on a path p' from e to d and are at distance $t + 1$ from e.[29] At time T, all decision

[29] We note that this construction implies that not only the content of each signal s_e is verifiable, but also the identity of the expert e that originated it. My results extend qualitatively to the case in which only the content of each signal s_e is verifiable.

makers $d \in D$ make their decisions $\hat{y}_d$.

The following result restates Proposition 4 for this generalized environment.

Proposition 11 *Suppose that $\boldsymbol{\alpha} > \mathbf{0}$, nature selects at random a non-empty set $E \subseteq \mathcal{N}$ of experts and a non-empty set $D \subseteq \mathcal{N}$ of decision makers, with a full support distribution P over $\wp^2_+(\mathcal{N})$, and that Assumption 1 holds: all signals s_e are informative of x and none is redundant. Then, the unique optimal network is the ordered line for any link cost $c \in (0, \bar{c})$, and there exists a threshold $\hat{c}$ such that the ordered line is pairwise stable for any $c \in (0, \hat{c})$.*

The core of the proof is to show that, for all realized sets of experts E and decision-makers D, the information of all the experts reach all decision makers when the network N is the ordered line. The logic behind this result is a similar to that of Proposition 9. Each possible decision-makers d knows how to interpret if she is not fully disclosed a signal s_e. Information on signals s_e of experts $e < d$ biased to the left is transmitted to d through players $i : e \leq i < d$ biased to the left. Suppose one such signal s_e is not disclosed to d. Then d reasons that the withheld information must be evidence that would move her decision $\hat{y}_d$ to the right, i.e., she formulates the (off-equilibrium-path) belief that $s_e = \min \omega_d(h^T)|_e$. This leads d to take the most rightward decision $\hat{y}_d$ possible compatibly with the information $\omega_d(h^T)|_e$ received. Such behavior deters all players $i < d$ from withholding any information about any signal s_e with $e \leq i < d$ along the path from e to d. An analogous, symmetric, argument concludes that all signals s_e of experts $e > d$ biased to the right are transmitted to d precisely, in equilibrium.

The above arguments conclude that the information of each experts reach all decision makers when N is the ordered line. The suboptimality of any other network, and the result that N is pairwise stable, follow from the same arguments used for Proposition 10, once noticed that, because the recognition probability P is full support on $\wp^2_+(\mathcal{N})$, every pair of one expert e and one decision maker d can be the only players in the disclosure game, $E = \{e\}$ and $D = \{d\}$, with positive probability P.

Because the assumption that the recognition probability P is full support on $\wp^2_+(\mathcal{N})$ plays a major role in the proof of Proposition 11, I conclude this section by considering the case in which every player is an expert and a decision maker with probability one, $E = D = \mathcal{N}$. The following example shows that, in this case, the ordered line need not be the unique optimal network.

Example 2 Suppose that there are 4 players at equidistant bliss points $b_i = ib$, for $i = 1, ..., 4$. All the players are simultaneously experts and decision makers: $E = D = N$, with equal weights $\alpha_i \equiv \alpha_{ij}$ for $j \neq i$. That is, every player i cares about the decision of every other player j equally. Suppose that the state is $x \sim U[0,1]$, and the signals s_i, $i = 1, ..., 4$ are i.i.d. from a continuous Bernoulli of shape x, that is $g(s_i|x) = k(x)x^{s_i}(1-x)^{1-s_i}$ for $s_i \in [0,1]$. As these distributions belong to an exponential family,[30] the expected value $E[x|s]$ of the state x given any profile of signals s is linear in $s_1, \ldots,$ and s_n (e.g., Jewel, 1974). As it is immediate that the sum of the signals $\Sigma(s) \equiv \sum_{i=1}^{n} s_i$ is a sufficient statistic for the joint distribution $g(s|x)$, it follows that $E[x|s]$ is linear in $\Sigma(s)$, and I write it as $E[x|s] = a\Sigma(s) + z$.

Consider the following line network N:

$$1 \text{———} 3 \text{———} 2 \text{———} 4$$

I will now show that there exists an equilibrium in which each player i transmits all information $\omega_i(h^t)$ precisely to his neighbor j on every path $p(e,d)$ at any history h^t at every time t she is called to play. Such an equilibrium is supported by off-equilibrium-path beliefs that assign probability one to $s_e = \min \omega_i(h^t)|_e$ if $b_e > b_i$ and $s_e = \max \omega_i(h^t)|_e$ if $b_e < b_i$, as has by now become customary in this paper. Hence, each player d at every history terminal history h^T plays $\hat{y}_d = E[x|\omega_d(h^T)] + b_d$ such that $E[x|\omega_d(h^T)] = a\Sigma(\hat{s}(\omega_d(h^T))) + z$, where $\hat{s}_e(\omega_d(h^T)) = \min \omega_d(h^T)|_e$ if $e > d$, $\hat{s}_d(\omega_d(h^T)) = s_d$, and $\hat{s}_e(\omega_d(h^T)) = \max \omega_d(h^T)|_e$ if $e < d$.

To see that this is an equilibrium, consider first players 1 and 4, who are called to play only at time $t = 0$. Evidently, player 1 is willing to transmit her signal s_1 precisely to player 3. This is because all decision makers i are such that $b_i > b_1$, and hence the logic of Milgrom (1981) applies. Analogously, also player 4 transmits her signal s_4 precisely to player 2.

Players 2 and 3 are called to play multiple times. At time 0, they are called to transmit information about their own signals to each of their neighbors, at time 1 they communicate information about their neighbors signals, and at time 2 about their neighbor at distance 2. Focus on 3, as the argument for 2 is symmetric. Again, the logic of Milgrom (1981) implies that 3 transmits every signal s_e precisely to player 1. Consider the choice of 3 to relay

[30] See Loaiza-Ganem and Cunningham (2019), where it is also derived that $k(x) = \frac{2\tanh^{-1}(1-2x)}{1-2x}$ if $x \neq 1/2$ and $k(x) = 2$ if $x = 1/2$.

information on either signal s_1 or s_3 to player 2. Player 3 anticipates that such information will also reach 4 precisely. Because player 3 cares about the decision of 2 and 4 equally, and the bias b_3 is exactly equal to the average of b_2 and b_4, player 3 relays her information about s_1 and s_3 precisely to 2.

In fact, although by withholding information about signal s_1, player 3 could bias the decision of 2 and 4 in the same direction, what she would gain by biasing 2's decision to the right would be more than lost by also biasing 4's decision, and vice versa. A fortiori, 3 also reveals s_3 precisely, as by withholding any information, she would wind up biasing 2's decision to the left and 4's decision to the right, both of them away from his bliss point $x + b_3$. The formal arguments are in Appendix.

6 Conclusion

This paper has studied strategic communication in networks by building on classical models of verifiable information transmission, such as Milgrom's (1981). The analysis has revealed that the unique optimal network is a line where players are ordered according to their preferences. This ordered line network is also pairwise stable, meaning no player has an incentive to unilaterally delete links, nor any pair of players to form an extra direct link in the network. These results contrast sharply with findings from earlier network studies such as those by Jackson and Wolinsky (1996). They identified star networks as both optimal and pairwise-stable in non-strategic communication settings with technological constraints such as information decay. In terms of network centrality, the contrast could not be more extreme. The star is the most centralized minimally connected network, whereas the line is the most decentralized one.

These results are particularly relevant in political economy applications, where strategic communication often shapes interactions between political agents. A star network will disrupt the flow of information, when the players have incentives to manipulate or withhold it. Instead, the line in which political agents form links with those with the closest views ensures that no verifiable information can be withheld in equilibrium. If that were the case, each decision maker will know which political side to blame for the withheld information, and would react by moving her decision in the opposite direction.

This study opens several avenues for future research. One potential extension could explore the implications of multi-dimensional states and decisions. Another area of interest is the endogenous formation of networks in environments of uncertainty, where players lack full information about each others' preferences. Future research could also incorporate repeated models of strategic communication, as well as the possibility that networks evolve over time.

References

[1] Acemoglu, Daron, Munther A. Dahleh, Ilan Lobel, and Asuman Ozdaglar. 2011. "Bayesian learning in social networks." *Review of Economic Studies*, 78: 1201–1236.

[2] Acemoglu, Daron, Kostas Bimpikis, and Asuman Ozdaglar. 2014. "Dynamics of Information Exchange in Endogenous Social Networks." *Theoretical Economics*, 9: 41–97.

[3] Alonso, Ricardo, Wouter Dessein, and Niko Matouschek. 2008. "When Does Coordination Require Centralization?" *American Economic Review*, 98(1): 145–179.

[4] Ambrus, Attila, Eduardo Azevedo, and Yuichiro Kamada. 2013. "Hierarchical Cheap Talk." *Theoretical Economics*, 8: 223–261.

[5] Bala, Venkatesh, and Sanjeev Goyal. 1998. "Learning from Neighbours," *Review of Economic Studies,* 65(3): 595–621.

[6] Bala, Venkatesh, and Sanjeev Goyal. 2000. "A Non-Cooperative Model of Network Formation." *Econometrica*, 68: 1181–1230.

[7] Battaglini, Marco. 2002. "Multiple Referrals and Multidimensional Cheap Talk." *Econometrica*, 70: 1379–1401.

[8] Bloch, Francis, Gabrielle Demange, and Rachel Kranton. 2018. "Rumors and Social Networks." *International Economic Review*, 59(2): 421–448.

[9] Box-Steffensmeier, Janet M., and Dino P. Christenson. 2014. "The Evolution and Formation of Amicus Curiae Networks." *Social Networks*, 36: 82–96.

[10] Bratton, Kathleen A., and Stella M. Rouse. 2011. "Networks in the Legislative Arena: How Group Dynamics Affect Cosponsorship." *Legislative Studies Quarterly*, 34(3): 423–460.

[11] Caldeira, Gregory A. 1985. "The Transmission of Legal Precedent: A Study of State Supreme Courts." *American Political Science Review*, 79(1): 178–194.

[12] Carpenter, Daniel P., Kevin M. Esterling, and David M. J. Lazer. 2004. "Friends, Brokers, and Transitivity: Who Informs Whom in Washington Politics?" *Journal of Politics*, 66(1): 22–246.

[13] Crawford, Vincent, and Joel Sobel. 1982. "Strategic Information Transmission." *Econometrica*, 50: 1431–145.

[14] Dessein, Wouter. 2002. "Authority and Communication in Organizations." *Review of Economic Studies*, 69: 811–838.

[15] Dye, Ronald A. 1985. "Disclosure of Nonproprietary Information." *Journal of Accounting Research*, 23(1): 123–145.

[16] Egorov, Georgy, and Konstantin Sonin. 2020. "Persuasion on Networks." *Mimeo.*

[17] Farrell, Joseph, and Robert Gibbons. 1989. "Cheap Talk with Two Audiences." *American Economic Review,* 79(5): 1214–23.

[18] Fiellin, Alan. 1962. "The Functions of Informal Groups in Legislative Institutions." *Journal of Politics*, 24(1): 72–91.

[19] Fowler, James H., Timothy R. Johnson, James F. Spriggs, Sangick Jeon, and Paul J. Wahlbeck. 2007. "Network Analysis and the Law: Measuring the Legal Importance of Precedents at the U.S. Supreme Court." *Political Analysis*, 15: 324–346.

[20] Galeotti, Andrea, Christian Ghiglino, and Francesco Squintani. 2013. "Strategic Communication Networks." *Journal of Economic Theory*, 148(5): 1751–1769.

[21] Galeotti, Andrea, Sanjeev Goyal, and Jurjen Kamphorst. 2006. "Network Formation with Heterogeneous Players." *Games and Economic Behavior*, 54(2): 353–372.

[22] Galperti, Simone, and Jacopo Perego. 2024. "Games with Information Constraints: Seeds and Spillovers," *Theoretical Economics,* forthcoming.

[23] Gieczewski, German. 2022. "Verifiable Communication on Networks." *Journal of Economic Theory*, 204: 105494.

[24] Gilligan, Thomas W., and Keith Krehbiel. 1987. "Collective Decision-Making and Standing Committees: An Informational Rationale for Restrictive Amendment Procedures." *Journal of Law, Economics and Organization*, 3: 287–335.

[25] Grossman, Sandford J. 1981. "The Informational Role of Warrantees and Private Disclosure About Product Quality." *Journal of Law and Economics,* 24: 461-83.

[26] Hume, David. 1741. "Part I, Essay VIII: Of Parties in General," in *Essays, Moral and Political,* by David Hume, Edinburgh: A. Kincaid.

[27] Ivanov, Maxim. 2010. "Communication via a Strategic Mediator." *Journal of Economic Theory*, 145: 869–884.

[28] Jackson, Matthew O., and Asher Wolinsky. 1996. "A Strategic Model of Social and Economic Networks." *Journal of Economic Theory*, 71(1): 44–74.

[29] Jackson, Matthew O., and Leeat Yariv. 2011. "Diffusion, Strategic Interaction, and Social Structure." In *Handbook of Social Economics,* Vol. 1A, edited by Jess Benhabib, Alberto Bisin, and Matthew O. Jackson, 645-678. Elsevier.

[30] Jewel, William S. 1974. "Credible Means are Exact Bayesian for Exponential Families." *ASTIN Bulletin*, 8: 77–90.

[31] Kamenica, Emir, and Matthew Gentzkow. 2011. "Bayesian Persuasion." *American Economic Review*, 101(6): 2590–2615.

[32] Kartik, Navin. 2009. "Strategic Communication with Lying Costs." *The Review of Economic Studies*, 76(4): 1359–95.

[33] König, Thomas, and Thomas Bräuninger. 1998. "The Formation of Policy Networks: Preferences, Institutions and Actors' Choice of Information and Exchange Relations." *Journal of Theoretical Politics*, 10(4): 445–71.

[34] Laumann, Edward O., and David Knoke. 1987. *The Organizational State: Social Choice in National Policy Domains*. Madison WI: The University of Wisconsin Press.

[35] Lipnowski, Eliot, and Doron Ravid. 2020. "Cheap Talk with Transparent Motives," *Econometrica,* 88(4): 1631-60.

[36] Lipnowski, Elliot, and Evan Sadler. 2019. "Peer-Confirming Equilibrium," *Econometrica,* 87(2): 567–591.

[37] Lizzeri, Alessandro. 1999. "Information Revelation and Certification Intermediaries." *The RAND Journal of Economics*, 30(2): 214–231.

[38] Loaiza-Ganem, Gabriel, and John P. Cunningham. 2019. "The Continuous Bernoulli: Fixing a Pervasive Error in Variational Autoencoders." In *Advances in Neural Information Processing Systems*: 13266–13276. Curran Associates, Inc.

[39] McPherson, Miller, Lynn Smith-Lovin, and James M. Cook. 2001. "Birds of a Feather: Homophily in Social Networks." *Annual review of sociology* 27(1): 415-444.

[40] Migrow, Dimitri. 2021. "Designing Communication Hierarchies." *Journal of Economic Theory*, 198: 105349.

[41] Milgrom, Paul R. 1981. "Good News and Bad News: Representation Theorems and Applications." *Bell Journal of Economics*, 12(2): 380–391.

[42] Morris, Stephen. 2001. "Political Correctness." *Journal of Political Economy*, 109(2): 231–265.

[43] Myerson, Roger B. 1991. *Game Theory: Analysis of Conflict.* Harvard University Press: Cambridge, MA.

[44] Onuchic, Paula and João Ramos, 2023. "Disclosure and Incentives in Teams." Mimeo, *University of Oxford.*

[45] Patty, John W., and Elizabeth M. Penn. 2014. "Sequential Decision Making and Information Aggregation in Small Networks." *Political Science Research and Methods*, 2(2): 243–271.

[46] Radner, Roy. 1993. "The Organization of Decentralized Information Processing." *Econometrica*, 61: 1109–46.

[47] Rose, Richard. 1964. "Parties, Factions and Tendencies in Britain," *Political Studies,* 12(1): 33-46.

[48] Sadler, Evan. 2020. "Diffusion Games," *American Economic Review,* 110(1): 225–270.

[49] Skinner, Douglas J. 1994. "Why Firms Voluntarily Disclose Bad News." *Journal of Accounting Research*, 32(1): 38–60.

[50] Squintani, Francesco. 2018. "Information Transmission in Political Networks." *Mimeo, University of Warwick.*

[51] Verrecchia, Robert E. 1983. "Discretionary Disclosure." *Journal of Accounting and Economics*, 5: 179–194.

[52] Vives, Xavier. 1984. "Duopoly Information Equilibrium: Cournot and Bertrand." *Journal of Economic Theory*, 34(1): 71–94.

[53] Zafonte, Mathew, and Paul A. Sabatier. 1998. "Shared Beliefs and Imposed Interdependencies as Determinants of Ally Networks in Overlapping Subsystems." *Journal of Theoretical Politics*, 10(4): 473–505.

[54] Zhang, Yan, A. J. Friend, Amanda L. Traud, Mason A. Porter, James H. Fowler, and Peter J. Mucha. 2008. "Community Structure in Congressional Cosponsorship Networks." *Physica A: Statistical Mechanics and Its Applications*, 387: 1–8.

Appendix: Omitted Proofs

Proof of Proposition 3. Part 1. Say the path $p(e,d)$ has no bias reversals, e.g. $b_i > b_d$ for all $i \neq d$. Then I claim there is a perfect Bayesian equilibrium, in which every such player i reveals her information $m_{ij}(h^t)$ to her successor j on the path $p(e,d)$, $m_{ij}(h^t) = \omega_i(h^t)$, in every history h^t at the time t in which she is called to play, and every player $j \neq e$ holds beliefs that the state is $\min \omega_j(h^t)$ with probability one at every history h^t at which j is called to play (j's beliefs at the other times is irrelevant), and the decision maker d chooses $y_d(h^T) = \min \omega_d(h^T) + b_d$, for every history h^T. As a result, the message $\hat{m}_{ij} = \{x\}$ travels along the path $p(e,d)$ from e to d on the equilibrium path, and d chooses $y_d(h^T) = x + d$, so that $E[Var(x|\omega_d(h^T), \mu)] = 0$.

In order to prove this, first note that the beliefs of every player $j \neq e$ are admissible in perfect Bayesian equilibrium, because admissibility constrains beliefs only on information sets $\omega_j(h^t)$ on the equilibrium path. For every state x, the unique information set on path is $\omega_j(h^t) = \{x\}$ and $\min \omega_j(h^t) = x$, consistently. Further, d's strategy y_d is clearly optimal given her beliefs.

Now consider the decision of any player $i \neq d$ on the path $p(e,d)$ at time $t = \ell(p(e,i))$ and at any history h^t, and let j be i's immediate successor on the path $p(e,d)$. I first note that $\omega_j(h^t) = X$ and $\omega_i(h^t) \subseteq \omega_j(h^{t+1})$ for any history h^{t+1} that contains h^t. Intuitively, j knows nothing before i communicates with her, and it cannot be that j knows more at time $t+1$ than i knows at time t, as anything that j knows she has been told by i. Further, because the expected utility $E[u_i(\hat{y}, x)|\omega_i(h^t), \mu]$ is an integral of quadratic loss functions $u_i(\hat{y}) = -(\hat{y} - x)^2$, it is a also a quadratic loss function.

I now show that player i does not gain by deviating from the equilibrium strategy $m_{ij}(h^t) = \omega_i(h^t)$. Proceeding backwards, consider any history h^{T-1} and the last player $i \neq d$ on the path $p(e,d)$. Recall that for any history h^T that includes h^{T-1}, it is the case that $\omega_i(h^{T-1}) \subseteq \omega_d(h^T)$, and that $E[u_i(\hat{y}, x)|\omega_i(h^{T-1}), \mu]$ is a quadratic loss function. Then, because $y\left(\omega_d(h^T)\right) = \min \omega_d(h^T) + b_d$, $\min \omega_d(h^T) \leq \min \omega_i(h^{T-1})$ by set inclusion, and $\min \omega_i(h^{T-1}) \leq E[x|\omega_i(h^{T-1}), \mu]$ by verifiability, I obtain that the bliss point $E[x|\omega_i(h^{T-1}), \mu] + b_i$ of i's expected utility $E[u_i(\hat{y}, x)|\omega_i(h^{T-1}), \mu]$ given information $\omega_i(h^{T-1})$ and equilibrium beliefs μ is such that $E[x|\omega_i(h^{T-1}), \mu] + b_i > \min \omega_i(h^{T-1}) + b_i > \min \omega_i(h^{T-1}) + b_d \geq \min \omega_d(h^T) + b_d$. Hence, the expected utility $E[u_i(\hat{y}, x)|\omega_i(h^{T-1}), \mu]$ of player i increases in $\hat{y}$, for $\hat{y} > E[x|\omega_i(h^{T-1})] + b_i$. As a result, i would like to maximize $\hat{y}$ and, by verifiability, this can be done by sending message $\hat{m}_{id}^{T-1} = \omega_i(h^{T-1})$.

For any $t = 0, ...T-1$, let i_t be the player on the path $p(e,d)$ at distance t from e. I have proved above that at any history h^{T-1} player i_{T-1} sends message $\hat{m}_{i_{T-1}i_T} = \omega_{i_{T-1}}(h^{T-1})$ to $i_T = d$. I now proceed by induction. For any $t < T-1$, suppose that at any history $h^{t+\tau}$, $\tau = 1, ..., T-1-t$, player $i_{t+\tau}$ sends message $\hat{m}_{i_{t+\tau}i_{t+\tau+1}} = \omega_{i_{t+\tau}}(h^{t+\tau})$ to his immediate successor $i_{t+\tau+1}$ on the path $p(e,d)$. Then the choice faced by player i_t at any history h^t is exactly the same faced by the last player $i \neq d$ on path $p(e,d)$ at any history h^{T-1}, considered above. The reason is because, again, $\omega_{i_t}(h^t) \subseteq \omega_{i_{t+1}}(h^{t+1})$ for any history h^{t+1} that contains h^t and because the induction argument implies that $\omega_{i_{t+1}}(h^{t+1}) = \omega_{i_d}(h^T)$ on the equilibrium history that contains h^{t+1}. So, again, because $b_{i_t} > b_d$, it is a best response for player i_t at history h^t to reveal all information and send message $\hat{m}_{i_t i_{t+1}} = \omega_{i_t}(h^t)$.

I now show that in every equilibrium μ, it is the case that the message $\hat{m}_{ij} = \{x\}$ travels along the path $p(e,d)$ from e to d on the equilibrium path, and the decision maker chooses $\hat{y} = x + b_d$, so that $E[Var(x|\omega_d(h^T), \mu)] = 0$ for all terminal histories h^T. Suppose not: there exists states x and terminal histories h^T such that $y(\omega_d(h^T)) \neq x + b_d$. In such a case, there must be a non-null set of states $\hat{X} \subseteq X$ such that for all $x \in \hat{X}$ the same information set $\hat{\omega}_d$ is reached with positive equilibrium μ probability, i.e., there exist different equilibrium path histories h^T that include all $x \in \hat{X}$ and such that $\omega_d(h^T) = \hat{\omega}_d$. Further, it must be that $y(\hat{\omega}_d) = E[x|\hat{\omega}_d, \mu] + b_d$. By the intermediate value theorem, there exist a non-null measure set $\tilde{X}$ of $x \in \hat{X}$ such that $x > E[x|\hat{\omega}_d, \mu]$. Pick any such a state x, and (using the same notation introduced above) suppose that any history h^{T-1}, player i_{T-1} knows that the state is x, i.e., $\omega_{i_{T-1}}(h^{T-1}) = \{x\}$. Then, because $x > E[x|\hat{\omega}_d, \mu]$ and $b_{i_{T-1}} > b_d$, player i_{T-1} prefers to induce decision $y(\{x\}) = x + b_d$ closer to her bliss point $x + b_{i_{T-1}} > x + b_d$ than $E[x|\hat{\omega}_d, \mu] + b_d$. Hence, player i_{T-1} strictly prefers to send $\hat{m}_{i_{T-1}} = \omega_{i_{T-1}}(h^{T-1}) = \{x\}$ over any message $\hat{m}_{i_{T-1}}$ that induces decision $E[x|\hat{\omega}_d, \mu]$. Using an induction argument analogous to the one above, it follows that or any $t = 0, ..., T-1$, player i_t strictly prefers to send message $\hat{m}_{i_t} = \omega_{i_t}(h^t)$ at any history in which $\omega_{i_t}(h^t) = \{x\}$ over any message that eventually induces decision $E[x|\hat{\omega}_d, \mu]$ with positive equilibrium μ probability. This include the expert $e = i_0$, therefore the information set $\hat{\omega}_d$ cannot reach the decision maker for any $x \in \tilde{X}$. But this contradicts that $\tilde{X}$ is a subset of $\hat{X}$.

Part 2. Suppose now that the path $p(e,d)$ has bias reversals. Consider terminal histories

h^T such that $\omega_d(h^T) = X$, and let $\hat{x}(X) = E[x|X, \mu]$. For the moment, suppose that $\hat{x}(X)$ is arbitrary (the proof will show that $\omega_d(h^T) = X$ must be on the equilibrium path, and hence that it is determined by Bayes rule.)

Pick any player $i \neq d$ on $p(e, d)$, and let t by the distance of i from e (for $i = e$, I let $t = 0$). Every player i's message $\hat{m}_{ij}^t$ is restricted so that $\hat{m}_{ij}^t \in \{\{x\}, X\}$. Hence, for any history h^t, the information $\omega_i(h^t)$ that player i has on x may either be that $\omega_i(h^t) = \{x\}$ or that $\omega_i(h^t) = X$, either i knows x or she knows nothing. If $\omega_i(h^t) = X$, then verifiability, $\omega_i(h^t) \subseteq \hat{m}_{ij}^t$ implies that $\hat{m}_{ij}^t = X$ and that $\omega_d(h^T) = X$ for every terminal history h^T that contains h^t.

Consider any player $i \neq d$ on $p(e, d)$ such that $b_i > b_d$. Consider any state $x \geq \underline{x}$ such that $\hat{x}(X) + 2(b_d - b_i) < x < \hat{x}(X)$, and suppose that $\omega_i(h^t) = \{x\}$. In any terminal history h^T such that $\omega_i(h^t) = \{x\}$, player d's decision is $\hat{y}_d = x + b_d$. Because $2(b_i - b_d) > \hat{x}(X) - x$, it follows that

$$\begin{aligned} u_{id}(\hat{x}(X) + b_d, x) - u_{id}(x + b_d, x) &= \\ &= -\alpha_{id}(\hat{x}(X) + b_d - x - b_i)^2 + \alpha_{id}(b_d - b_i)^2 \\ &= -\alpha_{id}(\hat{x}(X) - x)\left[(\hat{x}(X) - x) - 2(b_i - b_d)\right] > 0. \end{aligned}$$

Player i strictly prefers that d receives no information about x and plays $\hat{y}_d = \hat{x}(X) + b_d$, rather than d learns x and plays $\hat{y}_d = x + b_d$. Indeed, player i can insure that d is not disclosed x by sending message $\hat{m}_{ij}^t = X$, as this implies that $\omega_d(h^T) = X$ for every terminal history h^T that contains h^t. Hence, there cannot exist any equilibrium in which $\omega_d(h^T) = \{x\}$ on the equilibrium path, for any $x \geq \underline{x}$ such that $\hat{x}(X) + 2\max_{i>d}(b_i - b_d) < x < \hat{x}(X)$.

An analogous, symmetric, argument implies that there cannot exist any equilibrium in which $\omega_d(h^T) = \{x\}$ on the equilibrium path, for any $x \leq \overline{x}$ such that $\hat{x}(X) < x < \hat{x}(X) + 2\min_{i<d}(b_i - b_d)$.

As a consequence, the expected loss can be bounded as follows:

$$E[Var(x|\omega_d(h^T), \mu)] \geq \min_{\hat{x} \in X} \int_{\max\{\underline{x}, \hat{x}+2(b_1-b_d)\}}^{\min\{\hat{x}+2(b_n-b_d), \overline{x}\}} (\hat{x} - x)^2 f(x) dx > 0.$$

Part 3. The third result follows from the second one as a corollary, after noting that if $b_i - b_d > (<)0$ is sufficiently large, then the state space X is a subset of both sets $(\hat{x}(X) + b_d - b_i, \hat{x}(X))$ and $(\hat{x}(X), \hat{x}(X) + b_d - b_i)$. Hence, every player $i \neq d$ on $p(e, d)$ at distance t such that $b_i - b_d > (<)0$ knows that $\hat{x}(X) + b_d - b_i < x < \hat{x}(X)$, respectively that $\hat{x}(X) < x < \hat{x}(X) + b_d - b_i$ for every $x \in X$ and history h^t. As a result, the expected loss of each equilibrium μ is $E[Var(x|\omega_d(h^T), \mu)] \geq \int_X (\hat{x}(X) - x)^2 f(x)dx > 0$, and this can be so if and only if player d's decision is $y(\omega_d(h^T)) = \hat{x}(X) + b_d = E[x|\omega_d(h^T) = X, \mu] + b_d$ for almost every terminal history h^T. That is to say, player d almost always makes her decision (as if) without any information on x. ■

Proof of Proposition 5. The proof that for any c small enough the ordered line is a pairwise stable Nash network is immediate. Because the ordered line has no bias reversal paths, by Proposition 3, the state x is always relayed precisely along the path $p(e, d)$ regardless of the realized identities of players e and d. Because each realized expert e's information is useful to every realized decision maker d, there exists a cost threshold $\hat{c}_1 > 0$ such that for all link costs $c \in (0, \hat{c}_1)$, deleting links from the ordered line is detrimental.

Consider any network N such that there exists a pair of players e and d that are not connected through a path $p(e, d)$ without bias reversals. By Proposition 3, player d does not receive x precisely from e, so that $E[Var(x|\omega_d(h^T), \mu)] > 0$. Because $\alpha_{ed} > 0$, $\alpha_{dd} > 0$ and $P(e, d) > 0$, there exists a there exists a cost threshold $\hat{c}_2 > 0$ such that for all link costs $c \in (0, \hat{c}_2)$ player e and d each benefit more than c by sponsoring a direct link between themselves, so as to secure that d receives x from e, and that $E[Var(x|\omega_d(h^T), \mu)] = 0$.

Finally, consider any network N different from the ordered line but such that every player e and d is connected through a path with a bias reversal. The proof of Proposition 4 implies that N is not minimally connected, hence, it has a higher aggregate link cost. Hence picking $\hat{c} = \min\{\hat{c}_1, \hat{c}_2\}$ concludes proof that ordered line is the unique minimal aggregate cost pairwise stable network.

A minimal example of a pairwise stable network different from the ordered line is a 4-player is the circle portrayed in figure 2, in which both 1 and 4 are linked with 2 and 3. Every pair (e, d) is connected through a path without bias reversals: this is obvious for $d = 1$ and $d = 4$, whereas $d = 2$ (or $d = 3$) is directly connected to $e = 1$ ($e = 4$, resp.) and $e = 4$ ($e = 1$) and to $e = 3$ ($e = 2$) through $i = 4$ ($i = 1$). Hence for any $c > 0$, there is no pair of players who find it mutually beneficial to form any additional link. Removing the link from 1 to 2 (or the one from 1 to 3) removes the only path without bias reversals from 1 to 2 (or to 3, respectively), and makes it impossible that $d = 2$ ($d = 3$) receives precise information from $e = 1$. Hence, for any $\boldsymbol{\alpha} > \mathbf{0}$, there is a link cost threshold $\hat{c} > 0$ sufficiently small such that for all $c \leq \hat{c}$, both 1 and 2 (or 3, respectively) are willing to pay the cost c to ensure that $d = 2$ ($d = 3$) receives precise information from $e = 1$. A symmetric argument holds for deleting the link from 4 to 3 and the one from 4 to 2. ■

Proof of Proposition 6. Suppose that $\delta < 1$ but close to one.

Part 1. If every player i on the path $p(e, d)$ were to relay the message $\hat{m}_{ij} = \omega_i(h^t)$ for every information history h^t in which she is called to play, the message $\{x\}$ would reach player d along the path $p(e, d)$ with probability δ^T.

Say the path $p(e, d)$ has no bias reversals, e.g. $b_i > b_d$ for all i. For any δ, let $\hat{x}(\delta)$ solve $\left(1 - \delta^T\right) E[x] + \delta^T E[x|x \leq \hat{x}] = \hat{x}$. Further, for $\delta \to 1$, it is the case that $\hat{x} \to E[x|x \leq \hat{x}]$, and because $E[x|x \leq \hat{x}] < \hat{x}$ for any $\hat{x} > \underline{x}$ by the intermediate value theorem, it must be that $\hat{x} \to \underline{x}$. Take δ such that $\hat{x}(\delta) - \underline{x} < \min\{b_i - b_d\}$.

Consider the profile of communication strategies m such that each player $i \neq d$ on the path $p(d, e)$ at any history h^t in which she is called to play, discloses her information $\omega_i(h^t)$ to his neighbor j on path $p(e, d)$ if and only if $\min \omega_i(h^t) > \hat{x}$; i.e., $m_{ij}(h^t) = \omega_i(h^t)$ if $\min \omega_i(h^t) > \hat{x}$, and $m_{ei}(h^t) = X$ if $\min \omega_i(h^t) > \hat{x}$. Every player $j \neq e$ on the path $p(e, d)$ at any history history h^t in which she is called to play believes that $x = \min \omega_i(h^t)$ with probability one, unless $\omega_i(h^t) = X$, in which case she believes that $x = \hat{x}$ with probability one. The decision maker d plays $y_d\left(\omega_d\left(h^T\right), \mu\right) = \hat{x} + b_d$ if $\omega_d\left(h^T\right) = X$, and $y_d\left(\omega_d\left(h^T\right), \mu\right) = \min \omega_d(h^T) + b_d$ for every other terminal history information set $\omega_d(h^T)$.

This profile of strategies is an equilibrium, because d's decision is clearly sequentially rational, and no player i on the path $p(e, d)$ wishes to deviate. If d receives an information set $\omega_d\left(h^T\right) \neq X$, then she plays $y_{de}\left(\omega_e\left(h^T\right), \mu\right) = \min \omega_d\left(h^T\right) + b_d$. Because $x \in \omega_d\left(h^T\right)$, every set such that $x = \min \omega_d\left(h^T\right)$ yields the same outcome $\omega_d\left(h^T\right) = \{x\}$, and every set such that $x < \min \omega_d\left(h^T\right)$ yields a worse outcome for every player i than the set $\omega_d\left(h^T\right) = \{x\}$. For every state $x > \hat{x}$, each player i is obviously better off if d plays $x + b_d$ rather than $x' + b_d$ with $x' < x$, including $x' = \hat{x}$, so none of the players gains from deviating from equilibrium and they all disclose x whenever they know it. For every state $x < \hat{x}$, because $\hat{x} - \underline{x} < b_i - b_d$ for all players i on $p(e, d)$, none of them has any incentive in deviating. Each player i is better off if d plays $\hat{x} + b_d$ than if d plays $\hat{y} < \hat{x} + b_d$, because $x + b_i > \hat{x}(\delta) + b_d$ for all $x \in [\underline{x}, \hat{x}]$.

Because $\lim_{\delta\to 1} \hat{x}(\delta) = \underline{x}$, $E[Var(x|\omega_d(h^T), \mu)] = E(E[x|\omega_d(h^T), \mu] - x)^2 = \int_{x\leq\hat{x}}(E[x|\omega_d(h^T)] - x)^2 f(x)dx \to 0$ for $\delta \to 1$.

Part 2. Suppose now that the path $p\,(e, d)$ has bias reversals. Consider any player $i \neq d$ on $p\,(e, d)$. Because $\delta < 1$, the information $\omega_i(h^t) = X$ is held by player i at the time t in which she is called to play with strictly positive probability on the equilibrium path. Define $\hat{x}(X) = E[x|\omega_d(h^T) = X, \mu]$. Because on the equilibrium path i and d must hold common beliefs $E[x|\omega_i(h^t), \mu] = E[x|\omega_d(h^T) = X, \mu] = \hat{x}(X)$.

Suppose without loss of generality that $b_e > b_d$. Pick any player $i \neq d$ on $p\,(e, d)$ such that $b_i > b_d$, and let t by the distance of i from e (for $i = e$, I let $t = 0$). There are two possibilities to consider.

Suppose first that there exists a non-null measure set $\tilde{X}_i \subseteq (\hat{x}(X) + b_d - b_i, \hat{x}(X))$ such that for all $x' \in \tilde{X}_i$, player i learns that $\hat{x}(X) + b_d - b_i < x < \hat{x}(X)$ at some histories h^t that contain x' and that are on the equilibrium path with strictly positive probability μ. That is, player i's information $\omega_i(h^t)$ is such that there does not exist any state $x \notin (\hat{x}(X) + b_d - b_i, \hat{x}(X))$ from which a history $\hat{h}^t$ that does not contradict $\omega_i(h^t)$ can be reached with positive equilibrium μ probability. (Note that this is the case with $\tilde{X}_i = (\hat{x}(X) + b_d - b_i, \hat{x}(X))$ and every histories h^t that any contain $x' \in \tilde{X}_i$ when $i = e$, as the expert knows x at time 0.)

Then, i would prefer that d plays $\hat{y} = \hat{x}(X) + b_d$ rather than any action $\hat{y} < \hat{x}(X) + b_d$. This is because the bliss point $E[x|\omega_i(h^t), \mu] + b_i$ of i's expected utility $E[u_i\,(\hat{y}, x)\,|\omega_i(h^t), \mu]$ given information $\omega_i(h^t)$ and equilibrium beliefs μ is such that $E[x|\omega_i(h^t), \mu] + b_i > \hat{x}(X) + b_d - b_i + b_i = \hat{x}(X) + b_d$.

Indeed, i can secure that d plays $\hat{y} = \hat{x}(X) + b_d$ by blocking the transmission of any information on x, i.e., by sending message $\hat{m}_{ij} = X$ to his immediate successor j on the path $p(e, d)$ at time t. By doing so, i makes it impossible for any successor $j \neq d$ on the path $p(e, d)$ to send any message $\hat{m}_{jk}$ other than $\hat{m}_{jk} = X$ along the path $p(e, d)$. Hence, for every terminal history h^T on any equilibrium path that contain such histories h^t where i learns that $\hat{x}(X) + b_d - b_i < x < \hat{x}(X)$, the decision maker d's action $y_d(\omega_i(h^t), \mu) \geq \hat{x}(X) + b_d$. As a result, the realized loss is $(E[x|\omega_d(h^T), \mu] - x)^2 \geq (\hat{x}(X) - x)^2$. Integrating over $\tilde{X}_i$, and the histories h^t where i learns that $\hat{x}(X) + b_d - b_i < x < \hat{x}(X)$, the expected loss is:

$$E[Var(x|\omega_d(h^T), \mu)] \geq \int_{\tilde{X}_i} (\hat{x}(X) - x)^2 \mu(\hat{h}^t|x) f(x)dx > 0,$$

because $\tilde{X}_i$ has non-null measure.

Second, suppose that for almost all $x \in (\hat{x}(X)+b_d-b_i, \hat{x}(X))$, and μ-almost all histories h^t, player i does not learn that $\hat{x}(X)+b_d-b_i < x < \hat{x}(X)$ at h^t. In such a case, also player $\dot{d}$ will not learn that $\hat{x}(X)+b_d-b_i < x < \hat{x}(X)$ at any history h^T that contains any such history h^t, as everything that d learns must be known also to all players on the path $p(e,d)$ and equilibrium beliefs on the equilibrium path must be common across all players. As a result, it will be the case that the realized loss is $(E[x|\omega_d(h^T)]-x)^2 \geq \min\{(\hat{x}(X)+b_d-b_i-x)^2, (\hat{x}(X)-x)^2\}$ for almost all $x \in (\hat{x}(X)+b_d-b_i, \hat{x}(X))$, and the expected variance will be again strictly positive:

$$E[Var(x|\omega_d(h^T),\mu)] \geq \int_{\max\{\underline{x},\hat{x}(X)+b_d-b_e\}}^{\hat{x}(X)} \min\{(\hat{x}(X)+b_d-b_i-x)^2, (\hat{x}(X)-x)^2\} f(x)dx > 0.$$

Analogous arguments conclude that, for every player $i \neq d$ on $p(e,d)$ such that $b_i < b_d$, letting t by the distance of i from e. There are two possibilities. Either player i learns that $\hat{x}(X) < x < \hat{x}(X)+b_d-b_i$ on some non-null measure set $\tilde{X}_i$ and some histories h^t on path, and then

$$E[Var(x|\omega_d(h^T),\mu)] \geq \int_{\tilde{X}_i} (\hat{x}(X)-x)^2 \mu(\hat{h}^t|x) f(x)dx > 0,$$

or player i almost never learns that $\hat{x}(X) < x < \hat{x}(X)+b_d-b_i$ when this the case, and then

$$E[Var(x|\omega_d(h^T),\mu)] \geq \int_{\hat{x}(X)}^{\min\{\overline{x},\hat{x}(X)+b_d-b_e\}} \min\{(\hat{x}(X)-x)^2, (\hat{x}(X)+b_d-b_i-x)^2,\} f(x)dxf(x)dx > 0.$$

Because these arguments hold for any i and for any $\hat{x}(X)$, I obtain that, regardless of the (possibly off equilibrium path) equilibrium value of $\hat{x}(X)$, the expected loss $E[Var(x|\omega_d(h^T),\mu)]$ of any equilibrium μ is strictly positive, even as $\delta \to 1$. ■

Proof of Proposition 7. Suppose momentarily the optimal network is a tree. By the proof of Proposition 4, the ordered line is the only tree in which every pair of players e and d are connected through a path $p(e,d)$ without bias reversals. Using Proposition 6 and proceeding as in the proof of Proposition 4, I conclude that for all full support P, there exists a threshold $\bar{\delta}_1 < 1$ such that for all $\delta \in (\bar{\delta}_1, 1]$, the ordered line is the unique maximizer of welfare $W(N)$ among all trees N.

Now, let's compare the ordered line to networks that are not minimally connected. For δ sufficiently close to 1, unless the link cost c is too small, no connected network with loops can dominate the ordered line in terms of welfare, as the aggregate link cost is strictly higher. Likewise, for δ sufficiently close to 1, unless the link cost c is too high, any network that is not connected cannot dominate the ordered line, as some decision maker d would not receive some expert e's information.

So, there exist $\bar{\delta} < 1$ and an intermediate link cost range (c_{-}, c^{+}), with $\lim_{\delta \to 1} c_{-}(\delta) = 0$ and $\lim_{\delta \to 1} c^{+}(\delta) = \bar{\bar{c}}$, such that for all $c \in (c_{-}, c^{+})$ and $\delta \in (\bar{\delta}, 1]$, the unique optimal network N is the ordered line.

Turning to prove pairwise stability, because the ordered line has no bias reversal paths, Proposition 6 allows us to conclude that for all d, it is the case that $E[Var(x|\omega_d(h^T), \mu)] \to 0$ for $\delta \to 1$. Hence, for δ sufficiently close to one, there is no pair of players i and j who find it beneficial to form direct link, unless the cost c is too low. Proceeding as in the proof of Proposition 5, I obtain that for δ close enough to one, deleting links from the ordered line is detrimental unless the link c is too large. Hence, I obtain there exist a decay threshold $\bar{\delta} < 1$, and an intermediate cost range $(\hat{c}_{-}, \hat{c}^{+})$ with $\lim_{\delta \to 1} \hat{c}_{-}(\delta) = 0$, and $\lim_{\delta \to 1} \hat{c}^{+}(\delta) = \hat{c}$, such that for all $c \in (\hat{c}_{-}, \hat{c}^{+})$ and $\delta \in (\bar{\delta}, 1]$, the ordered line is pairwise stable. ■

Proof of Proposition 3. Suppose that $\delta = 1$ and consider a profile of lists ℓ_i, for $i \in \mathcal{N}$ such that $\ell_{ij}\ell_{ji} = 0$ for all pairs i, j, and $\ell_{ij} + \ell_{ji} = 1$ if and only if $|j - i| = 1$, i.e. i and j are consecutive indexes. The ordered line obtains, and by Proposition 3, each decision maker x decides fully informed. Because $c > 0$, no player i wishes to deviate from $\ell_{ij} = 0$ for any j. Unless c is too large, no player i wants to deviate from $\ell_{ij} = 1$ for any j either.

Suppose now that $\delta < 1$ and $P(e, d) = \frac{1}{n(n-1)}$ for all e, d. Suppose by contradiction that the ordered line obtains in a Nash equilibrium when $n \geq 6$ and the utility weights $\boldsymbol{\alpha} > \mathbf{0}$ are such that α_{ij} is arbitrarily small for all i and $j \neq i$, for some costs c. Consider the choices of players 3 and 4 to form a link with each other. By sponsoring a link with 4, player 3 gains access to the expert e's information when $e \geq 4$. For each such an expert e, the probability that x reaches player 3 is δ^{e-3}. Hence, because $P(e, d)$ is uniform across e, d, the total probability that x reaches 3 is $P(x|\ell_{34} = 1) = \frac{1}{n(n-1)} \sum_{e=4}^{n} \delta^{e-3}$. Player 3 can gain access to all experts $e \geq 4$ with less decay by forming a link with player 5, as in this case the total probability that x reaches 3 is $P(x|\ell_{35} = 1) = \frac{1}{n(n-1)}(2\delta + \sum_{e=6}^{n} \delta^{e-2})$, which is clearly larger than $P(x|\ell_{34} = 1)$. Hence, player 3 would not be willing to sponsor a link with 4, as she prefers to sponsor a link with 5. For the same reasons, player 4 is not willing to sponsor a link with 3, as she prefers to sponsor a link with 2.

To complete the analysis, I note that, for all $n < 6$ and utility weights $\boldsymbol{\alpha} > \mathbf{0}$ with α_{ij} arbitrarily small for all i and $j \neq i$, there is a range of intermediate costs for which the ordered line results as a Nash Equilibrium in which each link is sponsored by the "most moderate" player, i.e. $\ell_{i+1,1} = 1$ if $i \leq n/2$, $\ell_{i-1,1} = 1$ if $i \geq n/2$, and $\ell_{ij} = 0$ for all other i and j. For $n = 6$, instead, the optimal Nash equilibrium of the unilateral sponsorship game is as depicted in Figure 3, where the sponsor of each link is denoted by vertical dash. ■

Proof of Proposition 9. Part 1. Say the path $p(e, d)$ has no bias reversals, e.g. $i > d$ for all $i \neq d$ on $p(e, d)$. Then, proceeding as in the Proof of Proposition 3 there is a perfect Bayesian equilibrium, in which every $i \neq d$ plays $m_{ij}(h^t) = \omega_i(h^t)$ at every history h^t at the time t she is called to play, every player $j \neq e$ holds beliefs that $s = \min \omega_j(h^t)$ with probability one at every history h^t when j is called to play, and the decision maker d chooses $y_d(h^T) = y_d(\min \omega_d(h^T))$, for every history h^T. As a result, the message $\hat{m}_{ij} = \{s\}$ travels along the path $p(e, d)$ from e to d on the equilibrium path, and d chooses $y_d(h^T) = y_d(s)$, so that $E[L_d(y, x)|\omega_d(h^T); \mu] = E_s[EL_d(x|s)]$.

In order to prove this, first note that, as in Proof of Proposition 3, the beliefs of every player $j \neq e$ are admissible in perfect Bayesian equilibrium, and that d's strategy y_d is clearly optimal given her beliefs.

Now consider the decision of any player $i \neq d$ on the path $p(e, d)$ at time $t = \ell(p(e, i))$ and at any history h^t, and let j be i's immediate successor on the path $p(e, d)$. I first note that $\omega_j(h^t) = S$ and $\omega_i(h^t) \subseteq \omega_j(h^{t+1})$ for any history h^{t+1} that contains h^t. Further, because the expected utility $E[L_i(y, x)|\omega_i(h^t); \mu]$ is an integral of strictly concave loss functions $L_i(y, x)$, it is a also a strictly concave loss function.

Then, I show that player i does not gain by deviating from the equilibrium strategy $m_{ij}(h^t) = \omega_i(h^t)$. Proceeding backwards, consider any history h^{T-1} and the last player $i \neq d$ on the path $p(e, d)$. Recall that for any history h^T that includes h^{T-1}, it is the case that $\omega_i(h^{T-1}) \subseteq \omega_d(h^T)$. Then, because $y_d\left(h^T\right) = y_d(\min \omega_d(h^T))$, and $\min \omega_d(h^T) \leq \min \omega_i(h^{T-1})$ by set inclusion, I obtain that the bliss point $y_i(\omega_i(h^{T-1}); \mu)$ of i's expected utility $E[L_i(y, x)|\omega_i(h^{T-1}); \mu]$ given information $\omega_i(h^{T-1})$ and equilibrium beliefs μ is such that $y_i(\omega_i(h^{T-1}); \mu) \geq y_i(\omega_i(h^{T-1})) > y_d(\omega_i(h^{T-1})) \geq y_d(\omega_i(h^T))$, where the first inequality follows from verifiability, $s \in \omega_i(h^{T-1})$ and monotonicity of y_i in s, the second inequality follows from $i > d$. As the expected utility $E[L_i(y,x)|\omega_i(h^{T-1});\mu]$ is concave, it increases in y for $y > y_i(\omega_i(h^{T-1}))$. As a result, i would like to maximize y and, by verifiability, this can be done by sending message $\hat{m}_{id}^{T-1} = \omega_i(h^{T-1})$. Proceeding by induction as in Proof of Proposition 3, I obtain that each player $i \neq d$ plays $m_{ij}(h^t) = \omega_i(h^t)$ at every history h^t at the time t she is called to play.

I omit the proof that in every equilibrium μ, it is the case that the message $\hat{m}_{ij} = \{s\}$ travels along the path $p(e,d)$ from e to d on the equilibrium path, and the decision maker chooses $\hat{y} = y_d(s)$, so that $E[L_d(y,x|\omega_d(h^T),\mu)] = E_s[EL_d(x|s)]$ for all terminal histories h^T. This proof is an obvious generalization of the proof the same results in Proposition 3.

Part 2. Suppose now that the path $p(e,d)$ has bias reversals. Consider terminal histories h^T such that $\omega_d(h^T) = S$, which may or may not be on the equilibrium path. Consider decision $y_d(S;\mu)$, in case $\omega_d(h^T) = S$ if off the equilibrium path, then $y_d(S;\mu)$ is arbitrary, else it is determined by Bayes rule. Every player $i \neq d$'s message space is restricted to $M_{ij}^t = \{\{s\}, S\}$. Hence, for any history h^t, the information $\omega_i(h^t)$ that player i has on s may either be that $\omega_i(h^t) = \{s\}$ or that $\omega_i(h^t) = S$. If $\omega_i(h^t) = S$, then by verifiability $\hat{m}_{ij}^t = S$ and $\omega_d(h^T) = S$ for every terminal history h^T that contains h^t.

Pick any player $i > d$ on $p(e,d)$, and let t by the (possibly zero) distance of i from e. Consider the set $S_i(S;\mu) = \{s : y_d(s) < y_d(S;\mu) < y_i(s)\}$, for all $s \in S_i(S;\mu)$, player i would rather that d plays $\hat{y}_d = y_d(S;\mu)$ than $\hat{y} = y_d(s)$. Suppose that $s \in S_i(S;\mu)$, and suppose that $\omega_i(h^t) = \{s\}$. In any terminal history h^T such that $\omega_i(h^t) = \{s\}$, player d's decision is $\hat{y}_d = y_d(s)$. Because $y_d(s) < y_d(S;\mu) < y_i(s)$, player i strictly prefers that d receives no information about s and plays $\hat{y}_d = y_d(S;\mu)$, rather than d learns s and plays $\hat{y}_d = y_d(S;\mu)$. Indeed, player i can insure that s is not disclosed to d by sending message $\hat{m}_{ij}^t = S$, as this implies that $\omega_d(h^T) = S$ for every terminal history h^T that contains h^t. Hence, there cannot exist any equilibrium in which $\omega_d(h^T) = \{s\}$ on the equilibrium path, for any $s \in S_i(S;\mu)$. An analogous, symmetric, argument implies that there cannot exist any equilibrium in which $\omega_d(h^T) = \{s\}$ on the equilibrium path, for any $s \in S_i(S;\mu) = \{s : y_i(s) < y_d(S;\mu) < y_d(s)\}$ and any $i < d$. Hence, the expected loss can be bounded in a manner analogous to the proof of Proposition.3. ■

Proof of Proposition 10. The proof is omitted as it is the same as the proofs of Proposition 4 and 5, using Proposition 9 in lieu of Proposition 3. ■

Proof of Proposition 11. I first prove that if the network N is the ordered line, then there exists a perfect Bayesian equilibrium such that for all sets E of experts and D of decision makers, every $d \in D$ receives the signal s_e of every $e \in E$.

Define the set $I_-(E,D) = \{i : e \le i \le d$, for some $e \in E$ and $d \in D\}$, $I_+(E,D) = \{i : d \le i \le e$, for some $e \in E$ and $d \in D\}$. Because N is the ordered line, $I_+(E,D)$ identified the players i who are involved in information transmission and decision from experts in E on the right to decision makers in D on the left, and $I_-(E,D)$ vice versa. Let $I(E,D) = I_-(E,D) \cup I_-(E,D)$ be the set of players that are not idle in the game.

Consider a profile of strategies in which every player $i \in I_+(E,D)$ reveals all her information $\omega_i(h^t)|_e \subseteq S$, $m_{ij}(h^t)|_e = \omega_i(h^t)|_e$, for all $e \in E$ such that $e \ge i$ to her successor $j = i-1 \in I_+(E,D)$ in every history h^t at which she is called to play. (What i communicates for $s_e : e < i$ to $j = i-1 \in I_+(E,D)$ is irrelevant.) Symmetrically, every $i \in I_-(E,D)$ reveals all her information $\omega_i(h^t)|_e$, $m_{ij}(h^t)|_e = \omega_i(h^t)|_e$, for all $e \in E$ such that $e \le i$ to her successor $j = i+1 \in I_-(E,D)$ in every history h^t at which she is called to play. At period T, every player $d \in D$, chooses $y_d(h^T) = E[x|\mathbf{s}_{Ed}(\omega_d(h^T))] + b_d$, where $\mathbf{s}_{Ed}(\omega_d(h^T))$ is the profile of signals s_e such that $s_e = \min \omega_d(h^T)|_e$ for all $e \in E$ such that $e > d$ and $s_e = \max \omega_d(h^T)|_e$ for all $e \in E$ such that $e < d$. Note that because N is the ordered line, on the equilibrium path, every signal s_e of every $e \in E$ travels precisely to any $d \in D$, who chooses $y_d(h^T) = E[x|s_E] + b_d$, so that $E[Var(x|\omega_d(h^T),\mu)] = E[Var(x|s_E)]$.

For any player $e \in E$, let the equilibrium beliefs of every player $j \in I(E,D)$ in every history h^t at which j is called to play be such that $s_e = \min \omega_j(h^t)|_e$ with probability one if $j < e$, and that $s_e = \max \omega_j(h^t)|_e$ with probability one if $j > e$.

In order to prove that this is a perfect Bayesian equilibrium, first note that, as in the proof of Proposition 3, these beliefs are admissible. Consider any player $e \in E$ and player $j \in I(E,D)$ with $j \ne e$ and every signal s_e, the unique information set $\omega_j(h^t)$ at any history h^t on the equilibrium path at which j is called to play is such that $\omega_j(h^t)|_e = \{s_e\}$, and $\min \omega_j(h^t)|_e = \max \omega_j(h^t)|_e = \{s_e\}$ consistently. Further, d's strategy y_d is clearly optimal given her beliefs.

To continue with the proof, consider any $i \in I(E,D)$, note that because N is the ordered line, for any history h^t, $e \in E$ with $e \ge i$, and $j < i$, it is the case that $\omega_i(h^t)|_e \subseteq \omega_j(h^{t+1})|_e$ for any history h^{t+1} that contains any history h^t: as in the proof of Proposition 3, everything that $j < i$ learns about signal s_e for which $e \geq i$, she learns from i. Further, decompose player i's expected utility $E[L_i(\hat{\mathbf{y}}_D, x)\,|\omega_i(h^t), \mu] = \sum_{d \in D} \alpha_{id} E[L_i(\hat{y}_d, x)\,|\omega_i(h^t), \mu]$ at history t and note that for any $d \in D$, the expression $E[L_i(\hat{y}_d, x)\,|\omega_i(h^t), \mu]$ is a a quadratic loss function, as it is the integral of the quadratic loss functions $L_i(\hat{y}_d, x) = -(\hat{y}_d - x - b_i)^2$.

I now show that no player $i \in I_+(E, D)$ gains by deviating from the equilibrium strategy $m_{ij}(h^t)|_e = \omega_i(h^t)|_e$ for all $e \in E$ with $e \geq i$ and $j = i - 1 \in I_+(E, D)$. The proof that no player $i \in I_-(E, D)$ gain by deviating from the equilibrium strategy $m_{ij}(h^t)|_e = \omega_i(h^t)|_e$ for all $e \in E$ with $e \leq i$ and $j = i+1 \in I_-(E, D)$ is its mirror like image and hence omitted.

In order to show this result, I proceed backwards along the set $I_+(E, D)$, let $\hat{T}(E, D) = \max E - \min D$ be the length of the path from the most right-wing expert to the most leftist decision maker, consider any history $h^{\hat{T}(E,D)-1}$ and let $d = \min D$ and $i = \min D + 1$. Using a result just above above, for any history $h^{\hat{T}(E,D)}$ that includes $h^{\hat{T}(E,D)-1}$, it is the case that $\omega_i(h^{\hat{T}(E,D)-1})|_e \subseteq \omega_d(h^{\hat{T}(E,D)})|_e$ for all $e \in E\backslash\{d\}$, because $d < i$ and all $e \geq i$ for $e \in E\backslash\{d\}$.

Further, because $i > d$ and N is the ordered line, it follows that i is not on the path $p(e, d)$ from any $e \in E$ such that $e < d$. As consequence, i has no influence on the information that d receives about signals s_e such that $e < d$. In equilibrium, i knows that every signal signals s_e such that $e < d$ will reach d precisely. Because by construction, $\hat{T}(E, D) - 1$ is the last round of transmission to d of any information of any signal s_e such that $e > d$, player i knows that in equilibrium, d's decision will be $E[y_d(h^T)|\omega_i(h^{\hat{T}(E,D)-1})] = E[E[x|\left(\mathbf{s}^+_{Ed}(\omega_d(h^T)), \mathbf{s}^-_{Ed}\right)]|\omega_i(h^{\hat{T}(E,D)-1})] + b_d$, where $\mathbf{s}^-_{Ed}$ is any profile of signals $(s_e)_{e \leq d}$ and $\mathbf{s}^+_{Ed}(\omega_d(h^T))$ is the profile of signals $(s_e)_{e>d}$ such that $s_e = \min \omega_d(h^T)|_e$ for all $e > d$, and the external expectation is taken with respect to $\mathbf{s}^-_{Ed}$.

Proceeding as in the proof of Proposition 3, I now note that, for every $e > d$, $\min \omega_d(h^T)|_e \leq \min \omega_d(h^{\hat{T}(E,D)-1})|_e$ by set inclusion, $E[x|\left(\mathbf{s}^+_{Ed}(\omega_d(h^T)), \mathbf{s}^-_{Ed}\right)] \leq E[x|\left(\mathbf{s}^+_{Ed}(\omega_d(h^{\hat{T}(E,D)-1})), \mathbf{s}^-_{Ed}\right)]$ by single-crossingness, and $E[x|\left(\mathbf{s}^+_{Ed}(\omega_d(h^{\hat{T}(E,D)-1})), \mathbf{s}^-_{Ed}\right)] \leq E[x|\omega_d(h^{\hat{T}(E,D)-1})|_{e \geq d} \times \mathbf{s}^-_{Ed}, \mu]$ by verifiability. Integrating across $\mathbf{s}^-_{Ed}$, I obtain that $E[E[x|\left(\mathbf{s}^+_{Ed}(\omega_d(h^T)), \mathbf{s}^-_{Ed}\right)]|\omega_i(h^{\hat{T}(E,D)-1})] \leq E[x|\omega_d(h^{\hat{T}(E,D)-1}), \mu]$.

Hence, I have obtained that the bliss point $E[x|\omega_i(h^{\hat{T}(E,D)-1}), \mu] + b_i$ of i's expected utility $E[L_i(\hat{y}_d, x)\,|\omega_i(h^{\hat{T}(E,D)-1}), \mu]$ given information $\omega_i(h^{\hat{T}(E,D)-1})$ and equilibrium beliefs μ is such that $E[x|\omega_i(h^{\hat{T}(E,D)-1}), \mu] + b_i > E[E[x|\left(\mathbf{s}^+_{Ed}(\omega_d(h^T)), \mathbf{s}^-_{Ed}\right)]|\omega_i(h^{\hat{T}(E,D)-1})] + b_d =$

$E[y_d(h^T)|\omega_i(h^{\hat{T}(E,D)-1})]$, where the first inequality is because $b_i > b_d$.

As in the proof of Proposition 3, the expected utility $E[L_i(\hat{y}_d, x)|\omega_i(h^{\hat{T}(E,D)-1}), \mu]$ of player i from d's choice $\hat{y}_d$ increases in $\hat{y}_d$, for $\hat{y}_d > E[y_d(h^T)|\omega_i(h^{\hat{T}(E,D)-1})]$. As a result, i would like to maximize $\hat{y}_d$ and, by verifiability, this can be done by sending message $\hat{m}_{id}^{\hat{T}(E,D)-1}|_e = \omega_i(h^{\hat{T}(E,D)-1})|_e$ for all $e \geq i$.

The argument by induction is then similar to the one in the proof of Proposition 3. For any $t = 0, ...\hat{T}(E,D) - 1$, let i_t be the player in the set $I_+(E,D)$ at distance t from $\max E$. I have proved above that at any history $h^{\hat{T}(E,D)-1}$ player $i_{\hat{T}(E,D)-1}$ sends message $\hat{m}_{i_{\hat{T}(E,D)-1}i_{\hat{T}(E,D)}}|_e = \omega_{i_{\hat{T}(E,D)-1}}(h^{\hat{T}(E,D)-1})|_e$ to $i_{\hat{T}(E,D)} = d$ for all $e \in E$ with $e \geq i$. I now proceed by induction. For any $t < \hat{T}(E,D) - 1$, suppose that at any history $h^{t+\tau}$, $\tau = 1, ..., \hat{T}(E,D) - 1 - t$, player $i_{t+\tau}$ sends message $\hat{m}_{i_{t+\tau}i_{t+\tau+1}}|_e = \omega_{i_{t+\tau}}(h^{t+\tau})|_e$ to $i_{t+\tau+1}$ for all $e \in E$ with $e \geq i$.

Now, consider the choice faced by player i_t at any history h^t with respect to the transmission to i_{t+1} of information about s_e such that $e \in E$ and $e \geq i$. Again, for any $e \in E$ with $e \geq i$, and $d < i$, it is the case that $\omega_i(h^t)|_e \subseteq \omega_d(h^{t+1})|_e$ for any history h^{t+1} that contains history h^t. Because of the induction hypothesis, for any $e \in E$ with $e \geq i_t$, and $d < i_t$, it is the case that $\omega_d(h^T)|_e = \hat{m}_{i_t i_{t+1}}|_e$. Because N is the ordered line, all players $d' \in D$ that are connected to i_t through i_{t+1} are such that $d' < i_{t+1} < i$, and hence $b_d < b_i$. The same arguments applied to $i_{\hat{T}(E,D)-1}$ and d implies that also for all such d', the expected utility $E[L_i(\hat{y}_{d'}, x)|\omega_i(h^t), \mu]$ of player i_t from d''s choice $\hat{y}_{d'}$ increases in $\hat{y}_{d'}$ for $\hat{y}_{d'} > E[y_{d'}(h^T)|\omega_i(h^t)]$. So, again, it is a best response for player i_t at history h^t to reveal all information and send message $\hat{m}_{i_t i_{t+1}} = \omega_{i_t}(h^t)$.

I have therefore concluded that the stated conjectured equilibrium profile is indeed a Perfect Bayesian Equilibrium.

The proof that the ordered line N is the unique optimal network is then concluded from the following results. First, for every realized E and D it achieves the minimal expected loss $\sum_{d \in D} \alpha_{id} E[Var(x|\omega_d(h^T), \mu); N] = \sum_{d \in D} \alpha_{id} E[Var(x|s_E)]$. Second, for every other tree N, I proved in the proof of Proposition 4 that there exist singleton realizations $E = \{e\}$ and $D = \{d\}$ such that the expected loss is $E[Var(x|\omega_d(h^T), \mu); N] > E[Var(x|s_e)]$. Third, and final, for any link cost $c > 0$, adding links to the ordered line is wasteful, whereas for any link cost $c < \bar{c}$, deleting links from the ordered line is suboptimal, as it leads to decision makers $d \in D$ to lose all information about signal s_e for some $e \in E$, for some realizations of E and D. ■

Completion of Example 2. Consider the choice of player 3 of reporting information $\hat{m}_{32}|_3$ about signal s_3. The equilibrium payoff of player 3 from the choices of players 2 and 4 as a function of $\hat{m}_{32}|_3$ is:

$$\begin{aligned} u_3(\hat{m}_{32}|_3|s_3)|_{2,4} &= -\alpha_3 E[(a\Sigma(\min \hat{m}_{32}|_3, \mathbf{s}_{-3}) + z + b_2 - x - b_3)^2 \\ &\quad + (a(\max \hat{m}_{32}|_3, \mathbf{s}_{-3}) + z + b_4 - x - b_3)^2 \, |s_3], \end{aligned}$$

where the expectation is taken with respect to s_1, s_2, s_4 and x. Instead, the equilibrium payoff of player 3 from the choices of players 2 and 4 as a function of 3's message $\hat{m}_{32}|_1$ about signal s_1 is:

$$\begin{aligned} u_3(\hat{m}_{32}|_1|\omega_3(h^2))|_{2,4} &= -\alpha_3 E[(a\Sigma(\max \hat{m}_{32}|_1, \mathbf{s}_{-1}) + z + b_2 - x - b_3)^2 \\ &\quad + (a\Sigma(\max \hat{m}_{32}|_1, \mathbf{s}_{-1}) + z + b_4 - x - b_3)^2 \, |\omega_3(h^2)]. \end{aligned}$$

Consider the latter first. For any history h^2 at which 3 is called to reveal information $\omega_3(h^2)|_1$ about s_1 to 2, it is optimal for 3 not to withhold information, and send $\hat{m}_{32}|_1 = \omega_3(h^2)|_1$. In fact, simplifying the expression of u_3, and using the short-hand notation $\hat{\Sigma}_1 = \Sigma(\max \hat{m}_{32}|_1, \mathbf{s}_{-1})$, I obtain:

$$\begin{aligned} u_3(\hat{m}_{32}|_1)|_{2,4} &= -\alpha_3 E\left[\left(a\hat{\Sigma}_1 + z - x - b\right)^2 + \left(a\hat{\Sigma}_1 + z + b - x\right)^2 |\omega_3(h^2)\right] \\ &= -2\alpha_3 E\left[\left(a\hat{\Sigma}_1 + z - x\right)^2 |\omega_3(h^2)\right] - 2\alpha_3 b^2. \end{aligned}$$

On the equilibrium path, 2 knows s_1 at $t = 2$, and by definition of $E[x|\mathbf{s}]$, the expression $u_3(\hat{m}_{32}|_1)|_{2,4}$ is maximized by setting $a\hat{\Sigma}_1 + z = E[x|\mathbf{s}]$, i.e. by revealing s_1 through the message $\hat{m}_{32}$ such that $\hat{m}_{32}|_1 = \{s_1\}$. Off the equilibrium path, for any history h^2, the expression $u_3(\hat{m}_{32}|_1)|_{2,4}$ is maximized by sending message $\hat{m}_{32} = \omega_3(h^2)$. In fact, the expected state given information $\omega_3(h^2)$ is $E[E[x|\omega_3(h^2)|_1, \mathbf{s}_{-1}]|\omega_3(h^2)] < E[E[x| \max \omega_3(h^2)|_1, \mathbf{s}_{-1}]|\omega_3(h^2)] \leq E[E[x| \max \hat{m}_{32}|_1, \mathbf{s}_{-1}]|\omega_3(h^2)]$, where the first inequality follows from by the intermediate value theorem, and the second inequality because $\omega_3(h^2) \subseteq \hat{m}_{32}$.

A fortiori, player 3 maximizes the expression $u_3(\hat{m}_{32}|_3|s_3)|_{2,4}$ by revealing s_3 through the message $\hat{m}_{32}$ such that $\hat{m}_{32}|_3 = \{s_3\}$. By withholding information through a message

$\hat{m}_{32}$ such that $\{s_3\} \subsetneq \hat{m}_{32}|_3$, player 3 would bias the decisions of both 2 and 4 away from his bliss point $E[x|\mathbf{s}] + b_3$, for any realizations of $\mathbf{s}_{-3}$. In fact, player 3 would bias 2's decision $\hat{y}_2 = E[x|\min \hat{m}_{32}|_3, \mathbf{s}_{-3}] + b_2$ leftward away from $E[x|\mathbf{s}] + b_3$, and 4's decision $\hat{y}_4 = E[x|\max \hat{m}_{32}|_3, \mathbf{s}_{-3}] + b_2$ rightward also away from $E[x|\mathbf{s}] + b_3$. ■