Journal of Public Economic Theory

WILEY

JOURNAL OF PUBLIC ECONOMIC THEORY

ORIGINAL ARTICLE

Persuasion in Networks With Strategic Substitutes

Guopeng Li[1] | Yang Sun[2]

[1]School of Economics, Huazhong University of Science and Technology, Wuhan, China | [2]School of Economics, Southwestern University of Finance and Economcics, Chengdu, China

Correspondence: Yang Sun (sunyang789987@gmail.com)

Received: 24 June 2024 | **Revised:** 29 October 2024 | **Accepted:** 17 December 2024

Funding: This study received financial support from the National Natural Science Foundation of China (No. 72173049, 72203184) and the MOE project of Humanities and Social Sciences (No. 22YJC79111).

Keywords: Bayesian persuasion | downwardplaying mechanism | k-insulated set | local substitutes

ABSTRACT

We study Bayesian persuasion with local strategic substitutes in networks. A designer commits to a public signal to maximize total activity. Equilibria are characterized by the network's maximum k-insulated sets for each realization. We solve the optimal information structure and characterize beneficial persuasion. While agents individually prefer higher states, the designer's payoff is non-monotonic in the posterior mean due to substitution effects. This provides a rationale for downwardplaying mechanisms: revealing low states truthfully and mixing signals when high. Moreover, for tree, nested split, and core-periphery networks, the designer strictly benefits if the prior mean insulated set size is less than the highest state set size.

JEL Classification: C72, D82, D85

1 | Introduction

Bayesian persuasion models how an informed sender can influence receivers' actions via strategic communication. Kamenica and Gentzkow (2011) characterize optimal sender outcomes for a single receiver, extendable to multiple isolated receivers with public signaling. However, many real-world situations feature networked agents with interdependent payoffs. Recent work like Arieli and Babichenko (2019) and Mathevet, Perego, and Taneva (2020) shows how private signals can manipulate multiagent beliefs, but optimal information design in networks remains an open challenge. A key question is how an informed designer can construct public or private signals to induce desired actions given complex strategic interactions.

Consider the provision of public goods in a network setting, where agents' incentives to contribute are influenced by the actions of their neighbors. Examples include research collaborations, R&D spillovers, information acquisition, and infrastructure investments (see Bramoullé and Kranton 2007; Bramoullé, Kranton, and D'Amours 2014; Allouch 2015, for detailed discussions on public goods provision games in networks). In these contexts, agents' actions often exhibit strategic substitutes—an agent's incentive to contribute decreases as more of their neighbors contribute. A central planner or government entity may seek to persuade agents to contribute more to the public good by strategically revealing information about its value or productivity. For example, in R&D networks, firms may benefit from the innovations of their partners. However, the value of these spillovers may be uncertain, depending on factors such as market demand. An industry association could have additional information about these factors and might strategically disclose it to encourage investment in R&D. The optimal persuasion mechanism is complicated by the strategic interactions among firms in the network.

We model this public goods provision problem as a network game with strategic substitutes. Agents take binary actions—to contribute or not contribute—and their payoffs depend on both the unknown productivity of the public good and the choices of their neighbors. The central planner (designer) reveals public productivity-based signals to persuade more agents to contribute.

Our main contribution is to characterize how the structure of the underlying network shapes optimal information design in public goods provision games. We find that the size of the network's k-insulated sets, a concept proposed by Jagota, Narasimhan, and Šoltés (2001), plays a central role. In a k-insulated set, each node connects to at most k others in the set, while nodes outside connect to at least $k + 1$ nodes within it. The maximum k-insulated sets for each k directly determine equilibrium behavior under different posteriors. Finding the optimal combination of Bayes plausible posteriors and corresponding insulated sets maximizes the designer's payoff for a given prior.

Interestingly, for some network structures, the size of maximum k-insulated sets decreases with k. This implies more contributors when the perceived productivity of the public good is low. As a result, unlike single agent case, the designer's optimal strategy may reveal low productivity truthfully while randomizing signals when productivity is high. This "downwardplaying" approach might seem counterintuitive, but it can be optimal due to the strategic substitution effects in the network. A similar phenomenon occurs in the context of luxury fashion goods, where ads sometimes exhibit exclusivity or unflattering styles contrary to typical preferences (Simmel 1957).

We also characterize the conditions for beneficial persuasion. If the size of the insulated set corresponding to the prior mean is suboptimal, there must exist a distribution of Bayes-plausible posteriors that improves the designer's payoff. Moreover, for tree, nested split, and core-periphery networks, the designer always prefers to signal high productivity since the insulated set sizes grow monotonically with k in these structures. Thus, the designer benefits from persuasion when the size of the prior mean insulated set is smaller than that of the high-productivity insulated set. Finally, we characterize the conditions that the optimal mechanism is mixed in the sense that the designer never truthfully reveals any state, regardless of whether it is high or low.

This paper relates to vast literature in information design, particularly information revelation to multiple agents (see the fundamental works by Bergemann and Morris 2016, 2019; Mathevet, Perego, and Taneva 2020). Some works focus on the effect of persuasion in auction contexts (Alonso and Câmara 2016), voting games (Schnakenberg 2015; Alonso and Câmara 2016), and conflict prevention (Hennigs 2021). In this study, we investigate the impact of persuasion in public good provision games on networks. The most relevant work is Candogan (2022) studying persuasion in strategic complement games where the q-core indexes outcomes. This paper complements Candogan (2022) by focusing on strategic substitute games and showing the network's k-insulated set is key.[1] Arieli and Babichenko (2019) examine supermodular utilities with private signals. As a complementary, we focus on local strategic substitutes with public signals.

Furthermore, our research aligns with the literature on public good provision on networks, which originated with Bramoullé and Kranton (2007), where an agent's contribution substitutes for neighbors' contributions. Bramoullé, Kranton, and D'amours (2014) and Allouch (2015) analyze equilibrium uniqueness in these games, while Elliott and Golub (2019) focus on Pareto efficiency under substitutes. Our model connects the network public goods literature with Bayesian persuasion, elucidating how network interactions shape signaling incentives.

Another related area of literature is models of fashion originated by Karni and Schmeidler (1990). Most works in this area focus on pricing strategy (e.g., Adachi 2005; Bhattacharya et al. 2011; Cao et al. 2017), rather than advertising. Yoganarasimhan (2012) studies information revelation in the fashion goods market, proposing a new framework to explain communication strategies that cloak information on product tastefulness. In contrast, this paper utilizes a standard Bayesian persuasion model to rationalize the advertisements that downplay product quality.

The remainder of the paper proceeds as follows. Section 2 presents an illustrative example to build intuition and preview the main results. In Section 3, we formally introduce the model setup and key definitions. Our central results characterizing optimal information design are presented in Section 4. Section 5 is discussion and conclusion. All detailed proofs are relegated to the Appendix.

2 | An Illustrative Example

Consider a consumer deciding whether to purchase a fashion good of unknown quality $\omega \in \{\omega_L, \omega_H\} = \{-1, 1\}$. The utility of the consumer is simply the multiplication of the action and quality:

$$u(a, \omega) = a \cdot \omega,$$

where $a = 1$ if the consumer buys the good, and $a = 0$ otherwise. A seller aims to persuade the consumer to purchase through quality-based signals. If the consumer's prior assigns lower probability to high quality $\omega = 1$, the optimal strategy for the seller is to reveal the quality truthfully when $\omega = 1$ and confuse the consumer when $\omega = -1$ such that the consumer is indifferent between both actions upon receiving the high-quality signal.

Now, consider the case of multiple strategic substitute consumers. Each individual wants to differentiate from peers while still enjoying the good. Though high quality independently incentivizes purchase, the optimal mechanism may work oppositely: revealing low quality truthfully while mixing signals on high quality. This inversion is driven by substitution effects. The following carefully constructed example in Figure 1

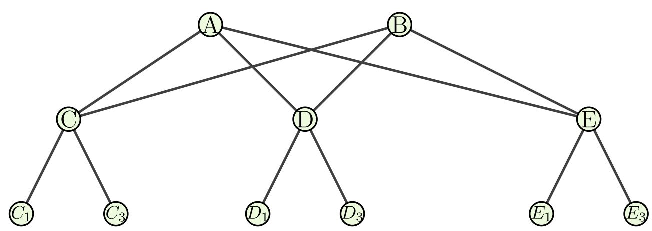


FIGURE 1 | Network **G**.

illustrates how consumer interactions modify signaling incentives away from the single agent case.[2]

There are 11 consumers embedded in an undirected network $\mathbf{G}$, shown in Figure 1. The payoff of each consumer i is

$$u_i(\mathbf{a}_{-i}, a_i, \omega) = a_i\left(\omega - \sum_j g_{ij}a_j\right),$$

where $\omega \in \{\omega_L, \omega_H\} = \{2.5, 3.7\}$. The common prior assigns equal probabilities to both quality levels, so the expected quality is 3.1. Therefore, a consumer purchases if and only if no more than 3 neighbors bought. It is easy to see that, in the equilibrium, there are 8 consumers buying: $\{C, D, E\}$ do not purchase while all others buy.

The crucial point in this example is that a lower expected quality increases buyers. Assume that the expected quality drops below 3, which implies that a consumer purchases if and only if no more than 2 neighbors did, there exists an (sender preferred) pure strategy equilibrium in which there are 9 buyers as A and B abstain from buying.

Unlike the single consumer case where the seller always prefers signaling high quality, this delicate network structure reverses the seller's preferences. The optimal strategy is downwardplaying: truthfully reveal the low quality $\omega_L = 2.5$ while mixing signals on high quality. Specifically, the optimal mechanism is

$$\pi(s_{AB}|\omega_L) = 1;\ \pi(s_{AB}|\omega_H) = \frac{5}{7};$$
$$\pi(s_{CDE}|\omega_L) = 0;\ \pi(s_{CDE}|\omega_H) = \frac{2}{7},$$

where s_{AB} and s_{CDE} represent recommendations for the subscripted consumers not to buy. This mechanism achieves an expected $\frac{62}{7}$ buyers, greater than 8 without persuasion. The seller benefits from truthfully revealing low quality in this network. The optimality of such downplaying mechanism provides a rationale for the use of exclusive or unflattering styles in luxury fashion advertising.

With an isolated consumer, the seller optimally reveals high quality truthfully since it directly incentives purchase. However, with multiple consumers, as illustrated in this example, the delicately constructed network results in the seller preferring to reveal the low quality. This reversal of preferences is driven by the strategic interactions between consumers. While an individual has a higher incentive to buy at high quality, this may discourage other consumers from buying due to substitution effects. The seller must account for this equilibrium response across consumers.

It is worth noting that the model of this fashion good example can be generalized to other settings, such as public goods provision games studied by Bramoullé and Kranton (2007) and Allouch and King (2019). In these contexts, the buyers can be seen as public good providers, while non-buyers are free riders. The strategic interactions and persuasion mechanisms discussed in this example are applicable to various scenarios where agents' actions exhibit substitution effects, including research collaboration, R&D spillovers, and infrastructure investment. In the following sections, we characterize the outcomes the designer can achieve through strategic signaling and the conditions for beneficial persuasion in various network topologies.

3 | Set Up

The primitives of the model consist of two essential ingredients: the network game and information design.

3.1 | Network Game

There is a set of agents, $N = \{1, \ldots, n\}$, embedded in a social network which is represented by an adjacency matrix $\mathbf{G} = (g_{ij})$. We consider an unweighted and undirected network here, where $g_{ij} = g_{ji} = 1$ indicates that agents i, j are neighbors, and $g_{ij} = g_{ji} = 0$ otherwise.[3] All agents $i \in N$ make a binary choice $a_i \in \{0, 1\}$ and the vector $\mathbf{a} = (a_1, \ldots, a_n)$ collects their choices. The agent is called active and inactive when he chooses 1 and 0, respectively. We impose the following assumption on the agents' preferences.

Assumption Best response function. Agent i's best responses to activity $\mathbf{a}_{-i}$ is given by

$$x_i^*(\mathbf{a}_{-i}) = \begin{cases} 1 & \text{if } \sum_j g_{ij}a_j \le \omega \\ 0 & \text{if } \sum_j g_{ij}a_j \ge \omega \end{cases}, \tag{1}$$

where $\omega \in \mathbb{R}$ is a realization of a payoff-relevant state.

The game exhibits strategic substitution—an agent desires to be active only when the total activity of his neighbors is smaller than threshold ω. The best response function is consistent with various classical models of public goods provision games in networks, such as Bramoullé and Kranton (2007), Bramoullé, Kranton, and D'amours (2014), Allouch (2015), Allouch and King (2019), and Gerke et al. (2024) which are applicable to research collaborations, R&D spillovers, information acquisition, and infrastructure investments. Below, we introduce two specific utility functions that satisfy this assumption.

- The utility of agent i is

$$u_i(\mathbf{a}_{-i}, a_i, \omega) = a_i\left(\omega - \sum_j g_{ij}a_j\right). \tag{2}$$

This utility form is a modification of Candogan (2022). Apparently, agent i chooses $a_i = 1$ when $\omega - \sum_j g_{ij}a_j \ge 0$ and $a_i = 0$ when $\omega - \sum_j g_{ij}a_j \le 0$.[4]

- The utility of agent i is

$$u_i(\mathbf{a}_{-i}, a_i, c) = b\left(a_i + \sum_j g_{ij} a_j\right) - c a_i,$$

where b is a concavely increasing function. Bramoullé and Kranton (2007) introduce this utility form, which is further explored by Allouch (2015) and Gerke et al. (2024), to study public goods provision games. Let $\bar{a}$ be the value that $b'(\bar{a}) = c$. Then, the parameter ω is the integer part of $\bar{a}$.

3.2 | Information Design

The payoff-relevant state $\omega \in \mathbb{R}$ is either high ω_H or low ω_L, whose distribution $\mu_0 \in \text{int}(\Delta(\{\omega_L, \omega_H\}))$ is the common prior of the agents.[5]

We consider the problem of an information designer who chooses a signaling mechanism π to influence the agents' choices. The designer is omniscient in the sense that he knows the true state ω advantage over the agents, and thus, the signaling mechanism π depends on ω. Following Candogan (2022), we focus on public signals observed by all agents. With public signals, Galperti and Perego (2023) prove revelation principle applies.[6] Therefore, we restrict attention to direct mechanisms,

$$\pi : \{\omega_L, \omega_H\} \to \Delta(A),$$

throughout this paper unless otherwise specified.

Given a direct mechanism π, for any $\mathbf{a} \in A$ such that $\sum_\omega \mu_0(\omega)\pi(\mathbf{a}|\omega) > 0$, define $\mu_{\mathbf{a}}(\omega) = \frac{\mu_0(\omega)\pi(\mathbf{a} \mid \omega)}{\sum_{\omega'} \mu_0(\omega')\pi(\mathbf{a} \mid \omega')}$ as the posterior belief of ω when the agents received signal $\mathbf{a}$. Let $\mathbb{E}_{\mu_{\mathbf{a}}} = \sum_\omega \mu_{\mathbf{a}}(\omega) \cdot \omega$ denote the posterior mean of ω under signal $\mathbf{a}$, and $\mathbb{E}_{\mu_0}$ is the prior mean when agents have no signal.

The direct mechanism π is incentive-compatible, denoted π^*, if for each $i \in N$ and $\mathbf{a} \in A$ such that $\sum_\omega \mu_0(\omega)\pi(\mathbf{a}|\omega) > 0$, we have

$$a_i = \begin{cases} 1 & \text{when } \sum_j g_{ij} a_j \leq \mathbb{E}_{\mu_{\mathbf{a}}} \\ 0 & \text{when } \sum_j g_{ij} a_j \geq \mathbb{E}_{\mu_{\mathbf{a}}} \end{cases}. \tag{3}$$

That is, every agent i obeys the designer's recommendation provided that he knows the others' recommendations $\mathbf{a}_{-i}$ via the public signal. The incentive compatible mechanism is a special class of Bayes correlated equilibria (Bergemann and Morris 2016) where the designer publicly announces the draw of action profile. In the following, we refer to this special category of Bayesian correlated equilibrium as the Bayes correlated equilibrium.

The designer's objective is to maximize the expected number of active agents. Formally, let Π^* denote the set of incentive-compatible direct mechanisms. The designer solves:

$$\max_{\pi^* \in \Pi^*} \sum_{\omega, \mathbf{a}, i} \mu_0(\omega)\pi^*(\mathbf{a}|\omega) \cdot a_i. \tag{4}$$

In the scenario of public goods provision, the designer wants to persuade more agents to provide the public goods. In the context of fashion good selling, the designer aims to maximize the total number of buyers.

In the single agent case, as shown by Kamenica and Gentzkow (2011), the designer can (weakly) improve his benefit by recommending 1 when $\omega \geq 0$ and incentive compatibly mixing recommendations $\{0, 1\}$ when $\omega < 0$. However, in the presence of a network, the designer needs a delicate mechanism tailored to the network structure to improve his benefit. Unlike single agents, some structures induce an optimal mechanism opposite to always revealing individual favorable information (high state).

4 | Main Result

Before analyzing the designer's problem, we first introduce a graph theory concept to describe the Bayes correlated equilibria. Let $N_i = \{j : g_{ij} = 1\}$ denote the set of i's neighbors in the network.

Definition 1. Given a network $\mathbf{G}$ and nonnegative integer k, a set of nodes $S \subseteq N$ is a k-insulated set if for each $i \in S$ we have $|N_i \cap S| \leq k$, and for each $i \notin S$ we have $|N_i \cap S| \geq k + 1$. A maximum k-insulated set is a k-insulated set with the largest cardinality. Let $\nu(k)$ denote the cardinality of maximum k-insulated set.

The concept of k-insulated set is introduced by Jagota, Narasimhan, and Šoltés (2001). They proved the existence of k-insulated set for any nonnegative integer k and any network. Allouch and King (2019) generalize this notion to characterize Nash equilibria in constrained public goods network games. When $k = 0$, a k-insulated set reduces to a maximal independent set in the network. By convention, we assume k-insulated set is empty if $k < 0$.

Let $\lfloor \mathbb{E}_{\mu_{\mathbf{a}}} \rfloor$ denote the integer part of posterior mean $\mathbb{E}_{\mu_{\mathbf{a}}}$. Then, $\lfloor \mathbb{E}_{\mu_{\mathbf{a}}} \rfloor$-insulated sets naturally characterize Bayes correlated equilibria, since an active agent must connect to at most $\lfloor \mathbb{E}_{\mu_{\mathbf{a}}} \rfloor$ other active agents, while an inactive agent must connect to at least $\lfloor \mathbb{E}_{\mu_{\mathbf{a}}} \rfloor + 1$ active agents. This observation is formalized by the following lemma.

Lemma 1. *Consider a direct mechanism π, then the mechanism is incentive-compatible if and only if for any $\mathbf{a} \in A$ such that $\sum_\omega \mu_0(\omega)\pi(\mathbf{a}|\omega) > 0$, the set of active agents $\{i : a_i = 1\}$ is a $\lfloor \mathbb{E}_{\mu_{\mathbf{a}}} \rfloor$-insulated set or a $(\lfloor \mathbb{E}_{\mu_{\mathbf{a}}} \rfloor - 1)$-insulated set.*

Lemma 1 characterizes the set of incentive-compatible direct mechanisms. In particular, for any realized state $\omega \in \{\omega_L, \omega_H\}$, any incentive-compatible direct mechanism induces a recommendation where all agents in an insulated set choose 1, while agents outside the insulated set are inactive. Importantly, for a given k, there may exist multiple k-insulated sets of different sizes in the network. However, we only need to focus on the maximum k-insulated sets since the designer prefers more active agents and can directly recommend action profiles. Any $\lfloor \mathbb{E}_{\mu_{\mathbf{a}}} \rfloor$-insulated set with fewer nodes is dominated by the maximum set and will never be implemented. That is, in the following analysis, we only focus on the designer's mostly preferred value.

Given a non-negative integer k, let $v(k)$ denote the size of maximum k-insulated sets, that is, $v(k) = \max_{S_k \subseteq N} |S_k|$ where S_k is a k-insulated set. By Lemma 1, the designer's optimization problem is to manipulate the beliefs to maximize the expected size of the belief-based maximum insulated sets.

Lemma 2. *The designer's problem* (4) *is equivalent to*

$$\max_{\pi} \sum_{\omega, \mathbf{a}} \pi(\mathbf{a}|\omega) v(\lfloor \mathbb{E}_{\mu_{\mathbf{a}}} \rfloor)$$
$$\text{s.t.} \quad \mu_0(\omega) = \sum_{\omega', \mathbf{a}} \pi(\mathbf{a}|\omega') \mu_{\mathbf{a}}(\omega), \forall\, \omega. \tag{5}$$

In Lemma 2, $\mathbf{a}$ *can be viewed as a general signal rather than just a recommendation. The posterior mean when the agent receives signal* $\mathbf{a}$ *is given by* $\mathbb{E}_{\mu_{\mathbf{a}}}$, *and the corresponding equilibrium is captured by the maximum* $\lfloor \mathbb{E}_{\mu_{\mathbf{a}}} \rfloor$*-insulated set.*[7] *The designer aims to maximize the expected size of maximum* $\lfloor \mathbb{E}_{\mu_{\mathbf{a}}} \rfloor$*-insulated set by manipulating "Bayes plausible" posterior beliefs.*

Let V be the concave closure of cardinality function v:

$$V(\mathbb{E}_{\mu}) = \sup\{z | (\lfloor \mathbb{E}_{\mu} \rfloor, z) \in co(v)\},$$

where $co(v)$ denotes the convex hull of the graph of v. Similar to Kamenica and Gentzkow (2011), V characterizes the optimal outcome the designer can achieve.

Proposition 1. *The largest expected number of active agents the designer can achieve is* $V(\mathbb{E}_{\mu_0})$.

The payoff achievable by the designer is jointly determined by the prior distribution μ_0 over $\{\omega_L, \omega_H\}$ and network structure. In particular, the insulated sets of the network are crucial in shaping the designer's payoff. Specifically, the maximum k-insulated sets for each k directly determine the equilibrium behavior of agents under different posterior means. The optimal combination of "Bayes plausible" posterior means gives the highest payoff the designer can induce for any given prior.

Figure 2 shows the cardinality of maximum insulated set and the corresponding concave closure V for the network from the illustrative example. The cardinality function v is a step function: the k-insulated set is empty with zero cardinality when $k < 0$; when $k = 0$ or 1, the maximum k-insulated set contains 8 nodes ($N \backslash \{C, D, E\}$); the maximum 2-insulated set has 9 nodes ($N \backslash \{A, B\}$), while the maximum 3-insulated set equals the maximum 0-insulated set with 8 nodes.[8] Figure 2b describes the designer's optimal payoffs, for an adequate range of states, under different prior mean values. Notably, the designer's payoff peaks when $\mathbb{E}_{\mu_0} \in [2, 3]$.

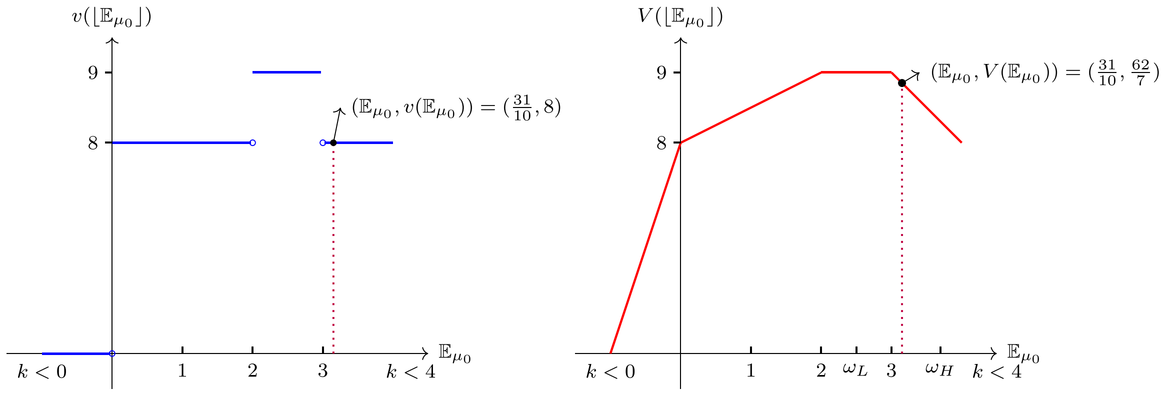
Remarks

1. *An empty network always induces the largest payoff for the designer, while a complete network always generates the lowest payoff, regardless of the prior distribution. With an empty network, every agent has the highest incentive to be active irrespective of others' choices since there are no substitution effects. In contrast, with a complete network, each agent's activity substitutes for others', naturally decreasing aggregate activity.*

2. *The designer's payoff is not monotone with network density.*[9] *Given a prior distribution of states* ω, *the two extreme structures—empty and complete networks—correspond to the highest and lowest payoffs. However, intermediate cases may not follow the intuition that sparser networks yield higher payoffs.*

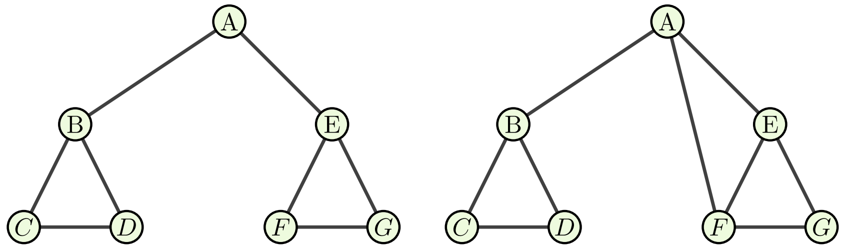
Consider two networks $\mathbf{G}$ and $\bar{\mathbf{G}}$ as shown in Figure 3, where $\bar{\mathbf{G}}$ is obtained by adding one link A-F to $\mathbf{G}$. Suppose $\omega_L, \omega_H \in [2, 3)$, so aggregate activity is determined by the cardinality of 2-insulated set. The 2-insulated set of $\mathbf{G}$ is $\{A, C, D, F, G\}$ with 5 nodes, while for $\bar{\mathbf{G}}$ it is $\{B, C, D, E, F, G\}$ with 6 nodes. Thus, the denser network $\bar{\mathbf{G}}$ induces a higher payoff than the sparser network $\mathbf{G}$, contradicting the payoff relation between the complete network (dense) and the empty network (sparse).

The key implication of Proposition 1 is that evaluating whether the designer benefits from persuasion and determining the



(a) Function $v(\cdot)$ (b) Function $V(\cdot)$

FIGURE 2 | The cardinalities of maximum insulated sets and designer's payoff.



(a) A sparse network $\mathbf{G}$ (b) A dense network $\bar{\mathbf{G}}$

FIGURE 3 | Dense network induces higher payoff.

value of an optimal signal requires the cardinalities of insulated sets under Bayes-plausible posterior beliefs. Kamenica and Gentzkow (2011) provide necessary and sufficient conditions on Bayes-plausible posteriors for beneficial persuasion in single-agent cases. With multiple agents, the characterization of beneficial persuasions must incorporate network properties due to substitution effects. In particular, whether the designer is able to benefit from persuasion depends on the relative size of insulated sets related to prior mean and the most preferred state. Let $d_{\max} = \max\{|N_i|_{i\in N}\}$ denote the highest degree in the network.

Corollary 1. *When $\mathbb{E}_{\mu_0}$ is not an integer, the designer strictly benefits from persuasion if and only if $v(\lfloor \mathbb{E}_{\mu_0} \rfloor) \neq \max\limits_{k\in[\lfloor\omega_L\rfloor,\lfloor\omega_H\rfloor]} \{v(\lfloor k \rfloor)\}$. Notably, beneficial persuasion is guaranteed when $\mathbb{E}_{\mu_0} < d_{\max} \leq \omega_H$, regardless of network structure.*

Corollary 1 implies that if the size of the prior-mean insulated set is not the designer's most preferred, there exists a Bayes-plausible posterior distribution that yields higher expected active agents. In particular, when $\mathbb{E}_{\mu_0} < d_{\max} \leq \omega_H$, the designer is guaranteed to benefit from persuasion without any network assumptions since $\lfloor \mathbb{E}_{\mu_0} \rfloor$-insulated set must be a proper subset of $\lfloor \omega_H \rfloor$-insulated set. This insight is similar to Kamenica and Gentzkow (2011), except the designer's most preferred state depends on network topology through the insulated set sizes. However, there is no monotonic relation between the k-insulated set size and the value of k. Whether the designer prefers high or low states is ambiguous, as shown in the leading example, where it depends on the network structure. Therefore, we must compare all the k-insulated sets for $k \in [\lfloor\omega_L\rfloor, \lfloor\omega_H\rfloor]$ against the prior-mean set.

Moreover, we impose a precondition that the prior mean $\mathbb{E}_{\mu_0}$ cannot be an integer. This is because the designer's payoff is the concave closure of a step function, with possible kink points at integer values in $[\omega_L, \omega_H]$.[10] If $\mathbb{E}_{\mu_0}$ is an integer, it could be a kink point, as illustrated in Figure 4 where $\omega_L = -1$, $\mathbb{E}_{\mu_0} = 0$ and $\omega_H = 2.5$. Moreover, the kink point $\mathbb{E}_{\mu_0}$ is on the concave closure of function $v(\cdot)$ with more than two steps. As a result, the designer can not strictly benefit from persuasion in this case. The single agent case in Kamenica and Gentzkow (2011) avoids this issue since the function $v(\cdot)$ is a two-step function, so no multiple kink points in the concave closure. Our multiagent setting introduces more than two possible outcomes that require the non-integer prior mean precondition for Corollary 1.

4.1 | Downwardplaying Mechanism

Definition 2. A mechanism $\pi : \{\omega_L, \omega_H\} \to \Delta\{s, s'\}$ is downwardplaying if

$$\pi(s|\omega_L) = 1;\ \pi(s|\omega_H) = q;$$
$$\pi(s'|\omega_L) = 0;\ \pi(s'|\omega_H) = 1 - q,$$

for some $q \in (0, 1)$.

A downwardplaying mechanism truthfully reveals the low state while sending randomized signals when the state is high. It can not be optimal with a single agent (binary choice), but may be optimal with multiple agents as analyzed above. One rationale for downwardplaying is that the low state results in a higher payoff for the designer due to substitution effects. Formally, as our illustration, with $\mathbb{E}_{\mu_0}$ not an integer, the optimal mechanism is downwardplaying if $v(\lfloor\omega_L\rfloor) > v(\lfloor\omega_H\rfloor)$ and function V only has one kink point in $[\lfloor\omega_L\rfloor, \lfloor\omega_H\rfloor]$. This rationale never applies in Candogan (2022)'s model with strategic complements, where higher state always increase active agents.

We provide some special network structures that avoid the downwardplaying rationale.

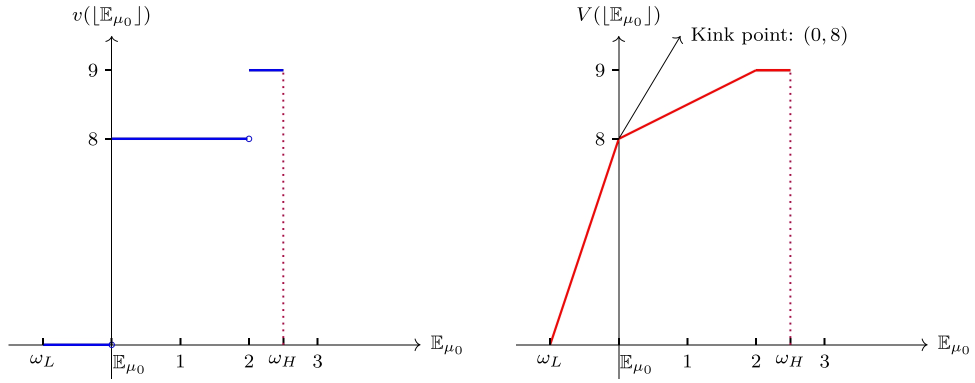
Definition 3. [Tree network] A network is a tree if it is connected and the number of nodes is one more than the number of links;

[Nested split network] A network is nested split if for any two nodes $i, j \in N$, either $N_i\backslash\{j\} \subseteq N_j\backslash\{i\}$ or the converse;

[Core-periphery network] A network G is a core-periphery network if there exists a set of core nodes $C \subseteq N$ and periphery nodes $P = N\backslash C$, such that: (1) for any $i, j \in C$, $g_{ij} = 1$; (2) for any $i, j \in P$, $g_{ij} = 0$.

Proposition 2. *Assume $\mathbb{E}_{\mu_0}$ is not an integer. When the network is a tree, nested split, or core-periphery, the designer strictly benefits from persuasion if and only if $v(\lfloor\mathbb{E}_{\mu_0}\rfloor) \neq v(\lfloor\omega_H\rfloor)$. Moreover, the optimal strategy is upwardplaying when $v(\lfloor\omega_L\rfloor) = v(\lfloor\mathbb{E}_{\mu_0}\rfloor) \neq v(\lfloor\omega_H\rfloor)$ and $v(\lfloor k\rfloor) \in \{v(\lfloor\omega_L\rfloor), v(\lfloor\omega_H\rfloor)\}, \forall\ k \in [\omega_L, \omega_H]$.*

Proposition 2 simplifies checking for beneficial persuasion compared to Corollary 1 for tree, nested split, and core-periphery networks. Rather than computing all k-insulated sets, only the prior mean and highest state sets are needed. This stark reduction is enabled by the monotonicity of insulated set size in the identified



(a) Function $v(\cdot)$ (b) Function $V(\cdot)$

FIGURE 4 | When $\mathbb{E}_{\mu_0}$ is an integer, it is possible that the designer can not benefit from persuasion even though $v(\lfloor \mathbb{E}_{\mu_0} \rfloor) \neq \max_{k \in [\lfloor \omega_L \rfloor, \lfloor \omega_H \rfloor]} \{v(\lfloor k \rfloor)\}$.

structures. Specifically, the number of nodes in k-insulated sets increases with k for these networks. This restores the single agent intuition that higher states are unambiguously better, so downwardplaying is never optimal regardless of the prior.

Moreover, even within the identified network classes, the designer's payoff is not necessarily monotone in density. In a star network, which is a tree, nested split, and core-periphery, more leaf agents increase the payoff. However, comparing the empty and complete networks, which are both nested splits and core-peripheries, shows an inverse relation—the denser complete network reduces the payoff. So there is not a universal density-payoff relation, even for these structured networks.

4.2 | Mixed Mechanism

Besides the reversal of the designer's preference due to strategic substitution, the leading example departs from Kamenica and Gentzkow (2011) in another key aspect: The designer's payoff function has multiple steps in the absence of persuasion, corresponding to the values of ω_L and ω_H. This distinction results in another uncommon optimal mechanism in the literature that never truthfully reveals any state, regardless of whether it is high or low, which is called a mixed mechanism.

Definition 4. A mechanism $\pi : \{\omega_L, \omega_H\} \to \Delta\{s, s'\}$ is mixed if

$$\pi(s|\omega_L) = q_1; \qquad \pi(s|\omega_H) = q_2;$$
$$\pi(s'|\omega_L) = 1 - q_1; \qquad \pi(s'|\omega_H) = 1 - q_2,$$

for some $q_1, q_2 \in (0, 1)$.

In this part, we investigate when a mixed mechanism is optimal for the sender. We know that the function $V(\cdot)$ is a piecewise linear function composed of several linear segments since $V(\cdot)$ is the concave closure of a step function $v(\cdot)$. The points where the slope of the function $V(\cdot)$ changes are called kink points. Specifically, the function $V(\cdot)$ is divided into t linear segments by the points $x_0, \ldots, x_{t+1}$ where $x_0 = (\omega_L, v(\omega_L))$ and $x_{t+1} = (\omega_H, v(\omega_H))$. The slope of $V(\cdot)$ on the left of x_i differs from the slope on the right for $i = 1, \ldots, t$. We call $x_0, x_1, \ldots, x_{t+1}$ as kink points, and use $((z_L, v(z_L)), (z_R, v(z_R)))$ to represent the segment on which $(\mathbb{E}_{\mu_0}, V(\mathbb{E}_{\mu_0}))$ lies. Suppose $\mathbb{E}_{\mu_0}, \omega_L, \omega_H$ are not integers, we have the following proposition.

Proposition 3. *When* $v(\lfloor \mathbb{E}_{\mu_0} \rfloor) \neq \max_{k \in [\lfloor \omega_L \rfloor, \lfloor \omega_H \rfloor]} \{v(\lfloor k \rfloor)\}$, *the optimal persuasion is a mixed mechanism if and only if* $\omega_L \neq z_L$ *and* $\omega_H \neq z_H$.

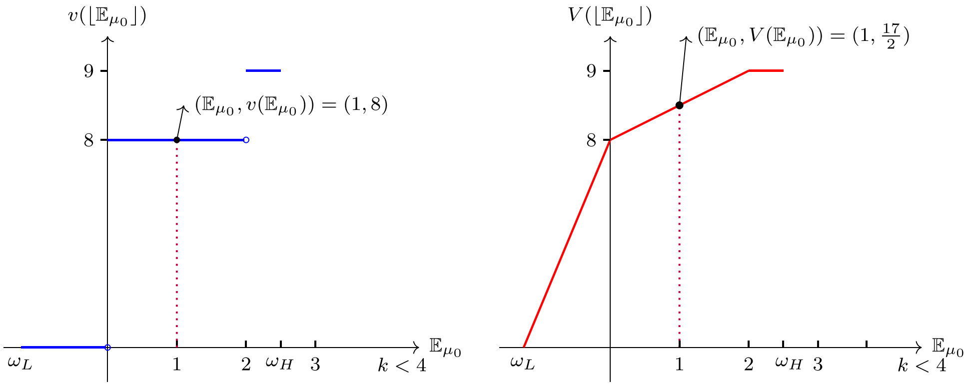
The optimality of mixed mechanism is due to multiple steps of function $v(\cdot)$, and we illustrate the intuition as follows. Suppose $\omega_L = -2$, $\omega_H = \frac{5}{2}$ and $\mathbb{E}_{\mu_0} = 1$, the case of which is illustrated by Figure 5. Function $V(\cdot)$ is a piecewise linear function composed of 3 linear segments, that is, $x_0 = (-2, 0)$, $x_1 = (0, 8)$ and $x_2 = (2, 9)$. Moreover, $(\mathbb{E}_{\mu_0}, V(\mathbb{E}_{\mu_0}))$ lies in the middle of one segment in which $z_L = 0 \neq \omega_L$ and $z_R = 2 \neq \omega_H$. The optimal mechanism is

$$\pi(s_{AB}|\omega_H) = \tfrac{2}{3}; \qquad \pi(s_{AB}|\omega_L) = \tfrac{1}{6};$$
$$\pi(s_{CDE}|\omega_H) = \tfrac{1}{3}; \qquad \pi(s_{CDE}|\omega_L) = \tfrac{5}{6}.$$

Intuitively, the mixed mechanism captures two opposing forces caused by the payoff function with more than two steps. On the one hand, the designer wants an upward state as $v(k) = 0$ when $k < 0$. On the other hand, the designer aims to manipulate the posterior expected state to 2, a kink point, when ω_H occurs. The mixed mechanism when the state is high ensures that agents cannot distinguish the real state upon receiving signal s_{CDE} and incentivizes the agents to obey the recommendation when the state is low. Generally, the mixed mechanism may be optimal when the designer aims to manipulate intermediate posterior means at kink points of the payoff function, which is formally characterized by the condition of $\omega_L \neq z_L$ and $\omega_H \neq z_H$. In particular, in this example, when $\omega_L \in (-16, 0)$ and $\omega_H \in (2, 3)$, the following mixed mechanism is optimal:

$$\pi(s_{AB}|\omega_H) = \frac{(1-\alpha)(1-\beta_2)}{(1-\alpha)(1-\beta_2) + \alpha(1-\beta_1)}; \qquad \pi(s_{AB}|\omega_L) = \frac{(1-\alpha)\beta_2}{(1-\alpha)\beta_2 + \alpha\beta_1};$$
$$\pi(s_{CDE}|\omega_H) = \frac{\alpha(1-\beta_1)}{(1-\alpha)(1-\beta_2) + \alpha(1-\beta_1)}; \qquad \pi(s_{CDE}|\omega_L) = \frac{\alpha\beta_1}{(1-\alpha)\beta_2 + \alpha\beta_1},$$

where $\alpha = 1 - \frac{\mathbb{E}_{\mu_0}}{2}$, $\beta_1 = \frac{\omega_H}{\omega_H - \omega_L}$, $\beta_2 = \frac{\omega_H - 2}{\omega_H - \omega_L}$. In this mixed mechanism, the designer manipulates two posterior means $z_L = 0$ and $z_H = 2$.



(a) Function $v(\cdot)$ (b) Function $V(\cdot)$

FIGURE 5 | The cardinalities of maximum insulated sets and designer's payoff.

5 | Discussion and Conclusion

In this paper, we study Bayesian persuasion with multiple strategic agents on a network exhibiting strategic substitutes. We characterize the optimal signaling mechanism for a designer seeking to maximize total agent activity. Proposition 1 expresses the maximum achievable payoff as a concave closure over network-insulated set cardinalities under different posteriors. This reduces optimal signaling to maximize insulated set sizes through Bayes-plausible beliefs, highlighting the pivotal network role in linking states to equilibria. An immediate corollary shows the designer benefits from persuasion if the prior-mean insulated set size is suboptimal. Notably, for tree, nested split, and core-periphery networks, persuasion strictly improves payoffs if the prior mean set is smaller than the highest state set. The structured topologies ensure monotonicity, eliminating complicated non-monotonic cases.

We conclude this paper by discussing some potential extensions of our model. First, we can relax the assumption of an undirected and unweighted network to consider undirected and weighted networks. The existence of k-insulated sets in weighted networks has been proved by Fu et al. (2022). Therefore, our main results can be extended to the case of weighted networks in a straightforward manner.

Second, our benchmark model assumes that an agent desires to contribute only when the total contribution of their neighbors is smaller than a threshold (cf. Equation 1). However, Andreoni (1998) argues that in many situations, there are fixed costs involved in producing public goods, and thus, there is a minimum threshold that contributions must meet before the benefits of the public goods can be consumed. Following this idea, we can extend our benchmark model by incorporating a lower bound that the agent is willing to contribute. Specifically, each agent's best response would be:

$$x_i^*(\mathbf{a}_{-i}) = \begin{cases} 1 & \text{if } \underline{\omega} \le \sum_j g_{ij} a_j \le \bar{\omega} \\ 0 & \text{otherwise} \end{cases}.$$

In this setting, each agent i is willing to contribute if an adequate number of neighbors (between $\underline{\omega}$ and $\bar{\omega}$) did. When the total number of contributors is less than $\underline{\omega}$, the agent is not willing to contribute because the aggregate contribution does not meet the threshold for the public goods to be consumed. When the total number of contributors is more than $\bar{\omega}$, as in our benchmark model, the agent is not willing to contribute because the marginal cost exceeds the marginal benefits. To characterize the optimal persuasion mechanisms in this extended model, we need to generalize the concept of ω-insulated sets to $(\bar{\omega}, \underline{\omega})$-insulated sets. In a $(\bar{\omega}, \underline{\omega})$-insulated set, each node in the set is connected to a number of other nodes in the set that falls within the range $[\underline{\omega}, \bar{\omega}]$. Nodes outside the set are connected to either fewer than $\underline{\omega}$ nodes or more than $\bar{\omega}$ nodes within the set. To determine whether the mediator can benefit from persuasion, we need to identify all $(\bar{\omega}, \underline{\omega})$-insulated sets in the network. However, the existence of nonempty $(\bar{\omega}, \underline{\omega})$-insulated sets is not guaranteed for all network structures. For example, in Figure 1, there is no (2, 2)-insulated set, where each node in the set would need to connect with exactly two other nodes in the set, and each node outside the set would need to connect with either at most 1 or at least 3 nodes in the set. Moreover, there is currently no known algorithm for efficiently identifying $(\bar{\omega}, \underline{\omega})$-insulated sets in general networks. Cheah and Corneil (1990) imply that identifying such sets is NP-hard, even in the special case where $\bar{\omega} = \underline{\omega}$. Given these challenges, we leave such a generalization for future study.

Acknowledgments

We thank Zhigang Cao, Wentao Fu, Junze Sun, and Kun Zhang for their constructive comments. We acknowledge financial supports from National Natural Science Foundation of China (No. 72173049, 72203184) and MOE project of Humanities and Social Sciences (No. 22YJC79111). All remaining errors are our own.

Data Availability Statement

Data sharing is not applicable to this article as no data sets were generated or analyzed during the current study.

Endnotes

[1] As Candogan (2022) states in the discussion: "Although it provides a comprehensive study of optimal public signaling mechanisms in settings where agents' actions exhibit strategic complementarities, it does not shed light on the structure of optimal public signaling mechanisms when the actions are strategic substitutes. We believe that this remains an interesting direction..."

[2]The network structure comes from Jagota, Narasimhan, and Šoltés (2001) which is used to demonstrate the non-monotonicity of k-insulated set.

[3]The model can be generalized to a more expansive framework in which the network is weighted (see Section 5 for specifics).

[4]The agent is indifferent between 0 and 1 when $\sum_j g_{ij} a_j = \omega$ in this case. If it is the tie case, we assume that agents choose the sender preferred equilibrium as in Kamenica and Gentzkow (2011).

[5]Here, we use $\mathrm{int}(X)$ to denote the interior of the set X and use $\Delta(X)$ to denote all distributions on X.

[6]Galperti and Perego (2023) study Bayesian communication games with information networks, by which agents may have additional information about their neighbors' signals. The public signal corresponds to complete information network which has been shown to satisfy the revelation principle.

[7]When $\lfloor \mathbb{E}_{\mu_{\mathbf{a}}} \rfloor$ is an integer, the corresponding equilibrium is captured by the maximum $\lfloor \mathbb{E}_{\mu_{\mathbf{a}}} \rfloor$-insulated set or the maximum $(\lfloor \mathbb{E}_{\mu_{\mathbf{a}}} \rfloor - 1)$-insulated set.

[8]For illustration, we assume $\omega_H < 4$ and omit cases with $k \geq 4$.

[9]This contrasts with Candogan (2022), who shows aggregate activity always increases with density under strategic complements.

[10]The points where the slope of the function $V(\cdot)$ changes are called kink points and we formally define it after Definition 4.

References

Adachi, T. 2005. "Third-Degree Price Discrimination, Consumption Externalities and Social Welfare." *Economica* 72, no. 285: 171–178. https://doi.org/10.1111/ecca.2005.72.issue-285.

Allouch, N. 2015. "On the Private Provision of Public Goods on Networks." *Journal of Economic Theory* 157: 527–552. https://doi.org/10.1016/j.jet.2015.01.007.

Allouch, N., and M. King. 2019. "Constrained Public Goods in Networks." *Journal of Public Economic Theory* 21, no. 5: 895–902. https://doi.org/10.1111/jpet.v21.5.

Alonso, R., and O. Câmara. 2016. "Persuading Voters." *American Economic Review* 106, no. 11: 3590–3605. https://doi.org/10.1257/aer.20140737.

Andreoni, J. 1998. "Toward a Theory of Charitable Fund-Raising." *Journal of Political Economy* 106, no. 6: 1186–1213. https://doi.org/10.1086/250044.

Arieli, I., and Y. Babichenko. 2019. "Private Bayesian Persuasion." *Journal of Economic Theory* 182, no. 3: 185–217. https://doi.org/10.1016/j.jet.2019.04.008.

Bergemann, D., and S. Morris. 2016. "Information Design, Bayesian Persuasion, and Bayes Correlated Equilibrium." *American Economic Review* 106, no. 5: 586–591. https://doi.org/10.1257/aer.p20161046.

Bergemann, D., and S. Morris. 2019. "Information Design: A Unified Perspective." *Journal of Economic Literature* 57, no. 1: 44–95. https://doi.org/10.1257/jel.20181489.

Bhattacharya, S., J. Kulkarni, K. Munagala, and X. Xu. 2011. "On Allocations With Negative Externalities." In *Internet and Network Economics: 7th International Workshop, WINE 2011, Singapore, December 11–14, 2011. Proceedings 7*, 25–36. Singapore: Springer.

Bramoullé, Y., and R. Kranton. 2007. "Public Goods in Networks." *Journal of Economic Theory* 135, no. 1: 478–494. https://doi.org/10.1016/j.jet.2006.06.006.

Bramoullé, Y., R. Kranton, and M. D'Amours. 2014. "Strategic Interaction and Networks." *American Economic Review* 104, no. 3: 898–930. https://doi.org/10.1257/aer.104.3.898.

Candogan, O. 2022. "Persuasion in Networks: Public Signals and Cores." *Operations Research* 70, no. 4: 2264–2298. https://doi.org/10.1287/opre.2021.2246.

Cao, Z., X. Chen, X. Hu, and C. Wang. 2017. "Approximation Algorithms for Pricing With Negative Network Externalities." *Journal of Combinatorial Optimization* 33, no. 2: 681–712. https://doi.org/10.1007/s10878-015-9988-1.

Cheah, F., and D. G. Corneil. 1990. "The Complexity of Regular Subgraph Recognition." *Discrete Applied Mathematics* 27, no. 1–2: 59–68. https://doi.org/10.1016/0166-218X(90)90129-Z.

Elliott, M., and B. Golub. 2019. "A Network Approach to Public Goods." *Journal of Political Economy* 127, no. 2: 730–776. https://doi.org/10.1086/701032.

Fu, W., D. Hua, X. Qian, and Y. Sun. 2022. "Constrained Public Goods in Weighted Networks With Heterogeneous Agents." *Economics Letters* 213, no. 1: 110332. https://doi.org/10.1016/j.econlet.2022.110332.

Galperti, S., and J. Perego. 2023. Information Systems. Technical report, Working Paper.

Gerke, S., G. Gutin, S.-H. Hwang, and P. R. Neary. 2024. "Public Goods in Networks With Constraints on Sharing." *Journal of Economic Theory* 219, no. 5: 105844. https://doi.org/10.1016/j.jet.2024.105844.

Grigorescu, E. 2004. "The Insulation Sequence of a Graph." *Discrete Applied Mathematics* 134, no. 1–3: 77–90. https://doi.org/10.1016/S0166-218X(03)00300-7.

Hennigs, R. 2021. "Conflict Prevention by Bayesian Persuasion." *Journal of Public Economic Theory* 23, no. 4: 710–731. https://doi.org/10.1111/jpet.v23.4.

Jagota, A., G. Narasimhan, and L. Šoltés. 2001. "A Generalization of Maximal Independent Sets." *Discrete Applied Mathematics* 109, no. 3: 223–235. https://doi.org/10.1016/S0166-218X(00)00215-8.

Kamenica, E., and M. Gentzkow. 2011. "Bayesian Persuasion." *American Economic Review* 101, no. 6: 2590–2615. https://doi.org/10.1257/aer.101.6.2590.

Karni, E., and D. Schmeidler. 1990. "Fixed Preferences and Changing Tastes." *American Economic Review* 80, no. 2: 262–267.

Mathevet, L., J. Perego, and I. Taneva. 2020. "On Information Design in Games." *Journal of Political Economy* 128, no. 4: 1370–1404. https://doi.org/10.1086/705332.

Schnakenberg, K. E. 2015. "Expert Advice to a Voting Body." *Journal of Economic Theory* 160, no. 4: 102–113. https://doi.org/10.1016/j.jet.2015.08.005.

Simmel, G. 1957. "Fashion." *American Journal of Sociology* 62, no. 6: 541–558. https://doi.org/10.1086/222102.

Yoganarasimhan, H. 2012. "Cloak or Flaunt? The Fashion Dilemma." *Marketing Science* 31, no. 1: 74–95. https://doi.org/10.1287/mksc.1110.0689.

Appendix A

Appendix

Proof of Lemma 1. Recall that π is incentive-compatible if for each $i \in N$ and $\mathbf{a} \in A$ such that $\sum_\omega \mu_0(\omega)\pi(\mathbf{a}|\omega) > 0$, we have

$$a_i = \begin{cases} 1 & \text{when } \sum_j g_{ij} a_j \leq \mathbb{E}_{\mu_{\mathbf{a}}} \\ 0 & \text{when } \sum_j g_{ij} a_j \leq \mathbb{E}_{\mu_{\mathbf{a}}} \end{cases}. \tag{A1}$$

Let $S = \{i : a_i = 1\}$ denote the set of active agents. We first discuss the case that $\lfloor \mathbb{E}_{\mu_{\mathbf{a}}} \rfloor$ is not an integer. Suppose, by way of contradiction, S is not a $\lfloor \mathbb{E}_{\mu_{\mathbf{a}}} \rfloor$-insulated set. Then either (i) there exists an agent $i \in S$ with $\sum_j g_{ij} a_j > \lfloor \mathbb{E}_{\mu_{\mathbf{a}}} \rfloor$, or (ii) there exists an agent $i \in N \backslash S$ with $\sum_j g_{ij} a_j \leq \lfloor \mathbb{E}_{\mu_{\mathbf{a}}} \rfloor$. If (i) happens, agent i's expected utility by taking action 1 is $u_i = \mathbb{E}_{\mu_{\mathbf{a}}} - \sum_j g_{ij} a_j < 0$. Therefore, i can increase its utility by deviating its action from 1 to 0. If (ii) happens, agent i's expected utility by taking action 1 is $u_i = \mathbb{E}_{\mu_{\mathbf{a}}} - \sum_j g_{ij} a_j > 0$ and thus i can increase its utility by deviating its action from 0 to 1.

The remaining case is that $\lfloor \mathbb{E}_{\mu_{\mathbf{a}}} \rfloor$ is an integer. In this case, S is either a $\lfloor \mathbb{E}_{\mu_{\mathbf{a}}} \rfloor$-insulated set or a $(\lfloor \mathbb{E}_{\mu_{\mathbf{a}}} \rfloor - 1)$-insulated set. The analysis is similar to the above case and we omit it here. □

Proof of Lemma 2. Given any posterior belief $\mu_{\mathbf{a}}$, we know that receivers only care about the posterior mean $\mathbb{E}_{\mu_{\mathbf{a}}}$. By lemma 1, we know that there are at most $v(\lfloor \mathbb{E}_{\mu_{\mathbf{a}}} \rfloor)$ active agents, which is the designer's payoff under $\mu_{\mathbf{a}}$. Therefore, the designer's objective is to maximize its expected payoff $\sum_{\omega,a} \pi(a|\omega) v(\lfloor \mathbb{E}_{\mu_a} \rfloor)$ by designing an optimal π. The information structure π should satisfy Bayes plausibility, that is, for any $\omega \in \Omega$, we should have $\mu_0(\omega) = \sum_{\omega',a} \pi(a|\omega') \mu_a(\omega)$. Formally, we can obtain the designer's optimization problem as follows.

$$\begin{aligned} &\max_{\pi} \sum_{\omega,a} \pi(a|\omega) v(\lfloor \mathbb{E}_{\mu_a} \rfloor) \\ &\text{s.t.} \quad \mu_0(\omega) = \sum_{\omega',a} \pi(a|\omega') \mu_a(\omega), \forall\, \omega. \end{aligned} \tag{A2}$$

□

Proof of Proposition 1. Let $\tau \in \Delta(\Delta(A))$ denote a distribution of the posterior belief and let $Supp(\tau)$ denote the support of τ. Using the methodology of Kamenica and Gentzkow (2011), we can convert the optimization problem (A2) to the following equivalent optimization problem.

$$\begin{aligned} &\max_{\tau} \quad \mathbb{E}_\tau v(\lfloor \mathbb{E}_\mu \rfloor) \\ &\text{s.t.} \quad \mu_0(\omega) = \sum_{Supp(\tau)} \tau(\mu) \mu(\omega), \forall\, \omega. \end{aligned} \tag{A3}$$

In problem (A3), the designer's objective is to maximize its expected payoff by mixing some posterior beliefs that satisfy Bayes plausibility. The plausibility condition $\mu_0(\omega) = \sum_{Supp(\tau)} \tau(\mu)\mu(\omega), \forall\, \omega$ is equal to $\mathbb{E}_{\mu_0} = \sum_{Supp(\tau)} \tau(\mathbb{E}_\mu) \mathbb{E}_\mu$ since we have only two states. Therefore, we can obtain the designer's problem as follows.

$$\max_{\tau} \quad \mathbb{E}_\tau v(\lfloor \mathbb{E}_\mu \rfloor). \tag{A4}$$

$$\text{s.t.} \quad \mathbb{E}_{\mu_0} = \sum_{Supp(\tau)} \tau(\mathbb{E}_\mu) \mathbb{E}_\mu. \tag{A5}$$

Applying the concavification method here, we know that the largest expected payoff the designer can achieve is $V(\mathbb{E}_{\mu_0})$. □

Proof of Proposition 2. We prove Proposition 2 through two steps. In first step, we will prove that if $v(\mathbb{E}_\mu)$ is an increasing step function about $\mathbb{E}_\mu$, then the designer benefits from persuasion if and only if $v(\lfloor \mathbb{E}_{\mu_0} \rfloor) \neq v(\lfloor \omega_H \rfloor)$. In second step, we will prove that if the network is a tree, nested split graph, or core-periphery, then $v(k)$ increases with integer k.

Step 1. In this step, we shall demonstrate that if $v(\mathbb{E}_\mu)$ is an increasing step function about $\mathbb{E}_\mu$, then the designer benefits from persuasion if and only if $v(\lfloor \mathbb{E}_{\mu_0} \rfloor) \neq v(\lfloor \omega_H \rfloor)$.

First, we give the formal definition of increasing step function.

Definition 5 Step Function. We say a function $f(x)$ is a step function in $[a, b]$ if there exists $t_1 = a < t_2 < \cdots < t_{Z-1} < t_Z = b$ such that for any $1 \leq z \leq Z - 1$, we have (i) $f(x_1) = f(x_2)$ for any x_1, x_2 with $x_1, x_2 \in [t_z, t_{z+1}]$, and (ii) $f(x_1) \neq f(x_2)$ for any $x_1 \in [t_z, t_{z+1}]$ and $x_2 \in [t_{z+1}, t_{z+2}]$ if $Z \geq 3$. We say a step function is an increasing step function if $f(x_1) \leq f(x_2)$ for any $x_1 < x_2$.

Next, we will define the steps of the function v as follows. Let $r_0 = -\infty$ and $r_{Z+1} = \infty$. Define $\mathcal{K} = \{r | v(r) \text{ is distinct from } v(r-1)\} = \{r_1, \ldots, r_z, \ldots, r_Z\}$ where r is integer. From the definition of $\mathcal{K}$, we know that $r_1 = 0$ and $r_Z = d_{\max}(\mathbf{g})$. We say $v(\cdot)$ is an increasing step function if $v(r_k) < v(r_{k+1})$ for $1 \leq z \leq Z$.

To prove sufficiency, we need to prove that if $v(\lfloor \mathbb{E}_{\mu_0} \rfloor) \neq v(\lfloor \omega_H \rfloor)$, then the designer can mix two posterior means satisfying Bayes plausibility to strictly increase its payoff. Let $E_{\mu_1} = E_{\mu_0} - \epsilon$ where $\epsilon > 0$ and ϵ is small enough, and let $E_{\mu_2} = \omega_H$. It is obvious that $E_{\mu_1} < E_{\mu_0} < E_{\mu_2}$, and thus $v(\lfloor \mathbb{E}_{\mu_1} \rfloor) = v(\lfloor \mathbb{E}_{\mu_0} \rfloor) < v(\lfloor \mathbb{E}_{\mu_2} \rfloor) = v(\lfloor \omega_H \rfloor)$ since $v(\mathbb{E}_{\mu_{\mathbf{a}}})$ is an increasing step function. Therefore, there exists a $\alpha \in (0, 1)$ such that $\mathbb{E}_{\mu_0} = \alpha E_{\mu_1} + (1 - \alpha) E_{\mu_2}$ and $V(\mathbb{E}_{\mu_0}) \geq \alpha v(\lfloor \mathbb{E}_{\mu_1} \rfloor) + (1 - \alpha) v(\lfloor \mathbb{E}_{\mu_2} \rfloor) > v(\lfloor \mathbb{E}_{\mu_0} \rfloor)$, that is, the designer strictly benefits from persuasion. To prove necessity, we only need to prove if $v(\lfloor \mathbb{E}_{\mu_0} \rfloor) = v(\lfloor \omega_H \rfloor)$, the designer can't benefit from persuasion. Since $v(\mathbb{E}_{\mu_{\mathbf{a}}})$ is an increasing step function, $v(\lfloor \omega_H \rfloor)$ is the largest payoff the designer can obtain, which directly means that the designer can't benefit from persuasion.

Step 2. We will prove that if the network is a (1) tree, (2) nested split, or (3) core-periphery, then $v(k)$ (weakly) increases with integer k.

(1) is due to Theorem 5.1 of Grigorescu (2004).

We prove (2) by way of contradiction. Note that given k, if the maximum k-insulated set includes agent j then it must contain any agent i with $d_i(\mathbf{g}) < d_j(\mathbf{g})$ since $N_i(\mathbf{g}) \subseteq N_j(\mathbf{g}) \cup \{j\}$ due to the definition of nested split graph. Let S_k denote a maximum k-insulated set where $0 \leq k < d_{\max}(\mathbf{g})$ and let S_{k+1} denote a maximum $(k+1)$-insulated set where $1 \leq k + 1 \leq d_{\max}(\mathbf{g})$. Suppose, by way of contradiction, $|S_k| > |S_{k+1}|$. Let $\hat{z} \in \arg\max_{z \in S_k} d_z(\mathbf{g})$ and $\bar{z} \in \arg\max_{z \in S_{k+1}} d_z(\mathbf{g})$. If $d_{\bar{z}}(\mathbf{g}) > d_{\hat{z}}(\mathbf{g})$, it is obvious that $|S_{k+1}| > |S_k|$. If $d_{\bar{z}}(\mathbf{g}) < d_{\hat{z}}(\mathbf{g})$, then there exists an agent $z \in S_k$ with $d_z(\mathbf{g}) = d_{\hat{z}}(\mathbf{g})$ such that $|N_z(\mathbf{g}) \cap S_{k+1}| \leq k$, which contradicts that S_{k+1} is a k-insulated set. Now, we consider the case $d_{\bar{z}}(\mathbf{g}) = d_{\hat{z}}(\mathbf{g})$. Note that each agent i with $d_i(\mathbf{g}) < d_{\hat{z}}(\mathbf{g})$ belongs to both S_k and S_{k+1}. Therefore, we only need to consider agents whose degree is equal to $d_{\hat{z}}(\mathbf{g})$. Let $M_1 \subseteq S_k$ denote the set in which each agent's degree is equal to $d_{\hat{z}}(\mathbf{g})$ and let $M_2 \subseteq S_{k+1}$ denote the set in which each agent's degree is equal to $d_{\hat{z}}(\mathbf{g})$. We need to show $|M_2| \geq |M_1|$. Suppose by way of contradiction, $|M_2| < |M_1|$. It follows from the definition of nested split graph that for any $i, j \in M_1 \cup M_2$, $N_i(\mathbf{g}) \setminus \{j\} = N_j(\mathbf{g}) \setminus \{i\}$. Therefore, $\mathbf{g}(M_1 \cup M_2)$ is either a complete subnetwork or an empty subnetwork. If $\mathbf{g}(M_1 \cup M_2)$ is an empty subnetwork, then each $i \in M_2$ satisfies that $|N_i(\mathbf{g}) \cap S_{k+1}| \leq k$ since $S_k \setminus M_1 = S_{k+1} \setminus M_2$. Therefore, S_{k+1} is not a maximum $(k+1)$-insulated set since there exists an agent $\xi \in M_1 \setminus M_2$ such that $|N_\xi(\mathbf{g}) \cap S_{k+1}| \leq k+1$. The same logic applies to the case that $\mathbf{g}(M_1 \cup M_2)$ is a complete subnetwork.

Finally, we prove (3). In a core-periphery network, agents can be partitioned into two groups $\hat{N}_1(\mathbf{g})$ and $\hat{N}_2(\mathbf{g})$ such that $N_i(\mathbf{g}) \subseteq \hat{N}_2(\mathbf{g})$ for each $i \in \hat{N}_1(\mathbf{g})$, and $N_j(\mathbf{g}) \subseteq \hat{N}_2(\mathbf{g}) \setminus \{j\}$ for each $j \in \hat{N}_2(\mathbf{g})$. We prove (3) by proving the following stronger lemma.

Lemma 3. *If* $\mathbf{g}$ *is a core-periphery network, then* $v(k+1) \in \{v(k), v(k)+1\}$ *for any* $0 \leq k \leq d_{\max}(\mathbf{g}) - 1$.

Proof of Lemma AL. For convenience, we construct a partition of $\hat{N}_2(\mathbf{g})$, denoted by $\mathcal{K}_0, \ldots, \mathcal{K}_{|\hat{N}_1(\mathbf{g})|}$, based on the number of friends in $\hat{N}_1(\mathbf{g})$. That is, for any $i \in \hat{N}_2(\mathbf{g})$, i belongs to $\mathcal{K}_m$ if $|N_i(\mathbf{g}) \cap \hat{N}_1(\mathbf{g})| = m$. If for some $0 \leq m \leq |\hat{N}_1(\mathbf{g})|$, nobody in $\hat{N}_2(\mathbf{g})$ has m friends who belong to $\hat{N}_1(\mathbf{g})$, then $\mathcal{K}_m = \emptyset$.

Finally, we prove that for each $1 \leq k \leq d_{\max}(\mathbf{g}) - 1$, we have $v(k+1) \in \{v(k), v(k)+1\}$. Let S_k denote a k-insulated set which contains $v(k)$ agents. We first prove a useful observation. □

Observation 1. For any k-insulated set S_k, there exists a k-insulated set $\hat{S}_k$ with $\hat{N}_1(\mathbf{g}) \subseteq \hat{S}_k$ and $|\hat{S}_k| \geq |S_k|$.

Given any one k-insulated set S_k, We design a novel algorithm to search a k-insulated set $\hat{S}_k$ with $\hat{N}_1(\mathbf{g}) \subseteq \hat{S}_k$ and $|\hat{S}_k| \geq |S_k|$.

Definition 6 Algorithm BS_k, T. Given any k-insulated set S_k, let $T = \hat{N}_1(\mathbf{g}) \setminus S_k$. Define the algorithm $B(S_k, T)$ as follows.

1. If $T = \emptyset$, then the algorithm stops. Otherwise then remove an arbitrary agent from T and put it into S_k.
2. If $\sum_{j \in S_k} g_{ij} \leq k$ for each $i \in S_K$, and $\sum_{j \in S_k} g_{ij} \geq k+1$ for each $i \notin S_k$, then return to Step 1.
3. If there exists an agent $i \in S_k$ such that $\sum_{j \in S_k} g_{ij} \geq k+1$, then select a $\bar{i} \in S_k$ such that $\bar{i} \in \arg\max_{i \in S_k} \sum_{j \in S_k} g_{ij}$. There exists such a $\bar{i} \in$ who is in $\hat{N}_2(\mathbf{g})$. Remove $\bar{i}$ from the current S_k.
4. If there exists an agent $i \in \hat{N}_2(\mathbf{g}) \setminus S_k$ such that $\sum_{j \in S_k} g_{ij} \leq k$, then select a $\underline{i} \in \hat{N}_2(\mathbf{g}) \setminus S_k$ such that $\underline{i} \in \arg\min_{i \in \hat{N}_2(\mathbf{g}) \setminus S_k} \sum_{j \in S_k} g_{ij}$. We know that $\sum_{j \in S_k} g_{ij} \leq k < k+1 \leq \sum_{j \in S_k} g_{\bar{i}j}$. Add $\underline{i}$ into the current S_k.
5. Return to step 2.

From the algorithm $B(S_k, T)$, we know that at each step t, the size of S_k is at least as large as the size of S_k at step $t-1$ and no agent in $S_k \cap \hat{N}_1(\mathbf{g})$ is removed from S_k. Therefore, we find a $\hat{S}_k$ with $\hat{N}_1(\mathbf{g}) \subseteq \hat{S}_k$ and $|\hat{S}_k| \geq |S_k|$ if we prove the algorithm $B(S_k, T)$ terminates in finite iterations for any initial k-insulated set S_k. We prove this by using a novel potential function(also called energy function) proposed by Jagota, Narasimhan, and Šoltés (2001).

For any k-insulated set S_k, define the potential function as follows.

$$E(S_k) = Z(\mathbf{g}(S_k)) - (k + 1/2)|S_k|,$$

where $Z(\mathbf{g}(S_k))$ is the number of links in the subgraph induced by S_k. We show that whenever $B(S_k, T)$ enter into step 2, it will return to step 1 in finite steps. Note that $E(S_k \setminus \{\bar{i}\}) - E(S_k) \leq E(S_k) - (k+1) + (k+1/2) = E(S_k) - 1/2$ and $E(S_k \cup \{\underline{i}\}) - E(S_k) \leq E(S_k) + k - (k+1/2) = E(S_k) - 1/2$. This implies that after $B(S_k, T)$ enters into step 3, either $B(S_k, T)$ enters into step 4 directly or the potential function will decrease at least 1/2 and then enter into step 4. Similarly, after $B(S_k, T)$ enters into step 4, $B(S_k, T)$ will enter into step 5 directly, or the potential function will decrease at least 1/2 and then $B(S_k, T)$ enters into step 4 after $B(S_k, T)$ enters into step 4. Therefore, each time $B(S_k, T)$ enters into step 2 and returns back after going through steps 3, 4, and 5, the potential function decreases at least 1/2. Since $E(S_k) = Z(\mathbf{g}(S_k)) - (k+1/2)|S_k|$ is bounded from below, $B(S_k, T)$ must return to Step 1 after finite steps whenever it enters into step 2. We also know that at each round of Step 1, the size of T will decrease at least 1. Since T is a finite set, the algorithm $B(S_k, T)$ will terminate in finite steps.

Therefore, without loss of generality, we can only focus on the k-insulated sets, which contain all agents belonging to $\hat{N}_1(\mathbf{g})$. Let $\hat{S}_k$ be a maximum k-insulated set that contains all agents belonging to $\hat{N}_1(\mathbf{g})$. It is easy to check that if an agent from $\mathcal{K}_z$ is in $\hat{S}_k$, then all agents belonging to $\mathcal{K}_{z-1}$ must be in $\hat{S}_k$. Let $\hat{z}$ denote the largest index such that all agents belonging to $\mathcal{K}_{z-1}$ are in $\hat{S}_k$. If $|\hat{S}_k \cap \mathcal{K}_{\hat{z}}| = |\mathcal{K}_{\hat{z}}|$, then we check $\mathcal{K}_{\hat{z}+1}$. If $\mathcal{K}_{\hat{z}+1} = \emptyset$, then $v(k+1) = v(k)$ since we can't put any agent in $\mathcal{K}_{\hat{z}+\tau}$ into $\hat{S}_k$ where $\hat{z}+1 \leq \hat{z}+\tau \leq |\hat{N}_1(\mathbf{g})|$. If $\mathcal{K}_{\hat{z}+1} \neq \emptyset$, then $v(k+1) = v(k)+1$ since we can put one and only one in $\mathcal{K}_{\hat{z}+1}$ into $\hat{S}_k$ to construct a $(k+1)$-insulated set. If $|\hat{S}_k \cap \mathcal{K}_{\hat{z}}| < |\mathcal{K}_{\hat{z}}|$, then we can put one and only one agent in $\mathcal{K}_{\hat{z}} \setminus \hat{S}_k$ into S_k to construct a $(k+1)$-insulated set, which implies that $v(k+1) = v(k)+1$. The proof of Lemma 3 is done.

Combining step 1 and step 2, we can directly complete the whole proof. □

Proof of Proposition 3. To make a mixed mechanism optimal, a necessary and sufficient condition is that the sender optimally chooses neither the posterior ω_L nor the posterior ω_H in problem (5). We first prove that if $\omega_L \neq z_L$ and $\omega_H \neq z_H$, then the optimal persuasion is a mixed mechanism. Select any two points $(\tau_1, V(\tau_1))$ and $(\tau_2, V(\tau_2))$ on $V(\cdot)$ where

$\omega_L \leq \tau_1 < \tau_2 \leq \omega_H$, and let $((\tau_1, V(\tau_1)), (\tau_2, V(\tau_2)))$ represent the linear segment formed by $(\tau_1, V(\tau_1))$ and $(\tau_2, V(\tau_2))$.

Note that the function $V(\cdot)$ is a concave closure, and thus, it is a concave function, which ensures that the linear segment $((\tau_1, V(\tau_1)), (\tau_2, V(\tau_2)))$ does not lie above the curve. Furthermore, $V(\cdot)$ is a piecewise linear concave function. This implies that if the linear segment $((\tau_1, V(\tau_1)), (\tau_2, V(\tau_2)))$ is between consecutive kink points (i.e., both $(\tau_1, V(\tau_1))$ and $(\tau_2, V(\tau_2))$ lies on the same segment (x_i, x_{i+1}) of $V(\cdot)$ for some i), then it coincides with the function. If the linear segment $((\tau_1, V(\tau_1)), (\tau_2, V(\tau_2)))$ is between nonconsecutive points (i.e., $(\tau_1, V(\tau_1))$ lies on the segment (x_i, x_{i+1}) for some i while $(\tau_2, V(\tau_2))$ lies on another segment $(x_{i'}, x_{i+1})$ where $i' > i$) lies below the function. Since $\omega_L < z_L$ and $\omega_H > z_H$, neither $(\omega_L, v(\omega_L))$ nor $(\omega_H, v(\omega_H))$ lies on the segment $((z_L, v(z_L)), (z_R, v(z_R)))$, and thus the mixed mechanism is optimal. With the same logic, we know that if either $\omega_L = z_L$ or $\omega_H = z_H$, the mixed mechanism is no longer an optimal mechanism. □