Games and Economic Behavior 131 (2022) 84–120

Contents lists available at ScienceDirect

Games and Economic Behavior

www.elsevier.com/locate/geb

Dynamic mechanism design on social networks ☆

Dawen Meng [a,*], Lei Sun [b,*], Guoqiang Tian [c,d,e]

[a] School of Economics, Shanghai University of Finance and Economics, China
[b] School of International Economics and Trade, Shanghai Lixin University of Accounting and Finance, China
[c] Department of Economics, Texas A&M University, United States of America
[d] Institute for Advanced Studies in Finance and Economics, Hubei University of Economics, China
[e] Institute for Advanced Research, Shanghai University of Finance and Economics, China

ARTICLE INFO

Article history:
Received 4 September 2019
Available online 9 November 2021

Keywords:
Dynamic mechanism design
Social network
Nonlinear pricing
Experience good
Key node
Key link
Network intervention

ABSTRACT

This paper studies a mechanism design problem with networked agents and stochastically evolving private information. In contrast to the canonical mechanism design theory focusing only on information asymmetry, we also pay attention to the topology of social network among agents. We find that the standard first-order approach for mechanism design is invalid in dynamic environment. As a remedy, a novel ironing technique is proposed, which produces a perfectly sorting allocation. Based on the optimal dynamic mechanism obtained, we define and compare some important nodes and edges in a network for different ranges of synergy parameter. We further discuss the network intervention problem, in which the principal can intervene to change the ex-ante distribution of individual types.

1. Introduction

The last decade has witnessed significant extensions of the traditional mechanism design theory to dynamic environments. It encompasses a wide range of problems in which an agent's private information evolves stochastically over time and the principal makes her decision in multiple periods. Another important strand of literature growing rapidly is the studies on social network analysis (SNA). Since humans are inherently social beings, an individual's well-being depends not only on his own activity, but also on that of others. A mechanism design problem often takes place in social settings where the agents' behaviors yield spill-overs on each other. The canonical mechanism design theory, however, pays its attention mainly to the information asymmetry between the principal and the agent(s), and ignores largely the role of network structure among agents. This article aims to fill this gap by embedding SNA into a dynamic mechanism design problem. We discuss the sequential nonlinear pricing of experience goods with networked consumers.

Nelson (1970) classifies products into search and experience goods according to the time point when consumers can accurately know their own preference. Search goods are products for which a prospective buyer knows his real preference,

☆ We are thankful to the valuable comments from an anonymous advisory editor and two anonymous referees. We also thank Zhigang Cao, Rongzhu Ke, Simon Loertscher, Zhuoran Lu, Huihua Nie, Alessandro Pavan, Xiang Sun, Satoru Takahashi, Guofu Tan, Jianxia Yang, Jie Zheng, participants of SAET (2018, Taibei), Camphor virtual summer school (2021) and seminar participants at BJTU, JUFE, SUIBE, ECUST, for helpful comments and discussions. Meng gratefully acknowledges financial support from the National Science Foundation of China (NSFC-71301094), Tian gratefully acknowledges financial support from the National Science Foundation of China (NSFC-71371117). All errors are our own.

* Corresponding authors.
E-mail addresses: devinmeng@hotmail.com (D. Meng), cathysunlei@126.com (L. Sun), gtian@tamu.edu (G. Tian).

https://doi.org/10.1016/j.geb.2021.10.007
0899-8256/

either by inspection or by world-of-mouth recommendation, prior to making his purchasing decision. Experience goods refer to the products or services whose attributes can only be ascertained upon one's self-experience.[1] The accurate preference for these goods is unknown to a consumer until his actual consumption is finished.[2] In a typical nonlinear pricing problem (e.g., Mussa and Rosen (1978) and Maskin and Riley (1984)), buyers are assumed to know their own preferences from the outset, the seller designs a static contract to elicit truthful reports and to maximize her own revenue. For experience goods, however, a dynamic mechanism design approach needs to be adopted since a buyer's private information is revealed gradually even to himself.

In this paper, we discuss three different effects in the consumption of experience goods: tolerance effect, network effect and bandwagon effect. The tolerance effect refers to one's inclination to try something different from what he used to consume. A given level of quality becomes less satisfying when one's past consumption has been greater. The network effect means that a user benefits from consumptions of his neighbours who choose the same quality. The bandwagon effect is a psychological phenomenon in which an individual gets fonder of certain quality because others have chosen it. It reflects one's desire to keep up with the Joneses. Tolerance effect reflects the self-influence, while network and bandwagon effects represent the interpersonal influences. Tolerance and bandwagon effects are intertemporal, while network effect is instantaneous.

As a research bridging dynamic mechanism design and SNA, we enrich the existing literature in several dimensions. First, we propose a novel ironing technique to obtain an optimal dynamic mechanism when the traditional method fails. The canonical static mechanism design theory relies heavily on the first-order approach (FOA), which includes three steps. First, identifying necessary (envelope) conditions for incentive compatibility that permit one to express information rents as a function of allocations and express the principal's objective as virtual surplus. Second, optimizing the principal's objective with respect to all possible allocations. Third, verifying whether or not the allocation obtained from the relaxed program is monotone with respect to types so that the allocation and a transfer (rent) satisfying all the local constraints constitute a global incentive-compatible mechanism. The implementability condition holds when certain appropriate primitive conditions on the distribution of types are met. In some cases where these conditions do not hold everywhere, one must resort to a horizontal "ironing" method to derive the optimal allocation, which flattens parts of the benchmark contract and leads to bunching for some types. Things are quite different for dynamic mechanism design problems. Various studies (see Mierendorff (2016), Battaglini and Lamba (2019), and Lu and Wang (2021), among many others) find that the neglected implementability conditions are, in general, violated in dynamic environments even if the primitive conditions on distributions are met. So the applicability of the FOA in dynamic environments is problematic. Though noticing this difficulty, the existing literature hasn't provide a successful remedy in the general sense. In this paper, we also find that the FOA is invalid, and the traditional horizontal ironing technique doesn't work anymore. The optimal dynamic mechanism requires the information rent to be sufficiently convex with respect to types (or equivalently, the slope of the allocation function to be sufficiently large). In view of these findings, we amend the FOA by ironing the relaxed mechanism with an upward-sloping line, and thus the optimal contract attained admits no bunching.

Second, we discuss the identifications of some important nodes and edges in a network within the mechanism design framework. How to identify the key player(s) and key link(s) are important problems in SNA. The existing literature explores these issues within the game-theoretic framework. The key player problem focuses on finding the node whose removal from the initial network would lead to the highest reduction in the aggregate equilibrium activity. Analogously, the key link problem aims to find a link, whose severance reduces total activity the most. In mechanism design setup, problems are different. Besides the key player and key link in traditional senses, the designer cares more about the node and edge having the highest influences on his own profit. In this paper, we define the agent whose isolation from the rest of the network changes the fashion trend (the absolute total quality consumed by all agents) the most as the top influencer, and the agent whose isolation hurts the principal the most as the key customer. We propose different centrality measures to index these agents, and give conditions under which they coincide with each other. The problem of finding a key link in mechanism design setup is also different from its game-theoretic counterpart. When adding a new link to, or sever an existing link from, a network, the designer must take into account both the impact on his own payoff and the incentives of the pair involved. That is, whether adding a new link or severing an existing link will be unanimously agreed upon by both parties. We define a newly added edge which increases the principal's payoff by the largest margin as the key link; the link which is mutually most beneficial to both parties as a stable link. Some negative results are found: the sets of key and stable links are incompatible with each other under weak strategic interactions. It is therefore impossible for the mechanism designer to build a most profit-enhancing link in an incentive compatible manner.

Third, we discuss a network intervention problem based on the optimal dynamic mechanism obtained. In contrast to the traditional paradigm studying the network intervention issue within a game-theoretic framework, the present paper embeds it into a mechanism design framework. We discuss how a profit-maximizing mechanism designer can intervene to change the distribution of individual types subject to a fixed budget constraint.

[1] As an old proverb says: the proof of the pudding is in the eating.

[2] Darby and Karni (1973) introduces the term credence goods and add this type to Nelson's classification. Credence goods are products or services, such as complex automobile repairs or medical services, whose accurate value is unknown to the common users even after consumption due to their lacking of sufficient technical expertise.

Related literature. This paper is related to several lines of research on dynamic mechanism design and social network analysis. The studies on dynamic mechanism design investigate how to implement dynamically efficient (surplus-maximizing) or dynamically optimal (revenue-maximizing) allocations in settings where the agents' private information evolves over time. This strand of research relies largely on the FOA, which is valid in many settings. See, among many others, Baron and Besanko (1984), Laffont and Tirole (1990, 1996), Courty and Li (2000), Battaglini (2005), Esö and Szentes (2007), Garrett and Pavan (2012), Pavan et al. (2014), Halac and Yared (2014), Bergemann and Strack (2015), Armstrong and Zhou (2016), Liu and Lu (2018), Bergemann et al. (2020), Liu et al. (2020), and Meng and Tian (2021) for successful applications of the FOA in various dynamic settings. In particular, Pavan et al. (2014) make an important methodological contribution by extending the traditional envelope condition to a general dynamic environment.

A recent line of research aims at extending the analysis to settings in which the FOA is invalid. Krähmer and Strausz (2011) study an optimal two-period procurement mechanism design problem in the joint presence of adverse selection and moral hazard. They show that in cases where moral hazard causes additional agency cost, the designer needs to adopt an horizontal ironing procedure to amend the relaxed mechanism. Battaglini and Lamba (2019) show that the mechanism obtained from the FOA generically fails to be dynamically incentive compatible. They also present conditions under which the designer can approximate the optimal mechanism with a simpler mechanism admitting a minimal loss in profits. Kakade et al. (2013) identify sufficient conditions under which the FOA fails, and they rely on numerical algorithms to obtain the optimal mechanism. Mierendorff (2016) gives a regularity condition on the type distribution under which the relaxed mechanism satisfies the neglected incentive constraint. For the case that the constraint cannot be neglected, they provide an amending method to adapt the contract in ways that leads to bunching allocations. In a recent work, Lu and Wang (2021) show that the FOA is invalid for a two-stage mechanism design problem with price search. They circumvent this difficulty by taking into account directly the original global incentive constraint, rather than decomposing it into envelope and implementability conditions. Our contribution to this strand of literature is to provide an innovative slope-ironing technique to obtain the optimal dynamic mechanism when the FOA fails.

Another related body of literature focuses on the identification of important nodes and edges in social network. In a network, different individuals play different roles, say, in learning, contagion, and information diffusion, depending on the positions they occupy. Some members are considered more "important" than others with regard to certain criterion. Social network theorists have proposed a number of indexes (say, degree centrality, betweenness, closeness, eigenvector centrality) to measure the relative importance of network locations (see Wasserman and Faust (1994) and Jackson (2008) for detailed review). While the measures given by sociologists are mainly geometric in nature, economists pay more attention to the behavioural foundation of these measures. The seminal work by Ballester et al. (2006) defines an *intercentrality measure*, which not only looks at how central an agent is, but also at how this agent contributes to the centrality of others. This measure helps identify the *key player* in a network game with strategic complementarity. A stream of subsequent researches apply the intercentrality measure to a variety of practical problems. See, among many others, Lindquist and Zenou (2019), König et al. (2019), Lee et al. (2021) and Denbee et al. (2021) for detailed discussions. While it is plausible to define the key player in the game-theoretic setups as the one having the largest influence on the aggregate equilibrium activity, the key agent problem in mechanism design context needs to take into account the influence on the designer's profit as well. The present paper proposes and compares measures for two types of important agents: the *top influencer* and the *key customer*, whose removals have the largest impacts on, respectively, the aggregate activity and total profit.

The "key-link" problem explores how a planner can optimally sever an existing edge from, or to add a new edge to, a network. In a local-aggregate model, Ballester et al. (2010) show that the key link depends on both the individual heterogeneity and the overall network structure. In contrast, Ushchev and Zenou (2020) find in a local-average model that the key link is determined only by the productivity of agents, and is independent of the network structure. In this paper, we extend the key-link problem to a framework of mechanism design, which considers both the principal's profit and bilateral incentives of the two agents involved.

The last class of research related to the present paper is about intervention in networks, which discuss how a utilitarian planner with limited resources can intervene to change individuals' incentives for taking the action. In a recent paper, Galeotti et al. (2020) show that a social planner's optimal intervention on individuals' marginal benefits can be decomposed into orthogonal principal components, which are determined by the network topology and are ordered according to their associated eigenvalues. They focus on the role of total budget and show that the optimal intervention is approximately proportional to the first principal component of network for sufficiently large budget. In contrast, this paper assumes a unit intervening budget, and pays attention to the roles of network synergy parameter and time horizon.

To our knowledge, few of the existing researches incorporates dynamic mechanism design and social network analysis, which are two primary concerns of this paper.[3] The rest of paper is organized as follows. Section 2 characterizes the optimal dynamic mechanism. Section 3 discusses the identification of some important nodes. Section 4 discusses the identification of important edges. Section 5 explores the network intervention problem. Section 6 concludes. The Appendix contains all omitted proofs.

[3] A notable exception is Jadbabaie and Kakhbod (2019). They study an optimal mechanism design problem on social networks. However, it differs from our analysis in two fundamental aspects: (i) It is within the framework of static mechanism design; (ii) The private information possessed by agents in their model is the strength of the network effect rather than the individual preference as in our paper.

2. The optimal dynamic mechanism

Consider a seller who wants to maximize her expected revenue from selling an experience good to n buyers within T periods, $N \equiv \{1, \cdots, n\}$ and $\mathcal{T} \equiv \{0, 1, \cdots, T\}$ denote, respectively, the sets of buyers and time indexes. The seller offers certain quality to buyers at zero cost and collect transfers from them, $x_{it} \in \mathbb{R}$ denotes the quality consumed by buyer i at period t, τ_{it} is the associated transfer, $\mathbf{x}_t \equiv (x_{it})_{i\in N}$ and $\boldsymbol{\tau}_t \equiv (\tau_{it})_{i\in N}$ denote, respectively, the vectors of qualities and transfers, $\mathbf{x}_{-it} \equiv (x_{jt})_{j\neq i}$ denotes the consumption vector of agents other than i.[4]

Buyers are connected via a social network g with a symmetric and irreducible adjacency matrix $\mathbf{G} \equiv [g_{ij}]$,[5] where $g_{ij} = g_{ji} = 1$ if i and j are adjacent, and $g_{ij} = 0$ otherwise. By convention, self-loops are not allowed, i.e., $g_{ii} = 0, \forall i \in N$. $E \equiv \{(i, j) \in N \times N | g_{ij} = 1\}$ denotes the set of edges, $\mathcal{G}_0 \equiv \{(i, j) | i \neq j, g_{ij} = 0\}$ denotes the set of void edges. $N_i \equiv \{j \in N | g_{ij} = 1\}$ represents the neighborhood of individual i, $\mathcal{G}_0^i \equiv \{j | j \neq i, g_{ij} = 0\}$ is the set of potential neighbors of i. To facilitate further analysis, we introduce some centrality measures adopted in prior studies of SNA.

- The *degree centrality* is defined as the number of links incident to node i: $\rho_i \equiv e_i^\top \mathbf{G}\mathbb{1} = e_i^\top \mathbf{G}^2 e_i$, with e_i denoting the ith standard basis vector in $\mathbb{R}^n$, and $\mathbb{1}$ denoting the $n-$dimensional vector of ones.
- The *eigenvector centrality* $EC_i \equiv p_{i1}$, is defined as the ith element of the normalized principal eigenvector $\mathbf{p}_{\cdot 1}$ of the adjacency matrix $\mathbf{G}$. Given our assumption that $\mathbf{G}$ is irreducible (connected), EC_i is well-defined.[6]
- The *subgraph centrality* $SC_i \equiv \sum_{k=0}^{\infty} \alpha_k g_{ii}^{[k]}$ is the weighted sum of all closed walks of different lengths in $\mathbf{G}$ starting and ending both at vertex i, penalizing the contribution of $k-$length walks by α_k, where $g_{ij}^{[k]}$ denotes the (i, j) entry of matrix $\mathbf{G}^k$. α_k varies with the problems considered.
  (*a*) If $\alpha_k = \frac{\beta^k}{k!}$, then $SC_i = e_i^\top \exp(\beta\mathbf{G})e_i$. It is called *exponential subgraph centrality*;
  (*b*) If $\alpha_k = \beta^k$, then $SC_i = e_i^\top \mathbf{M}(\beta, \mathbf{G})e_i = m_{ii}(\beta, \mathbf{G})$, where $\mathbf{M}(\beta, \mathbf{G}) \equiv (\mathbf{I} - \beta\mathbf{G})^{-1}$, $m_{ij}(\beta, \mathbf{G})$ denotes its (i, j) entry. It is called *resolvent subgraph centrality*.
- *Katz-Bonacich centrality* $b_i(\beta, \mathbf{G}) \equiv e_i^\top \mathbf{M}(\beta, \mathbf{G})\mathbb{1}$ counts the weighted sum of walks emanating from a node i. Given a vector $\boldsymbol{\alpha} \in \mathbb{R}_+^n$, $b_i(\beta, \boldsymbol{\alpha}, \mathbf{G}) \equiv e_i^\top \mathbf{M}(\beta, \mathbf{G})\boldsymbol{\alpha}$ is defined as the $\boldsymbol{\alpha}-$weighted Katz-Bonacich centrality.
- The *incercentrality*

$$c_i(\beta, \mathbf{G}) = \frac{b_i^2(\beta, \mathbf{G})}{m_{ii}(\beta, \mathbf{G})} = b_i(\beta, \mathbf{G}) + \sum_{j\neq i} \frac{b_i(\beta, \mathbf{G})m_{ij}(\beta, \mathbf{G})}{m_{ii}(\beta, \mathbf{G})} \tag{1}$$

  defined by Ballester et al. (2006), counts the sum of i's Katz-Bonacich centrality and its contributions to every other player's Katz-Bonacich centralities.

Each buyer has a quadratic per-period utility function

$$u_{it}(\theta_{it}, x_{it}, \mathbf{x}_{-it}) \equiv \theta_{it}x_{it} - \frac{1}{2}x_{it}^2 + \frac{\beta}{2}\sum_{j\neq i} g_{ij}x_{it}x_{jt} - \tau_{it}, \tag{2}$$

where $\theta_{it} \in \mathbb{R}$ stands for consumer i's preference in period t. Here, we extend the well-known model of Mussa and Rosen (1978) by allowing θ_{it} and x_{it} to be either positive or negative. It means that certain quality is not desired by everybody. Quality sold at the same price are ranked from top to bottom by consumers with positive θ_{it} and in the reverse order by those with negative θ_{it}.[7] The first two terms $\theta_{it}x_{it} - x_{it}^2/2$ stand for agent i's standalone utility depending only on his own consumption; the third term $\beta\sum_{j\neq i} g_{ij}x_{it}x_{jt}/2$ represents the network value derived from his interactions with others. A consumer benefits from consumption of his neighbors if they choose qualities of the same sign as him.

The buyer's value evolves according to a stochastic process:

$$\theta_{it+1} = \theta_{it} + s\left(x_{it} - \beta\sum_{j\neq i} g_{ij}x_{jt}\right) + \epsilon_{it}, \forall i \in N, \forall t \in \mathcal{T}\backslash\{T\}. \tag{3}$$

[4] Throughout this paper, a bold lowercase letter represents a vector, while a bold capital letter represents a matrix.

[5] By graph theory, an undirected graph g is connected if and only if its adjacency matrix $\mathbf{G}$ is irreducible.

[6] According to the Perron-Frobenius theorem, when matrix $\mathbf{G}$ is irreducible, $\mathbf{p}_{\cdot 1}$ is unique and satisfies $p_{i1} > 0, \forall i \in N$.

[7] Consider the adoption of social networking service. Positive quality corresponds to the stickiness of a social media, while negative quality correspond to informativeness of it. A social media with strongly positive x_{it} (like Facebook) contains strong ties (families, colleagues, close friends, schoolmates, etc.), is highly sticky, and allows for in-depth communication among users. A social media with strongly negative x_{it} (like Twitter) contains weak ties (strangers, fans, etc.), but it is effective to disseminate information due to the large number of people it can reach. Consumers with positive θ_{it} attach greater importance to communication with their friends and families, while negative θ_{it} are those who care more about sharing ideas and catching up with trends around the world.

$\partial\theta_{it+1}/\partial\theta_{it} > 0$ indicates the persistence of preference. $\partial\theta_{it+1}/\partial x_{it} = s \in (-1, 0)$ captures the *tolerance effect* of consumption: a customer is getting less satisfied with the quality he used to consume due to one's inclination to try something new.[8] $\beta > 0$ is a synergy parameter reflexing the strength of both the instantaneous and intertemporal interactions between linked individuals. $\partial^2 u_{it}/\partial x_{it}\partial x_{jt} = \beta/2 > 0, \forall j \in N_i$ represents the instantaneous *strategic complementarity* across individuals. $\partial\theta_{it+1}/\partial x_{jt} = -s\beta > 0, \forall j \in N_i$ represents the intertemporal *bandwagon effect*: a consumer shifts his own preference towards the direction of qualities chosen by his reference group due to one's desire to keep up with the Joneses. Throughout this paper, we assume that $\beta \in (0, 1/\rho(\mathbf{G}))$, with $\rho(\mathbf{G})$ denoting the spectral radius of matrix $\mathbf{G}$.[9] ϵ_{it} is a stochastic noise term. Let $\boldsymbol{\theta}_t \equiv (\theta_{it})_{i\in N}$, $\boldsymbol{\epsilon}_t \equiv (\epsilon_{it})_{i\in N}$, then we can write (3) in matrix form:

$$\boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}_t + \mathbf{a}_i^\top \mathbf{x}_t + \boldsymbol{\epsilon}_t, \tag{4}$$

where $\mathbf{a}_i^\top$ is the ith row of matrix $\mathbf{A} \equiv s(\mathbf{I} - \beta\mathbf{G})$, with $\mathbf{I}$ denoting the identity matrix.

The initial types $\{\theta_{i0}\}_{i\in N}$ are drawn in an independent manner from uniform distributions over symmetric supports $[\underline{\theta}_i, \overline{\theta}_i]$ with $\underline{\theta}_i + \overline{\theta}_i = 0$. $\overline{\theta}_i$ measures the uncertainty in agent i's valuation. We denote by $f_i(\theta) \equiv 1/(\overline{\theta}_i - \underline{\theta}_i)$ and $F_i(\theta) = (\theta - \underline{\theta}_i)/(\overline{\theta}_i - \underline{\theta}_i)$, the density and distribution functions of θ_{i0}. Hyperrectangle $\mathcal{D} \equiv \prod_{i\in N}[\underline{\theta}_i, \overline{\theta}_i]$ denotes the domain of types. The noise terms $\{\epsilon_{it}\}$ are statistically independent across either i or t:

$$(\epsilon_{it}, \epsilon_{js}) \text{ are independent for either } i \neq j \text{ or } s \neq t. \tag{5}$$

Also, $\{\epsilon_{it}\}$ are identically distributed for a given t with p.d.f $g_t(\cdot)$ and c.d.f $G_t(\cdot)$ over $(-\infty, +\infty)$. It is assumed that $\mathbb{E}(\epsilon_{it}) = 0$ and $\mathbb{V}ar(\epsilon_{it}) = \sigma_t^2, \forall i \in N, \forall t \in \mathcal{T}\backslash\{T\}$. The noise terms are independent of the initial types, that is,

$$(\theta_{i0}, \epsilon_{jt}) \text{ are independent for any pair } (i, j) \in N \times N \text{ and any } t \in \mathcal{T}\backslash\{T\}. \tag{6}$$

In each period t, θ_{it} is private information to buyer i, while its distribution are common knowledge. The seller offers and commits to a dynamic mechanism $\mathcal{M} \equiv \{x_{it}(\widehat{\boldsymbol{\Theta}}_t), \tau_{it}(\widehat{\boldsymbol{\Theta}}_t)\}_{i\in N, t\in\mathcal{T}}$ stipulating qualities and transfers contingent on $\widehat{\boldsymbol{\Theta}}_t \equiv (\hat{\theta}_{js})_{j\in N, s\leqslant t}$, the history of reports up to the time considered.

The overall game unfolds as follows. At the outset of the game, the seller offers to all consumers a dynamic mechanism $\mathcal{M}$, each consumer decides whether to accept or reject it. It is assumed that each buyer can fully commit to the long-term mechanism. So a buyer accepting the mechanism initially are not allowed to exit in the future. At the beginning of each period, the principal will disclose to all buyers the whole history of reports in previous periods, so that the mechanism depends on the realized public history at each date. At date t, agent i observes only his own type θ_{it}, and then submits his report $\hat{\theta}_{it}$ to the principal if he accepts the contract initially. The qualities and transfers $\{x_{it}, \tau_{it}\}_{i\in N, t\in\mathcal{T}}$ are then executed as promised ex-ante, and then the game enters next period.

We denote by

$$\tilde{u}_{it}(\widehat{\boldsymbol{\Theta}}_{t-1}, \theta_{it}, \hat{\boldsymbol{\theta}}_t) \equiv \begin{bmatrix} \theta_{it}x_{it}(\widehat{\boldsymbol{\Theta}}_{t-1}, \hat{\boldsymbol{\theta}}_t) - \frac{1}{2}x_{it}^2(\widehat{\boldsymbol{\Theta}}_{t-1}, \hat{\boldsymbol{\theta}}_t) \\ +\frac{\beta}{2}\sum_{j\neq i} g_{ij}x_{it}(\widehat{\boldsymbol{\Theta}}_{t-1}, \hat{\boldsymbol{\theta}}_t)x_{jt}(\widehat{\boldsymbol{\Theta}}_{t-1}, \hat{\boldsymbol{\theta}}_t) - \tau_{it}(\widehat{\boldsymbol{\Theta}}_{t-1}, \hat{\boldsymbol{\theta}}_t) \end{bmatrix}$$

the instantaneous utility of type θ_{it} under reports profile $\hat{\boldsymbol{\theta}}_t \equiv (\hat{\theta}_{jt})_{j\in N}$ and a public history $\widehat{\boldsymbol{\Theta}}_{t-1}$. We abuse notation to write the truthful instantaneous utility of type θ_{it} as $u_{it}(\widehat{\boldsymbol{\Theta}}_{t-1}, \boldsymbol{\theta}_t) \equiv \tilde{u}_{it}(\widehat{\boldsymbol{\Theta}}_{t-1}, \theta_{it}, \theta_{it}, \boldsymbol{\theta}_{-it})$, where $\boldsymbol{\theta}_{-it} \equiv (\theta_{jt})_{j\neq i}$ denotes the vector of types other than i. Assuming that there is no time discounting, we write the truthful continuation utility of type θ_{it} from date t onward as

$$\mathcal{U}_{it}(\widehat{\boldsymbol{\Theta}}_{t-1}, \theta_{it}) \equiv \left\{ \begin{array}{c} \mathbb{E}_{\boldsymbol{\theta}_{-it}}\left[u_{it}(\widehat{\boldsymbol{\Theta}}_{t-1}, \theta_{it}, \boldsymbol{\theta}_{-it}) \middle| \mathcal{I}_{it}\right] \\ +\sum_{s=t+1}^{T} \mathbb{E}_{\boldsymbol{\theta}_{-it}, \{\boldsymbol{\theta}_\tau\}_{\tau=t+1}^{s}}\left[u_{is}(\widehat{\boldsymbol{\Theta}}_{t-1}, \theta_{it}, \boldsymbol{\theta}_{-it}, \boldsymbol{\theta}_{t+1}, \cdots, \boldsymbol{\theta}_s) \middle| \mathcal{I}_{it}\right] \end{array} \right\},$$

where $\mathcal{I}_{it} \equiv \{\widehat{\boldsymbol{\Theta}}_{t-1}, \theta_{it}\}$ is the information set available to agent i upon public history $\widehat{\boldsymbol{\Theta}}_{t-1}$ and truthful declaration at date t. Assumptions (5) and (6) guarantee that $\{\theta_{it}\}_{i\in N}$ are independent across i for a given $\widehat{\boldsymbol{\Theta}}_{t-1}$. So agent i's private information available at time t conveys no information on the prediction of $\boldsymbol{\theta}_{-it}$. That is, $\mathbb{E}_{\boldsymbol{\theta}_{-it}}[u_{it}(\widehat{\boldsymbol{\Theta}}_{t-1}, \theta_{it}, \boldsymbol{\theta}_{-it})|\mathcal{I}_{it}] = \mathbb{E}_{\boldsymbol{\theta}_{-it}} u_{it}(\widehat{\boldsymbol{\Theta}}_{t-1}, \theta_{it}, \boldsymbol{\theta}_{-it})$.

[8] Say, users often develop tolerance to some prescription medications, such as pain killer, antibiotic and analgesic drugs, due to constant exposure to them. After repeated use, one has to increase the dosage or switch to more potent taking ways, such as snorting or vein injecting, to achieve the same efficacy as before.

[9] This condition is commonly assumed in the network literature (see Ballester et al. (2006)). It guarantees the Neumann series expansion: $(\mathbf{I} - \beta\mathbf{G})^{-1} = \sum_{k=0}^{\infty} \beta^k \mathbf{G}^k$.

The expected continuation utility of type θ_{it} reporting to be $\hat{\theta}_{it}$ at time t is

$$\widetilde{\mathcal{U}}_{it}(\widehat{\boldsymbol{\Theta}}_{t-1}, \theta_{it}, \hat{\theta}_{it}) \equiv \left\{ \begin{array}{c} \mathbb{E}_{\boldsymbol{\theta}_{-it}} \tilde{u}_{it}(\widehat{\boldsymbol{\Theta}}_{t-1}, \theta_{it}, \hat{\theta}_{it}, \boldsymbol{\theta}_{-it}) + \\ \mathbb{E}_{\boldsymbol{\theta}_{-it}} \mathbb{E}_{\theta_{it+1}} \left[\mathcal{U}_{it+1}(\widehat{\boldsymbol{\Theta}}_{t-1}, \hat{\theta}_{it}, \boldsymbol{\theta}_{-it}, \theta_{it+1}) \middle| \widehat{\mathcal{I}}_{it} \right] \end{array} \right\}, \tag{7}$$

$\widehat{\mathcal{I}}_{it} \equiv \{\widehat{\boldsymbol{\Theta}}_{t-1}, \theta_{it}, \hat{\theta}_{it}\}$ is the information set of agent i when he misreports his type at date t. The distribution of one's future type θ_{it+1} depends on his true type θ_{it}, and the current consumption profile of all agents, so expression (7) takes the form

$$\widetilde{\mathcal{U}}_{it}(\widehat{\boldsymbol{\Theta}}_{t-1}, \theta_{it}, \hat{\theta}_{it}) = \left\{ \begin{array}{c} \mathbb{E}_{\boldsymbol{\theta}_{-it}} \tilde{u}_{it}(\widehat{\boldsymbol{\Theta}}_{t-1}, \theta_{it}, \hat{\theta}_{it}, \boldsymbol{\theta}_{-it}) + \\ \mathbb{E}_{\boldsymbol{\theta}_{-it}} \mathbb{E}_{\theta_{it+1}} \left[\mathcal{U}_{it+1}(\widehat{\boldsymbol{\Theta}}_{t-1}, \hat{\theta}_{it}, \boldsymbol{\theta}_{-it}, \theta_{it+1}) \middle| \theta_{it}, \mathbf{x}_t(\widehat{\boldsymbol{\Theta}}_{t-1}, \hat{\theta}_{it}, \boldsymbol{\theta}_{-it}) \right] \end{array} \right\}.$$

To guarantee each agent's truthful declaration upon any public history, and his participation from the outset, the following IR and IC constraints need to be met:

$$IR_{i0} : \mathcal{U}_{i0}(\theta_{i0}) \geqslant 0, \forall i \in N, \forall \theta_{i0} \in [\underline{\theta}_i, \overline{\theta}_i], \tag{8}$$

$$IC_{it} : \mathcal{U}_{it}(\widehat{\boldsymbol{\Theta}}_{t-1}, \theta_{it}) \geqslant \widetilde{\mathcal{U}}_{it}(\widehat{\boldsymbol{\Theta}}_{t-1}, \theta_{it}, \hat{\theta}_{it}), \forall \theta_{it}, \hat{\theta}_{it}, \widehat{\boldsymbol{\Theta}}_{t-1}, \forall i \in N, \forall t \in \mathcal{T}. \tag{9}$$

We have the following lemma regarding condition IC_{it}.

Lemma 1.

- *IC_{it} implies the following envelope (EN_{it}) and implementability conditions (IM_{it}): $\forall \widehat{\boldsymbol{\Theta}}_{t-1}, \forall \theta_{it}$:*

$$EN_{it} : \frac{\partial \mathcal{U}_{it}(\widehat{\boldsymbol{\Theta}}_{t-1}, \theta_{it})}{\partial \theta_{it}} = \mathbb{E}_{\boldsymbol{\theta}_{-it}, \{\boldsymbol{\theta}_s\}_{s \geqslant t+1}} \left[\sum_{s \geqslant t} x_{is}(\widehat{\boldsymbol{\Theta}}_{t-1}, \boldsymbol{\theta}_t, \cdots, \boldsymbol{\theta}_s) \middle| \mathcal{I}_{it} \right], \tag{10}$$

$$IM_{it} : \frac{\partial}{\partial \hat{\theta}_{it}} \mathbb{E}_{\boldsymbol{\theta}_{-it}, \{\boldsymbol{\theta}_\tau\}_{\tau \geqslant t+1}} \left[\sum_{s \geqslant t} x_{is}(\widehat{\boldsymbol{\Theta}}_{t-1}, \boldsymbol{\theta}_t, \cdots, \boldsymbol{\theta}_s) \middle| \mathcal{I}_{it} \right] \geqslant 0; \tag{11}$$

- *Conversely, if EN_{it} and condition*

$$IM'_{it} : \frac{\partial}{\partial \hat{\theta}_{it}} \mathbb{E}_{\boldsymbol{\theta}_{-it}, \{\boldsymbol{\theta}_\tau\}_{\tau \geqslant t+1}} \left[\sum_{s \geqslant t} x_{is}(\widehat{\boldsymbol{\Theta}}_{t-1}, \hat{\theta}_{it}, \boldsymbol{\theta}_{-it}, \cdots, \boldsymbol{\theta}_s) \middle| \widehat{\mathcal{I}}_{it} \right] \geqslant 0, \forall \widehat{\boldsymbol{\Theta}}_{t-1}, \theta_{it}, \hat{\theta}_{it}, \tag{12}$$

are satisfied, then IC_{it} holds.

Proof. See Appendix A. □

The seller's expected profit is written as

$$\begin{aligned} \mathcal{V} &= \mathbb{E}_{\{\boldsymbol{\theta}_s\}_{s \geqslant 0}} \sum_{i \in N} \sum_{t \in \mathcal{T}} \left[\theta_{it} x_{it}(\boldsymbol{\Theta}_t) - \frac{1}{2} x_{it}^2(\boldsymbol{\Theta}_t) + \frac{\beta}{2} \sum_{j \neq i} g_{ij} x_{it}(\boldsymbol{\Theta}_t) x_{jt}(\boldsymbol{\Theta}_t) - u_{it}(\boldsymbol{\Theta}_t) \right] \\ &= \mathbb{E}_{\{\boldsymbol{\theta}_s\}_{s \geqslant 0}} \sum_{i \in N} \sum_{t \in \mathcal{T}} \left[\theta_{it} x_{it}(\boldsymbol{\Theta}_t) - \frac{1}{2} x_{it}^2(\boldsymbol{\Theta}_t) + \frac{\beta}{2} \sum_{j \neq i} g_{ij} x_{it}(\boldsymbol{\Theta}_t) x_{jt}(\boldsymbol{\Theta}_t) \right] \\ &\quad - \sum_{i \in N} \mathbb{E}_{\theta_{i0}} \left\{ \mathbb{E}_{\boldsymbol{\theta}_{-i0}, \{\boldsymbol{\theta}_s\}_{s \geqslant 1}} \left[\sum_{t \in \mathcal{T}} u_{it}(\boldsymbol{\Theta}_t) \middle| \mathcal{I}_{i0} \right] \right\} \\ &= \mathbb{E}_{\{\boldsymbol{\theta}_s\}_{s \geqslant 0}} \sum_{t \in \mathcal{T}} \left[\boldsymbol{\theta}_t^\top \mathbf{x}_t(\boldsymbol{\Theta}_t) - \frac{1}{2} \mathbf{x}_t^\top(\boldsymbol{\Theta}_t) (\mathbf{I} - \beta \mathbf{G}) \mathbf{x}_t(\boldsymbol{\Theta}_t) \right] - \sum_{i \in N} \mathbb{E}_{\theta_{i0}} \mathcal{U}_{i0}(\theta_{i0}). \end{aligned}$$

Following the standard first-order approach in mechanism design literature, we neglect momentarily the implementablility conditions and consider a relaxed problem:

$$[\mathscr{P}_r] : \max_{\{\mathbf{x}_t(\boldsymbol{\Theta}_t)\}_{t \in \mathcal{T}}} \mathcal{V}, s.t. : EN_{it}, IR_{i0}, \forall i \in N, \forall t \in \mathcal{T}. \tag{13}$$

Lemma 2. *The optimal relaxed mechanism entails:*

- *a quality scheme*

$$\boldsymbol{x}_t^*(\boldsymbol{\theta}_0, \boldsymbol{\theta}_t) = \frac{1}{1-(T-t)s}\left(\boldsymbol{I} - \beta \boldsymbol{G}\right)^{-1}\left[\boldsymbol{\theta}_t + \boldsymbol{\vartheta}(\boldsymbol{\theta}_0)\right], \tag{14}$$

where $\boldsymbol{\vartheta}(\boldsymbol{\theta}_0) \equiv (\vartheta_i(\theta_{i0}))_{i\in N}$,

$$\vartheta_i(\theta_{i0}) = \begin{cases} \theta_{i0} - \underline{\theta}_i & if \;\; \theta_{i0} \in \left[\underline{\theta}_i, \frac{\theta_i}{2}\right) \\ -\theta_{i0} & if \;\; \theta_{i0} \in \left[\frac{\theta_i}{2}, \frac{\overline{\theta}_i}{2}\right] \\ \theta_{i0} - \overline{\theta}_i & if \;\; \theta_{i0} \in \left(\frac{\overline{\theta}_i}{2}, \overline{\theta}_i\right] \end{cases}; \tag{15}$$

- *and an associated profit*[10]

$$\mathcal{V}^*(T, \beta, \overline{\boldsymbol{\theta}}, \boldsymbol{G}) = \frac{1}{12}\frac{T+1}{1-Ts}\overline{\boldsymbol{\theta}}^\top \widehat{\boldsymbol{M}}\overline{\boldsymbol{\theta}} + \frac{1}{2}\sum_{t=0}^{T-1}\frac{\sigma_t^2(T-t)}{1-(T-t-1)s}Tr\boldsymbol{M},$$

where $\widehat{\boldsymbol{M}} \equiv diag\{m_{11}, \cdots, m_{nn}\}$, m_{ii}*s are diagonal entries of matrix* $\boldsymbol{M} \equiv (\boldsymbol{I} - \beta\boldsymbol{G})^{-1}$, $\overline{\boldsymbol{\theta}} \equiv (\overline{\theta}_1, \cdots, \overline{\theta}_n)^\top$, $Tr(\cdot)$ *denotes the trace of a square matrix.*

Proof. See Appendix B. □

We next check whether or not the relaxed mechanism obtained from $[\mathscr{P}_r]$ satisfies all the ignored constraints IM'_{it}. In static environment, these constraints are satisfied automatically whenever the monotone hazard rate conditions hold. In our dynamic model, however, these constraints are found to be violated. So, we must resort to some "ironing procedure", which flattens parts of the benchmark schedule, to derive the optimal incentive compatible contract. The critical problem is how to determine the optimal ironing lines.

Proposition 1. *Suppose that the time horizon* T *satisfies*

$$T \geqslant \frac{1}{2s}\left[1 - \sqrt{1 - 4s(1+s)\left(2 + \frac{\sum_{j\neq i} m_{ij}}{m_{ii}}\right)}\right], \forall i \in N, \tag{16}$$

then we have the following results regarding the optimal mechanism:

- *The optimal quality is*

$$\overline{\boldsymbol{x}}_t^*(\boldsymbol{\theta}_0, \boldsymbol{\theta}_t) = \frac{1}{1-(T-t)s}\left(\boldsymbol{I} - \beta\boldsymbol{G}\right)^{-1}\left[\boldsymbol{\theta}_t + \bar{\boldsymbol{\vartheta}}(\boldsymbol{\theta}_0)\right], \tag{17}$$

where $\bar{\boldsymbol{\vartheta}}(\boldsymbol{\theta}_0) \equiv \left(\bar{\vartheta}_i(\theta_{i0})\right)_{i\in N}$,

$$\overline{\vartheta}_i(\theta_{i0}) = \begin{cases} \theta_{i0} - \underline{\theta}_i & if\, \theta_{i0} \in \left[\underline{\theta}_i, \frac{(1+\varphi)\underline{\theta}_i}{1-\varphi}\right) \\ \varphi(\theta_{i0} - \overline{\theta}_i) & if\, \theta_{i0} \in \left[\frac{(1+\varphi)\underline{\theta}_i}{1-\varphi}, \overline{\theta}_i\right] \end{cases} or \begin{cases} \varphi(\theta_{i0} - \underline{\theta}_i) & if\, \theta_{i0} \in \left[\underline{\theta}_i, \frac{(1+\varphi)\overline{\theta}_i}{1-\varphi}\right) \\ \theta_{i0} - \overline{\theta}_i & if\, \theta_{i0} \in \left[\frac{(1+\varphi)\overline{\theta}_i}{1-\varphi}, \overline{\theta}_i\right] \end{cases}. \tag{18}$$

where $\varphi \equiv \frac{T(1-Ts)}{(T+1)(1-Ts+s)} - 1$.

- *The associated optimal profit obtained is*

$$\overline{\mathcal{V}}^*(T, \beta, \overline{\boldsymbol{\theta}}, \boldsymbol{G}) = \left[\begin{array}{c} \frac{T+1}{2(1-Ts)(1-\varphi)^2}\overline{\boldsymbol{\theta}}^\top\left(\varphi^2\boldsymbol{M} + \frac{1+4\varphi^2}{3}\widehat{\boldsymbol{M}}\right)\overline{\boldsymbol{\theta}} \\ +\frac{1}{2}\sum_{t=0}^{T-1}\frac{\sigma_t^2(T-t)}{1-(T-t-1)s}Tr(\boldsymbol{M}) \end{array}\right]. \tag{19}$$

Proof. See Appendix C. □

[10] With a slight abuse of notation, here we write the surplus attained as a function of $T, \beta, \overline{\boldsymbol{\theta}}$ and $\mathbf{G}$.

In the static mechanism design literature, the phenomenon of countervailing incentives arises when the information rent function is $U-$shaped. When it is increasing (resp. decreasing) in the agent's type, the agent has an incentive to understate (resp. overstate) his private information, the optimal contract will prescribe an allocation below (resp. above) the first-best level. For intermediate type realizations, the principal sets a pooling contract that is type-invariant. The neglected monotonicity condition is met provided the monotone hazard rate property (MHRC) holds. In cases of violation, one needs to resort to an horizontal ironing technique to derive the optimal allocation. In our dynamic setup, however, we need to iron the contract even when MHRC holds.[11] More importantly, the traditional horizontal ironing procedure doesn't work any more. Now we must iron the allocation with an upward-sloping line. As a result, the optimal allocation obtained admits no bunching.

To see why a dynamic contract is more difficult to implement than its static counterpart, we give a simple example with two periods ($\mathcal{T} = \{0, 1\}$) and a single agent ($n = 1$). $\mathbf{G}$ is now degenerate into a scalar 0. The relaxed qualities given by expressions (14) and (15) simplify to $x_1^*(\theta_0, \theta_1) = \theta_1 + \vartheta(\theta_0, c^*)$, $x_0(\theta_0) = \frac{\theta_0 + \vartheta(\theta_0, c^*)}{1-s}$, with $c^* = 0$. The envelope condition of period-1 is $\frac{\partial \mathcal{U}_1}{\partial \theta_1}(\theta_0, \theta_1) = x_1^*(\theta_0, \theta_1) = \theta_1 + \vartheta(\theta_0, 0)$, which implies $\mathcal{U}_1(\theta_0, \theta_1) = \frac{1}{2}[\theta_1 + \vartheta(\theta_0, 0)]^2 + \phi(\theta_0)$ for some function $\phi(\cdot)$ depending only on θ_0. The incentive compatible condition of period-0 is thus represented as:

$$\mathcal{U}_0(\theta_0) = \max_{\hat{\theta}_0} \left\{ \theta_0 x_0(\hat{\theta}_0) - \tau_0(\hat{\theta}_0) + \mathbb{E}_{\theta_1} \left[\mathcal{U}_1(\hat{\theta}_0, \theta_1) \middle| \theta_0, \hat{\theta}_0 \right] \right\} = \rho(\theta_0) + \frac{\theta_0^2}{2},$$

where $\rho(\theta_0) \equiv \max_{\hat{\theta}_0}[\theta_0 \alpha(\hat{\theta}_0) + \beta(\hat{\theta}_0)]$, $\alpha(\hat{\theta}_0) \equiv (1+s)x_0(\hat{\theta}_0) + \vartheta(\hat{\theta}_0, 0)$, $\beta(\hat{\theta}_0) \equiv \phi(\hat{\theta}_0) - \tau(\hat{\theta}_0) + [\vartheta(\hat{\theta}_0, 0) + s x_0(\hat{\theta}_0)]^2/2 + \sigma_0^2/2$. In a static mechanism design problem, the information rent needs to be a convex function or equivalently, the allocation needs to be nondecreasing, in type. The dynamic incentive compatibility, however, requires a stronger condition. The convexity of $\rho(\cdot)$[12] implies that $\mathcal{U}_0(\cdot)$ is not only convex but also is more convex than $\theta_0^2/2$. Or equivalently, $x_0(\theta_0)$ ought to increase faster than $\theta_0/2$.[13] Therefore, the relaxed allocations with bunching interval fails to meet the dynamic incentive compatibility anymore.

Removing the virtual valuation $\overline{\boldsymbol{\vartheta}}(\boldsymbol{\theta}_0)$ from expression (17), we get the first-best allocation $\mathbf{x}_t^{FB}(\boldsymbol{\theta}_t) = \frac{1}{1-(T-t)s}(\mathbf{I} - \beta\mathbf{G})^{-1}\boldsymbol{\theta}_t$. It serves as a natural benchmark to which the second-best allocation $\overline{\mathbf{x}}_t^*(\boldsymbol{\theta}_0, \boldsymbol{\theta}_t)$ is compared. The expected distortion of the second-best trade profile relative to the benchmark solution is characterized in terms of the weighted Katz-Bonacich centrality:

$$\begin{aligned} d_{it}(\beta, \overline{\boldsymbol{\theta}}, T, s, \mathbf{G}) &\equiv \mathbb{E}|\mathbf{x}_{it}^{FB}(\boldsymbol{\theta}_t) - \overline{\mathbf{x}}_{it}^*(\boldsymbol{\theta}_0, \boldsymbol{\theta}_t)| \\ &= \frac{1}{1-(T-t)s}\mathbb{E}|e_i^\top(\mathbf{I} - \beta\mathbf{G})^{-1}\overline{\boldsymbol{\vartheta}}_0(\boldsymbol{\theta}_0)| \\ &= -\frac{\varphi b_i(\beta, \overline{\boldsymbol{\theta}}, \mathbf{G})}{[1-(T-t)s](1-\varphi)}, \end{aligned} \tag{20}$$

the last equality follows from $\mathbb{E}\vartheta_i(\theta_{i0}) = \pm\frac{\varphi\overline{\theta}_i}{1-\varphi}$. Given the results in Proposition 1, we now discuss the impacts of the primitive parameters and the underlying network structure on the allocative distortions.

Proposition 2. *The distortion $d_{it}(\beta, \overline{\boldsymbol{\theta}}, T, s, \boldsymbol{G})$ depends on parameters T, t, s and β: $\partial d_{it}/\partial T < 0$, $\partial d_{it}/\partial t > 0$, $\partial d_{it}/\partial s > 0$, $\partial d_{it}/\partial \beta > 0$. Moreover, the ranking of d_{it} across i depends on β,$\overline{\theta}_i$ and the structure of $\boldsymbol{G}$:*

- *For small β, i.e., $\beta \approx 0$, we have:*
  - (i) *If $\overline{\theta}_i$s are heterogeneous across i, then the agent with the largest uncertainty has the largest distortion, i.e., $\arg\max_{i\in N} d_{it}(\beta, \overline{\boldsymbol{\theta}}, T, s, \boldsymbol{G}) = \arg\max_{i\in N} \overline{\theta}_i$;*
  - (ii) *if $\overline{\theta}_i \equiv \overline{\theta}, \forall i \in N$, then the node with the largest vertex degree has the largest distortion, i.e., $\arg\max_{i\in N} d_{it}(\beta, \overline{\boldsymbol{\theta}}, T, s, \boldsymbol{G}) = \arg\max_{i\in N} \rho_i$.*
- *For large β, i.e., $\beta \approx 1/\rho(\boldsymbol{G})$, the node with the largest eigenvector centrality has the largest allocative distortion, i.e., $\arg\max_{i\in N} d_{it}(\beta, \overline{\boldsymbol{\theta}}, T, s, \boldsymbol{G}) = \arg\max_{i\in N} p_{i1}$.*

Proof. See Appendix D. □

Due to the intertemporal consumption-preference dependence, a one-shot distortion on allocation propagates through both the types process and the network. It thus produces a more persistent and more widespread influence on the efficiency

[11] Monotone hazard rate conditions $\frac{d}{d\theta}\frac{1-F_i(\theta)}{f_i(\theta)} \leqslant 0 \leqslant \frac{d}{d\theta}\frac{F_i(\theta)}{f_i(\theta)}$ are trivially satisfied under our assumption of uniform distribution.

[12] Notice that $\rho(\cdot)$ is the pointwise maximum of a class of affine functions, so it is obviously convex.

[13] It follows from the convexity of $\rho(\theta_0)$ that $\rho'(\theta_0) = \alpha(\theta_0) \equiv (1+s)x_0(\theta_0) + \vartheta_0(\theta_0, 0) = (1+s)x_0(\theta_0) + (1-s)x_0(\theta_0) - \theta_0 = 2x_0(\theta_0) - \theta_0$ is increasing in θ_0. So we have $x_0'(\theta_0) > 1/2$.

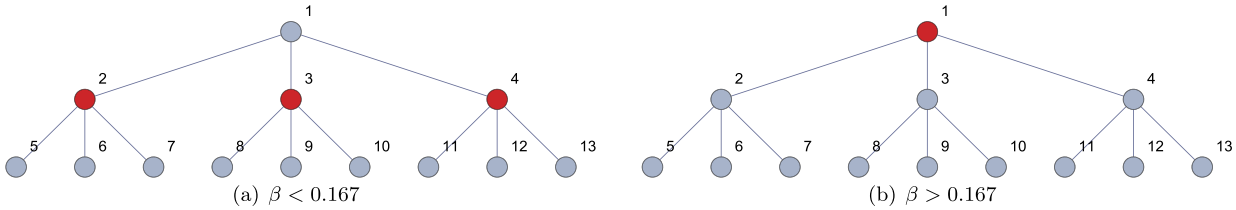


Fig. 1. Nodes with the largest allocative distortions in a three-tier tree

achieved than its static counterpart. The traditional rent extraction vs. efficiency trade off thus calls for a less distorted allocation profile. When either the number of periods remaining $(T-t)$ becomes larger, or the intertemporal dependence (s) becomes stronger,[14] an allocative distortion has greater influence on the future efficiencies. So the principal needs to provide a mechanism with smaller distortions. If $s=0$, there is no intertemporal impact, the distortion d_{it} is identical across t, and attains its maximum $b_i(\beta,\overline{\boldsymbol{\theta}},\mathbf{G})/(T+2)$.

Since parameter β captures the strength of both the strategic complementarities and the bandwagon effects, the distortion due to asymmetric information are therefore augmented by a larger β. Note that for β close to zero, the allocative distortions are ranked first by the individual uncertainty $\overline{\theta}_i$, and then by vertex degree ρ_i in case of identical $\overline{\theta}_i$s; for large β, however, the ranking of distortions is independent of any individual uncertainty and is determined solely by the eigenvector centrality of nodes. We proceed to discuss how the change in an individual's uncertainty affects allocative distortions of all agents. Let $\Delta d_{ijt}(\beta,\Delta\overline{\theta},T,s,\mathbf{G}) \equiv d_{it}(\beta,\overline{\boldsymbol{\theta}}+\Delta\overline{\boldsymbol{\theta}}_j,T,s,\mathbf{G}) - d_{it}(\beta,\overline{\boldsymbol{\theta}},T,s,\mathbf{G})$ be the change of i's distortion upon an increment in j's uncertainty, where $\Delta\overline{\boldsymbol{\theta}}_j = (0,\cdots\Delta\overline{\theta}\cdots,0)^\top$ denotes a vector with $\Delta\overline{\theta}>0$ at its jth position, and zeros at all other positions. We then have the following results.

Proposition 3. *The uncertainty-distortion impacts are always positive:* $\Delta d_{ijt}(\beta,\Delta\overline{\theta},T,s,\boldsymbol{G})>0,\forall i,\forall j,\forall t$. *As* $\beta\to 0$, $\Delta d_{ijt}(\beta,\Delta\overline{\theta},T,s,\boldsymbol{G}) = O[\beta^{k(i,j)}],\forall t\in\mathcal{T}$, *where* $k(i,j)\equiv\min\{s|g_{ij}^{[s]}>0\}$ *is the geodesic distance between nodes i and j*[15]*; as* $\beta\to 1/\rho(\boldsymbol{G})$, $\arg\max_i \Delta d_{ijt}(\beta,\Delta\overline{\theta},T,s,\boldsymbol{G}) = \arg\max_j \Delta d_{ijt}(\beta,\Delta\overline{\theta},T,s,\boldsymbol{G}) = \arg\max_{s\in N} p_{s1},\forall t\in\mathcal{T}$.

Proof. See Appendix E. □

This theorem shows that a change in the initial uncertainty of an agent has permanent impacts on the allocative distortions in each subsequent period. For a fixed t, this impact exhibits a "ripple effect" whenever β is sufficiently small: change in $\overline{\theta}_j$ has the most prominent influence on the distortion of agent j himself, then it propagates throughout the network. It decays radially outward at a rate β. On the concentric circle with radius r $\{i\in N|k(i,j)=r\}$, the impact has an order of $O(\beta^r)$. In contrast, when β is close to its upper bound $1/\rho(\mathbf{G})$, the interpersonal impact between two agents is independent of their geodesic distance. A node with the largest eigenvector centrality influence, and is also influenced by, others in the most prominent manner.

Example 1. Fig. 1 marks with red nodes (To see the colored figures, the reader is referred to the web version of this article.) the agents having the largest allocative distortions for a three tier tree graph and an identical $\overline{\theta}_i=\overline{\theta},\forall i\in N$. The second tier nodes 2, 3, 4 have the largest vertex degree ($\rho_1=3,\rho_2=\rho_3=\rho_4=4,\rho_5=\cdots=\rho_{13}=1$), while the root node 1 has the largest eigenvector centrality ($p_{11}=0.5,p_{21}=p_{31}=p_{41}=0.408,p_{51}=\cdots=p_{13,1}=0.166$). Fig. 2 shows that when β is close to zero, the second tier nodes 2, 3 and 4 have the largest distortion; when $\beta\approx 1/\rho(\mathbf{G})$, however, the root node 1 admits the largest distortion. Fig. 3 depicts m_{ij}s as functions of β. Panel (a) shows that when β is close to zero, a change in agent 3's uncertainty has the largest impact on his own distortion. The impacts of agents 1, 8,9 and 10, whose geodesic distance to 3 are all 1, are in the same order of $O(\beta)$ (note that curves $m_{31}(\beta,\mathbf{G})$ and $m_{38}(\beta,\mathbf{G})$ are very close for sufficiently small β). The second tier nodes 2 and 4 have a common geodesic distance $k(2,3)=k(4,3)=2$ and impacts of order $O(\beta^2)$. The third tier nodes 5,6,7,11,12,13 have a common distance 3, and thus have impacts of order $O(\beta^3)$. Panel (b) of Fig. 3 shows that for $\beta\approx 1/\rho(\mathbf{G})=0.408$, the root node 1 with the largest eigenvector centrality has the largest influence on the distortion of node 3.[16]

Next, we turn to two extreme cases where the first-best allocation is restored. Taking limits as $s\to -1$ or $T\to\infty$ in expressions (19) and (20), we immediately get $\varphi\to 0$, and thus the following results.

Proposition 4. *The first-best allocation is achieved when either* $s\to -1$ *or* $T\to\infty$.

[14] Note that a smaller s has a large magnitude and thus represents a stronger intertemporal dependence.
[15] Note that $k(i,i)=0,\forall i\in N$.
[16] Note that the impact of 1 is even larger than that of 3 itself.

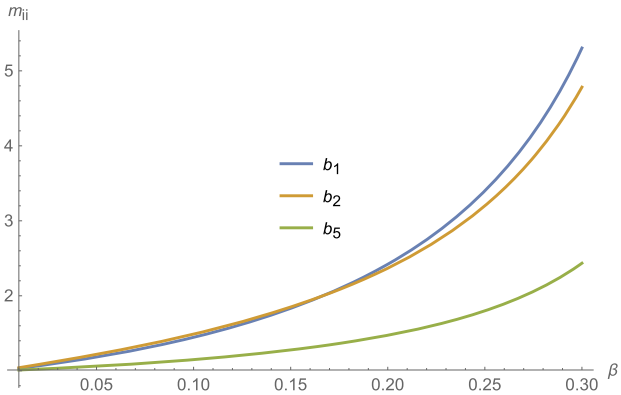


Fig. 2. $b_i(\beta, \mathbf{G})$ for network in Fig. 1.

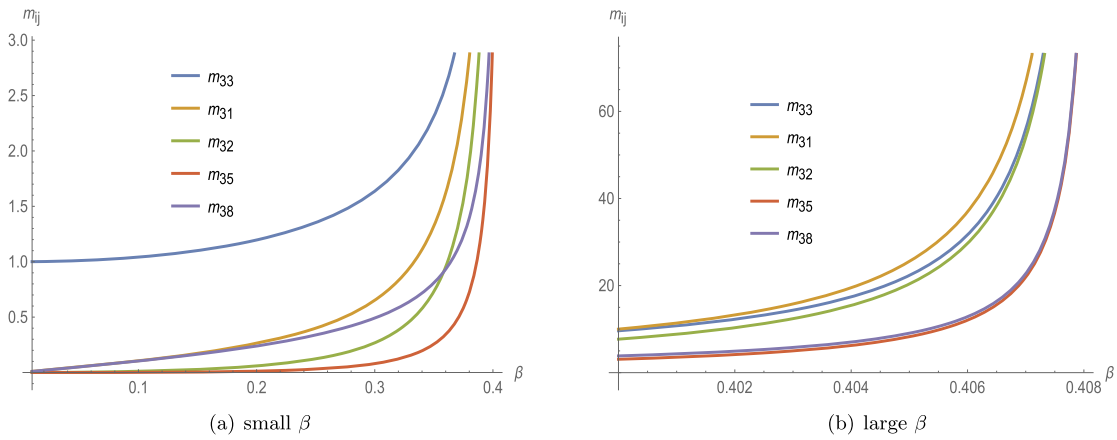


Fig. 3. $m_{ij}(\beta, \mathbf{G})$ for network in Fig. 1.

- *If* $s \to -1$, $d_{it}(\beta, \overline{\boldsymbol{\theta}}, T, s, \boldsymbol{G}) \to 0$,

$$\overline{\mathcal{V}}^*(\beta, \overline{\boldsymbol{\theta}}, T, s, \boldsymbol{G}) \to \frac{1}{6}\overline{\boldsymbol{\theta}}^\top \widehat{\boldsymbol{M}}\overline{\boldsymbol{\theta}} + \frac{1}{2}\sum_{t=0}^{T-1}\sigma_t^2 Tr\boldsymbol{M};$$

- *if* $T \to \infty$ *and* $\sum_{t=0}^{\infty}\sigma_t^2 < \infty$, *we also have* $d_{it}(\beta, \overline{\boldsymbol{\theta}}, T, s, \boldsymbol{G}) \to 0$,

$$\overline{\mathcal{V}}^*(\beta, \overline{\boldsymbol{\theta}}, T, s, \boldsymbol{G}) \to -\frac{1}{2s}\left(\frac{1}{3}\overline{\boldsymbol{\theta}}^\top \widehat{\boldsymbol{M}}\overline{\boldsymbol{\theta}} + \sum_{t=0}^{\infty}\sigma_t^2 Tr\boldsymbol{M}\right).$$

Condition $\sum_{t=0}^{\infty}\sigma_t^2 < \infty$ guarantees the boundedness of total surplus under infinite horizon. It requires that the noises $\{\epsilon_{it}\}$ vanish gradually as time passes. Proposition 4 gives novel results of full surplus extraction (FSE) beyond the logic à la Crémer and McLean (1985, 1988). Their seminal work shows that if agents have correlated private information, then the uninformed principal can exploit this correlation to cross-check their truthful reports and thus extracts all social surplus in a Bayesian mechanism. Our result instead focuses on the role of intertemporal influence s and time horizon T, and thus provides a new explanation on FSE.

If the intertemporal effect is at its strongest possible level $s = -1$, even very small current distortion may produce an unboundedly large impact on future efficiency. This will discourage the principal's distortion motive and restore the first-best allocation, i.e., $d_{it} \equiv 0$.

The time horizon T works as follows. The impact of an agent's initial misreport $\hat{\theta}_{i0}$ on his expected lifetime total allocation works through three channels: (a) it affects directly the allocation of the initial period, i.e., $x_{i0}^*(\hat{\theta}_{i0}, \boldsymbol{\theta}_{-i0})$; (b) the impact on $\mathbf{x}_0^*(\hat{\theta}_{i0}, \boldsymbol{\theta}_{-i0})$ propagates to the subsequent periods through the stochastic process of type-allocation independence, and thus affects the expectation on future allocations; (c) it also affects the allocations of all future periods because of the history-contingent nature of dynamic mechanism. We mark these channels in the following expression (21):

$$I_i(\theta_{i0},\hat{\theta}_{i0}) \equiv \mathbb{E}_{\boldsymbol{\theta}_{-i0},\{\boldsymbol{\theta}_s\}_{s\geqslant 1}}\left[\underset{\underset{(a)}{\downarrow}}{x^*_{i0}(\hat{\theta}_{i0},\boldsymbol{\theta}_{-i0})} + \sum_{t=1}^{T}\underset{\underset{(c)}{\downarrow}}{x^*_{it}(\hat{\theta}_{i0},\boldsymbol{\theta}_{-i0},\boldsymbol{\theta}_t)}\,\middle|\,\theta_{i0},\underset{\underset{(b)}{\downarrow}}{\mathbf{x}^*_0(\hat{\theta}_{i0},\boldsymbol{\theta}_{-i0})}\right]. \tag{21}$$

Since channels (a) and (b) work only through their impacts on the first-period allocations, their impacts vanish gradually as the time horizon becomes large. The impact through channel (c) is, however, persistent. Expressions (83) to (85) in Appendix C help to decompose $\partial I_i(\theta_{i0},\hat{\theta}_{i0})/\partial\hat{\theta}_{i0}$ into three effects corresponding to three channels:

$$\frac{\partial I_i}{\partial\hat{\theta}_{i0}}(\theta_{i0},\hat{\theta}_{i0}) = \underbrace{\frac{m_{ii}[1+\vartheta_i'(\hat{\theta}_{i0})]}{1-Ts}}_{\substack{\parallel\\ ⓐ}} + \underbrace{\frac{Tsm_{ii}[1+\vartheta_i'(\hat{\theta}_{i0})]}{[1-(T-1)s](1-Ts)}}_{\substack{\parallel\\ ⓑ}} + \underbrace{\frac{Tm_{ii}\vartheta_i'(\hat{\theta}_{i0})}{1-(T-1)s}}_{\substack{\parallel\\ ⓒ}}.$$

As $T\to\infty$, terms ⓐ and ⓑ tend to zero, while term ⓒ goes to $-m_{ii}\vartheta_i'(\hat{\theta}_{i0})/s$. Hence, $IM_0' : \partial I(\theta_{i0},\hat{\theta}_{i0})/\partial\hat{\theta}_{i0}\geqslant 0$ requires $\vartheta_i'(\hat{\theta}_{i0})\geqslant 0, \forall\hat{\theta}_{i0}$. Following the well-known results in countervailing incentive literatures (e.g., Lewis and Sappington (1989), Maggi and Rodriguez-Clare (1995)), when the allocation changes its sign across the entire type space, the informational rent is $U-$shaped with both extreme types earning no rents. The tradeoff between rent extraction and efficiency calls for efficient allocations at two extreme types, and downward (resp. upward) distortion near the top (resp. bottom) type. That is, the virtual valuation satisfies $\vartheta_i(\theta_{i0}) = F_i(\theta_{i0})/f_i(\theta_{i0}) > 0$ near $\underline{\theta}_i$, and $\vartheta_i(\theta_{i0}) = [F_i(\theta_{i0})-1]/f_i(\theta_{i0}) < 0$ near $\overline{\theta}_i$. Any "second best" path of virtual valuation starting at the bottom with the efficient value and ending up at the top with also the efficient value is inevitably decreasing for some intermediate realizations of θ_{i0}. This clearly yields a contradiction with $\vartheta_i'(\hat{\theta}_{i0})\geqslant 0, \forall\hat{\theta}_{i0}\in[\underline{\theta}_i,\overline{\theta}_i]$ required by IM_{i0}'. As a result, the principal will set $\vartheta_i(\theta_{i0})\equiv 0$ and thus restores the complete information allocations.

Indeed, the first-best allocation $\mathbf{x}_t^{FB}(\boldsymbol{\theta}_t)$ satisfies conditions $IM_{i0}', \forall i\in N$. To see this, note that $\mathbf{x}_t^{FB}(\boldsymbol{\theta}_t), t\geqslant 1$ are independent of $\hat{\theta}_{i0}$, so channel (c) vanishes, $\hat{\theta}_{i0}$ affects $I_i(\theta_{i0},\hat{\theta}_{i0})$ only through its impact on $\mathbf{x}_0^{FB}(\hat{\theta}_{i0},\boldsymbol{\theta}_{-i0})$ (channels (a) and (b)). However, $\mathbf{x}_0^{FB}(\hat{\theta}_{i0},\boldsymbol{\theta}_{-i0}) = \frac{1}{1-Ts}(\mathbf{I}-\beta\mathbf{G})^{-1}(\hat{\theta}_{i0},\boldsymbol{\theta}_{-i0})$ itself is infinitesimally small as time horizon goes to infinity. It means that an agent's allocations under both truthful declaration and misreport are negligibly small. So an agent has little to gain from misreports, the first-best allocation is therefore approximately incentive compatible.

3. Identification of important nodes

With the results of Proposition 1 in hand, we now proceed to discuss the impact of network structure on the total surplus accrues to the principal and the aggregate quality consumed by the agents. To focus on the role of network topology, we assume in this section that there is no individual heterogeneity, i.e., $\overline{\theta}_i\equiv\overline{\theta}, \forall i\in N$. It follows directly from (19) that the principal's optimal revenue takes the form[17]

$$\overline{\mathcal{V}}^*(\beta,\mathbf{G}) = \phi_1\mathbb{1}^\top(\mathbf{I}-\beta\mathbf{G})^{-1}\mathbb{1} + \phi_2 Tr(\mathbf{I}-\beta\mathbf{G})^{-1}, \tag{22}$$

where

$$\phi_1 \equiv \frac{1}{2}\left(\frac{\varphi\overline{\theta}}{1-\varphi}\right)^2\frac{T+1}{1-Ts}, \phi_2 \equiv \frac{\overline{\theta}^2(1+4\varphi^2)(T+1)}{6(1-\varphi)^2(1-Ts)} + \frac{1}{2}\sum_{t=0}^{T-1}\frac{(T-t)\sigma_t^2}{1-(T-t-1)s}. \tag{23}$$

The expectation of lifetime aggregate quality consumed by all consumers is

$$\mathcal{Q}(\beta,\mathbf{G}) \equiv \mathbb{E}\sum_{t\in\mathcal{T}}\sum_{j\in N}\overline{x}^*_{jt}(\boldsymbol{\theta}_0,\boldsymbol{\theta}_t) = \sum_{t\in\mathcal{T}}\mathbb{1}^\top\mathbb{E}\overline{\mathbf{x}}^*_t(\boldsymbol{\theta}_0,\boldsymbol{\theta}_t) \tag{24}$$

$$= \sum_{t\in\mathcal{T}}\mathbb{1}^\top\mathbb{E}_{\boldsymbol{\theta}_0}\mathbb{E}_{\boldsymbol{\theta}_1}\left[\cdots\mathbb{E}_{\boldsymbol{\theta}_t}\left[\overline{\mathbf{x}}^*_t(\boldsymbol{\theta}_0,\boldsymbol{\theta}_t)|\mathcal{I}_{t-1}\right]\cdots\middle|\mathcal{I}_0\right] \tag{25}$$

$$= \pm\sum_{t\in\mathcal{T}}\frac{\varphi\overline{\theta}}{(1-Ts)(1-\varphi)}\mathbb{1}^\top(\mathbf{I}-\beta\mathbf{G})^{-1}\mathbb{1} \tag{26}$$

$$= \pm\frac{(T+1)\varphi\overline{\theta}}{(1-Ts)(1-\varphi)}\mathbb{1}^\top(\mathbf{I}-\beta\mathbf{G})^{-1}\mathbb{1}. \tag{27}$$

(24) $\Rightarrow$ (25) follows from repeated applications of the Law of Iterated Expectation. (25) $\Rightarrow$ (26) follows from $\mathbb{1}^\top\mathbb{E}\overline{\mathbf{x}}^*_0(\boldsymbol{\theta}_0) = \pm\frac{\varphi\overline{\theta}}{(1-Ts)(1-\varphi)}\mathbb{1}^\top(\mathbf{I}-\beta\mathbf{G})^{-1}\mathbb{1}$, the sign depends on whether we choose $\mathbf{d}^* = \overline{\theta}\mathbb{1}$ or $\mathbf{d}^* = -\overline{\theta}\mathbb{1}$ as the ironing vector.

[17] To economize on notation, we omit all arguments but β and $\mathbf{G}$.

Now, we come to the identification of important nodes in network **G**. This is an important and long-standing problem in social network literature. The criterion for "importance" varies with the problem considered. In our setup, the seller needs to pay attention to two types of important agents: the *top influencer*(TI) and the *key customer*(KC). The $TI(i^*)$ refers to a buyer whose removal from the network leads to the largest change in the fashion trend, namely, the aggregate quality consumed by all agents in equilibrium. $KC(j^*)$ is the one whose removal results in the largest reduction in the profit obtained by the seller from the residual network. They are solutions to, respectively, the following problems[18]:

$$[\mathscr{T}] : \max_{i\in N} \left| \mathcal{Q}(\beta, \mathbf{G}) - \mathcal{Q}(\beta, \mathbf{G}^{-i}) \right| \text{ and } [\mathscr{K}] : \max_{i\in N} \left[\overline{\mathcal{V}}^*(\beta, \mathbf{G}) - \overline{\mathcal{V}}^*(\beta, \mathbf{G}^{-i}) \right],$$

where $\mathbf{G}^{-i}$ is the adjacency matrix obtained from **G** by replacing with zeros all of its ith row and ith column entries:

$$\mathbf{G} = \begin{bmatrix} 0 & \mathbf{g}_i^\top \\ \mathbf{g}_i & \mathbf{G}_{-i,-i} \end{bmatrix}, \mathbf{G}^{-i} = \begin{bmatrix} 0 & \mathbf{0}^\top \\ \mathbf{0} & \mathbf{G}_{-i,-i} \end{bmatrix}.$$

$[\mathscr{T}]$ and $[\mathscr{K}]$ are both finite optimization problems, which admit at least one solution. The key problem is to obtain simple measures for TI and KC.

In game-theoretic framework, a planner conventionally wishes to target, among the population, the most influential individual who has the largest influence on the aggregate equilibrium activities. For example, in the public health context, a policy maker needs to select a subset of population members to immunize or quarantine in order to optimally contain an epidemic. In the criminal justice context, a planner needs to select a small number of players in a criminal network to neutralize (e.g., by arresting, exposing or discrediting) in order to maximally disrupt the network's ability to mount coordinated action. In the present paper, we pursue this line of research by identifying the top influencer in a nonlinear pricing problem.

In contrast to the network game model, the principal pays more attention to his total profit in our mechanism design environment. An important node from the seller's optimality concern, rather than from the agents' strategic consideration, needs to be identified. So we define and measure the KC whose removal leads to the largest loss to the seller's profit. This gives rise to a natural question: do TI and KC coincide with each other? Put differently, whether or not the objective of regulating the overall network activities is in accord with the interests of a profit-maximizing planner himself? We find that the answer is generally no. But it is in the affirmative for some special cases. Next, we provide simple and direct indexes for TI and KC.

Proposition 5. *The top influencer is the one with the highest intercentrality measure $c_i(\beta, \boldsymbol{G})$, while the key customer is the one with the highest index*

$$h_i(\beta, \boldsymbol{G}) \equiv \phi_1 c_i(\beta, \boldsymbol{G}) + \phi_2 \frac{m_{ii}^{[2]}(\beta, \boldsymbol{G})}{m_{ii}(\beta, \boldsymbol{G})}, \tag{28}$$

where $m_{ii}^{[2]}(\beta, \boldsymbol{G})$ is the ith diagonal entry of matrix $\boldsymbol{M}^2 \equiv (\boldsymbol{I} - \beta \boldsymbol{G})^{-2}$, ϕ_1 and ϕ_2 is given by (23).

Proof. See Appendix F. □

In a complete information network game, Ballester et al. (2006) define the *key player* as the one whose removal induces the largest aggregate activity reduction, and show that the intercentrality measure c_i helps determine it. In our mechanism design setup, the *top influencer* is also indexed by c_i. The *key customer* is, however, defined from the perspective of the seller, and is indexed by centrality measure h_i. Indexes c_i and h_i are related, but they do not always provide the same ranking. Both TI and KC vary with the synergy parameter β, which serves as a decay factor penalizing the contribution of longer walks. For extreme case of $\beta \to 0$, we find that rankings produced by c_i and h_i are aligned, so TI and KC coincide with each other.

Proposition 6. *Suppose that β is close to zero, then we have: (i) If **G** is an irregular network, TI and KC coincide with each other: they are both the agent(s) with the highest vertex degree; (ii) if **G** is regular and $g_{ii}^{[3]}$ is not identical across i, then $TI(i^*)$ and $KC(j^*)$ are different: $i^* \in \arg\min_{i\in N} g_{ii}^{[3]}$, $j^* \in \arg\max_{i\in N} g_{ii}^{[3]}$, where $g_{ii}^{[3]}$ is the ith diagonal element of matrix $\boldsymbol{G}^3$.*

Proof. See Appendix G. □

We illustrate part (ii) with the following example.

[18] Since $\mathcal{Q}(\beta, \mathbf{G})$ may be either positive or negative depending on the choice of ironing vector $\mathbf{d}^*$, we use the absolute value to denote the overall impact of an agent's removal.

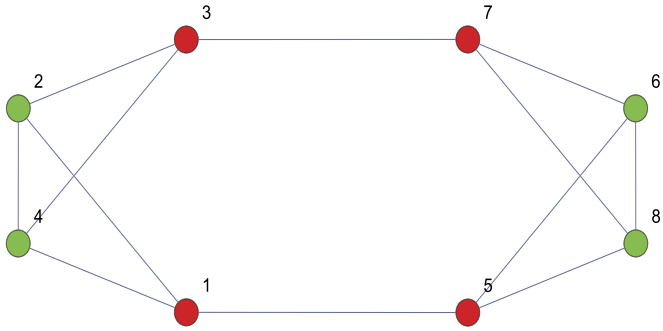


Fig. 4. Discrepancy of TI and KC in a 3-regular network.

Table 1
Indexes of nodes in Fig. 4.

i	1	2
ρ_i	3	3
$g_{ii}^{[3]}$	2	4
c_i	1.9735	1.9695
h_i	59.239	59.5844

Example 2. Fig. 4 gives an eight-nodes 3−regular network with $T = 10, \beta = 0.1, s = -0.3, \overline{\theta} = 10, \sigma_t \equiv 1, \forall t \in \mathcal{T}\backslash\{T\}$. The red and green nodes denote the TIs and KCs, respectively. Table 1 reports the indexes $\rho_i, g_{ii}^{[3]}, c_i$ and h_i.

In the opposite extreme case where $\beta \to 1/\rho(\mathbf{G})$, we also have the coincidence of TI and KC.

Proposition 7. *(i) Suppose β is close to $1/\rho(\boldsymbol{G})$,*

$$\Delta_i + \underline{\delta} \geqslant \Delta_j + \overline{\delta}, \forall i \in \mathbb{P}_1, \forall j \in \mathbb{P}_2, \tag{29}$$

then TI and KC are both within $\mathbb{P}_1$, where $\mathbb{P}_1 \equiv \arg\max_{i\in N} \Delta_i$, $\mathbb{P}_2 \equiv \arg\max_{i\in N\backslash\mathbb{P}_1} \Delta_i$,

$$\Delta_i \equiv -\frac{1}{p_{i1}^2}\sum_{j\neq 1}\frac{p_{ij}^2}{\lambda_1-\lambda_j}, \delta_i \equiv \frac{2}{p_{i1}}\sum_{j\neq 1}\frac{p_{ij}(\boldsymbol{p}_{\cdot j}^\top \mathbb{1})}{(\lambda_1-\lambda_j)(\boldsymbol{p}_{\cdot 1}^\top \mathbb{1})},$$

$\overline{\delta} = \max_{i\in N}\delta_i, \underline{\delta} = \min_{i\in N}\delta_i$, $\boldsymbol{p}_{\cdot 1} \equiv \{p_{i1}\}_{i\in N}$ is the principal eigenvector of $\boldsymbol{G}$, $\lambda_1 > \lambda_2 \geqslant \cdots \geqslant \lambda_n$ denote the eigenvalues of $\boldsymbol{G}$ in decreasing order. (ii) If we further assume that

$$\underline{\Delta}_i > \overline{\Delta}_j, \forall i \in \mathbb{Q}_1, \forall j \in \mathbb{Q}_2, \tag{30}$$

then TI and KC are both within $\mathbb{Q}_1$, where $\underline{\Delta}_i \equiv \frac{1}{\lambda_1-\lambda_2}\left(1-\frac{1}{p_{i1}^2}\right)$, $\overline{\Delta}_i \equiv \frac{1}{\lambda_1-\lambda_n}\left(1-\frac{1}{p_{j1}^2}\right)$, $\mathbb{Q}_1 \equiv \arg\max_{i\in N} p_{i1}$ and $\mathbb{Q}_2 \equiv \arg\max_{i\in N\backslash\mathbb{Q}_1} p_{i1}$.

Proof. See Appendix H. □

Notice that $\{\mathbf{p}_{\cdot j}\}_{j=1}^n$ constitutes a canonical basis of space $\mathbb{R}^n$, so we have $\sum_{j=1}^n(\mathbf{p}_{\cdot j}^\top\mathbb{1})^2 = n\sum_{j=1}^n\cos^2(\mathbf{p}_{\cdot j}, \mathbb{1}) = n$. If the eigenvector centralities are evenly distributed across nodes, we have $\mathbf{p}_{\cdot 1}^\top\mathbb{1} \approx \sqrt{n}$ and $\mathbf{p}_{\cdot j}^\top\mathbb{1} \approx 0, \forall j \neq 1$. Hence, δ_i is negligibly small compared with Δ_i, so condition (29) holds. Condition (30) is met when the gap between the second largest and the smallest eigenvalues $\lambda_2 - \lambda_n$ is sufficiently small. Note that a connected graph satisfies $\lambda_2 = \lambda_n$ if and only if it is a complete graph K_n. So condition (30) holds when $\mathbf{G}$ is akin to K_n, or put differently, when $\mathbf{G}$ is sufficiently dense.[19]

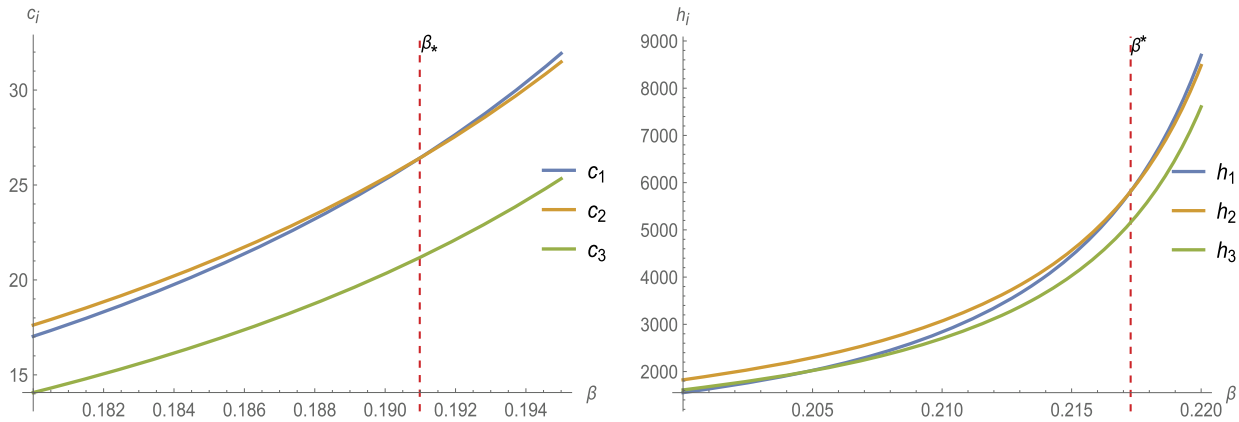
Example 3. Consider the eleven-nodes bridge network $\mathbf{G}$ in Fig. 6.[20] Let $T = 30, s = -0.2, \overline{\theta} = 20, \sigma_t \equiv 1, \forall t \in \mathcal{T}\backslash\{T\}$. There are three different locations in this network. Type I: player 1; type II: players 2, 6, 7 and 11; and type III: players 3, 4, 5, 8, 9, and 10. Type I and type III players have four direct links, while type II players have five. Player 1 bridges together two fully intra-connected communities of five players each. Table 2 computes some important measures of $\mathbf{G}$.

[19] Notice that K_n is the densest one among the class of simple graphs since every pair of its distinct vertices is connected.

[20] This example was given by Ballester et al. (2006) to illustrate the identification of key player in a static complete information network game. We adopt this example to facilitate comparison with prior studies.

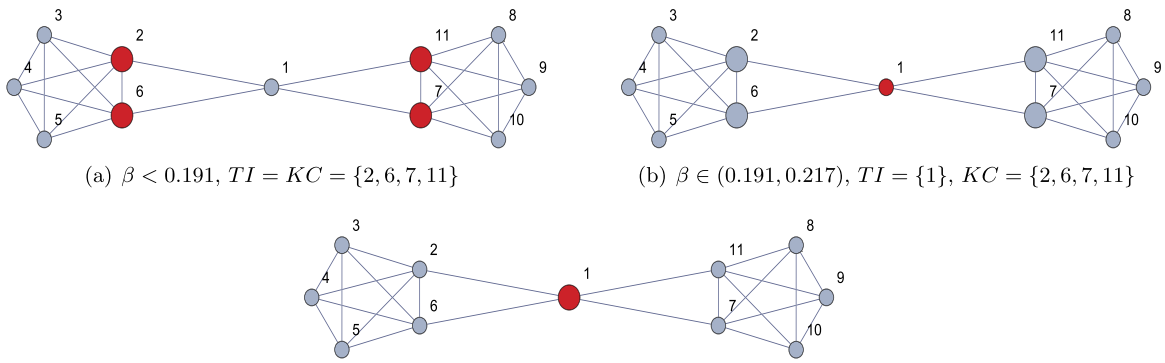
Table 2
Indexes of nodes in Fig. 6.

i	1	2	3
ρ_i	4	5	4
p_{i1}	0.303	0.334	0.278
δ_i	−0.035	−0.034	0.040
Δ_i	−2.230	−3.515	−5.251
$\underline{\Delta}_i$	−24.460	−19.744	−29.629
$\overline{\Delta}_i$	−1.492	−1.204	−1.807



(a) $c_i(\beta)(\beta_* = 0.191)$ (b) $h_i(\beta)(\beta^* = 0.217)$

Fig. 5. $c_i(\beta)$ and $h_i(\beta)$ for bridge network given in Fig. 6.



(a) $\beta < 0.191$, $TI = KC = \{2, 6, 7, 11\}$ (b) $\beta \in (0.191, 0.217)$, $TI = \{1\}$, $KC = \{2, 6, 7, 11\}$

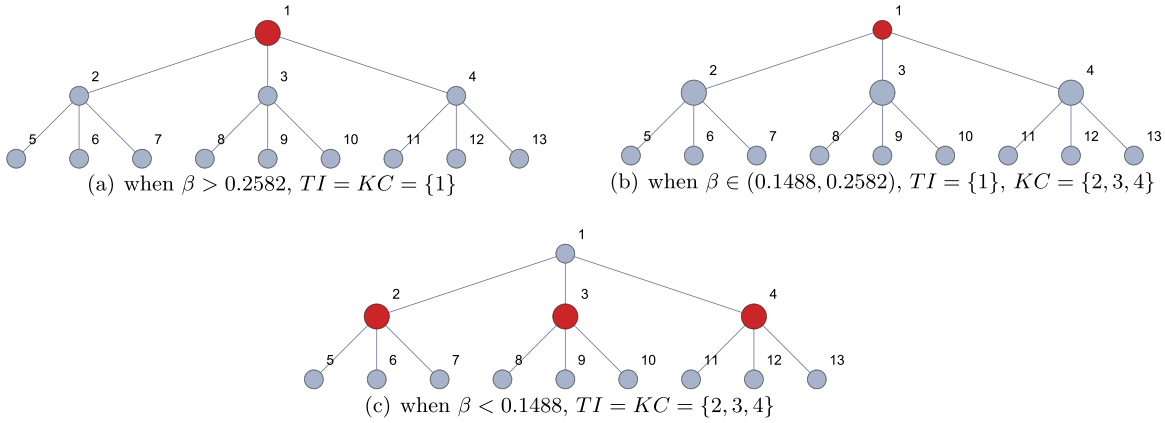
(c) $\beta > 0.217$, $TI = KC = \{1\}$

Fig. 6. TIs and KCs for different values of β.

Fig. 5 depicts c_i and h_i as functions of the decay factor β. It shows that when $\beta \in (\beta^*, 1/\lambda_1(\mathbf{G})] = (0.217, 0.227]$, type I agent (1) is both the TI and the KC; when $\beta \in (\beta_*, \beta^*) = (0.191, 0.217)$, 1 is the TI, while type II agents $(2, 6, 7, 11)$ are the KC; when $\beta \in (0, \beta_*) = (0, 0.191)$, type II nodes are both the TI and the KC. Fig. 6 marks the TI by red nodes, and the KC by nodes of large size, for different ranges of β.

Note that for small β ($\beta < \beta_*$), Type II agents having the largest vertex degree ($\rho_i = 5$) are both the TI and the KC. This validates the (i) part of Proposition 6. We see from Table 2 that $\mathbb{P}_1 = \{1\}, \mathbb{P}_2 = \{2, 6, 7, 11\}$, $\mathbb{Q}_1 = \{2, 6, 7, 11\}, \mathbb{Q}_2 = \{1\}$. It is easily verified that condition (29) holds, so TI and KC coincide with each other whenever $\beta \approx 1/\rho(\mathbf{G}) = 0.227$. This provides a justification for part (i) of Proposition 7. One can see, however, that part (ii) of Proposition 7 fails to hold in this example. For $\beta \approx 1/\rho(\mathbf{G})$, although TI and KC coincide, they are not the agents with the largest eigenvector centrality, i.e., $i^*, j^* \notin \mathbb{Q}_1 = \{2, 6, 7, 11\}$. We can check that condition (30) doesn't hold since $\underline{\Delta}_i = -19.744 \leqslant \overline{\Delta}_j = -1.492, \forall i \in \mathbb{Q}_1, j \in \mathbb{Q}_2$.

Example 4. Note that condition (30) is sufficient but not necessary for TI and KC to coincide with the node having the largest eigenvector centrality. See the three-tier tree given by Fig. 1, with its important indexes given in Table 3. It is



(a) when $\beta > 0.2582$, $TI = KC = \{1\}$

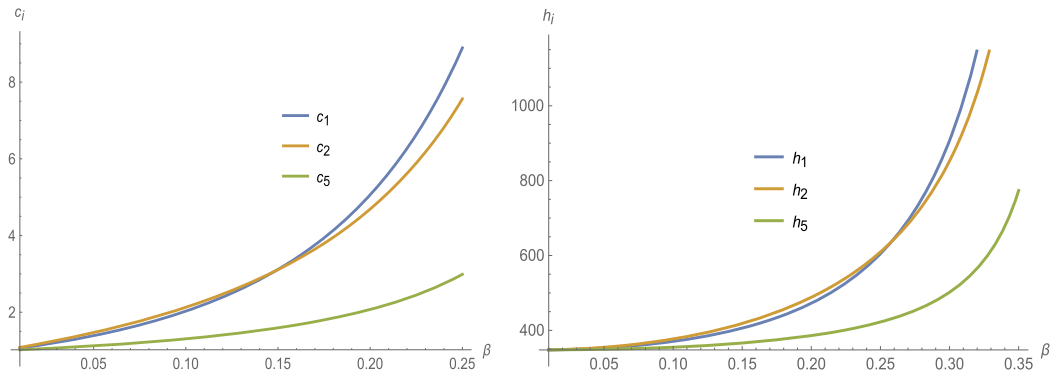
(b) when $\beta \in (0.1488, 0.2582)$, $TI = \{1\}$, $KC = \{2, 3, 4\}$

(c) when $\beta < 0.1488$, $TI = KC = \{2, 3, 4\}$

Fig. 7. TIs and KCs for different values of β.

Table 3
Indexes of nodes in Fig. 7.

i	1	2	5
ρ_i	3	4	1
p_{i1}	0.5	0.4082	0.1667
Δ_i	−1.0206	−3.4701	−17.3506
δ_i	−0.4082	−0.0981	0.6045
$\underline{\Delta}_i$	−4.1815	−6.9692	−48.7846
$\overline{\Delta}_i$	−0.6124	−1.0206	−7.1443



(a) $c_i(\beta)$

(b) $h_i(\beta)$

Fig. 8. $c_i(\beta)$ and $h_i(\beta)$ for a tree given by Fig. 1.

easily seen from Table 3 that $\mathbb{P}_1 = \mathbb{Q}_1 = \{1\}$, $\mathbb{P}_2 = \mathbb{Q}_2 = \{2, 3, 4\}$. Condition (29) holds since $\Delta_i + \underline{\delta} = -1.0206 - 0.4082 \geqslant -3.4701 + 0.6045 = \Delta_j + \overline{\delta}, \forall i \in \mathbb{P}_1, j \in \mathbb{P}_2$. Condition (30), however, fails since $\underline{\Delta}_i = -4.1815 < -1.0206 = \overline{\Delta}_j, \forall i \in \mathbb{Q}_1, j \in \mathbb{Q}_2$. Nevertheless, we find part (ii) of Proposition 7 still holds: when β is close to $1/\rho(\mathbf{G}) = 0.4082$, TI and KC both coincide with node 1. Fig. 8 depicts c_i and h_i as functions of β. For different ranges of β, Fig. 7 marks with, respectively, red and large-sized nodes, the TI and KC.

4. Identification of important edges

In addition to individual nodes, the seller also needs to pay attention to some important edges in a network. In the presence of strategic complements, the seller would always like to enhance communication among customers by adding new links. In some situations, however, he has to expand the network in an optimal way due to the limitation of resources. If he is allowed to add only one edge to the initial network, he will choose the *key link* among all void edges. Formally, we have the following definition.

Definition 1. The key link is a pair of nodes (i^*, j^*) belong to $\mathcal{K} \equiv \arg\max_{(i,j)\in\mathcal{G}_0}[\overline{\mathcal{V}}^*(\mathbf{G}^{+ij}) - \overline{\mathcal{V}}^*(\mathbf{G})]$, where $\mathbf{G}^{+ij} = \mathbf{G} + e_i e_j^\top + e_j e_i^\top$ is a new adjacency matrix obtained from $\mathbf{G}$ by adding ones to its (i, j) and (j, i) entries.

In the important nodes problem, the planner considers only the equilibrium and the corresponding efficiency. In contrast, when building a new link, she needs to consider the bilateral incentives of the individuals involved as well. In practical applications, the formation of a new edge usually requires unanimous consent from both parties. Say, a plan to open a new airline route between two cities needs to receive approval from both the municipal authorities.

We assume the link-building decision is made before the customers fully learn their own types. So it is the change of the ex-ante rents that determines their incentives for consenting the new link. From $d_i^* = \overline{\theta}, \forall i \in N$,[21] $\mathbb{E}\theta_{i0} = 0$, and

$$\overline{\vartheta}_i(\theta_{i0}, d_i^*) = \begin{cases} \theta_{i0} + \overline{\theta} & if \;\; \theta_{i0} \in \left[-\overline{\theta}, -\frac{(1+\varphi)\overline{\theta}}{1-\varphi}\right] \\ \varphi(\theta_{i0} - \overline{\theta}) & if \;\; \theta_{i0} \in \left(-\frac{(1+\varphi)\overline{\theta}}{1-\varphi}, \overline{\theta}\right] \end{cases},$$

we have

$$\begin{aligned}
\mathcal{U}'_{i0}(\theta_{i0}) &= \mathbb{E}_{\boldsymbol{\theta}_{-i0}} \sum_{t=0}^{T} \mathbb{E}_{\boldsymbol{\theta}_t} \bar{x}_{it}^*(\boldsymbol{\theta}_0, \boldsymbol{\theta}_t) = (T+1)\mathbb{E}_{\boldsymbol{\theta}_{-i0}} \bar{x}_{i0}^*(\boldsymbol{\theta}_0) \\
&= \frac{T+1}{1-Ts} e_i^\top (\mathbf{I} - \beta\mathbf{G})^{-1} \mathbb{E}_{\boldsymbol{\theta}_{-i0}} \left[\boldsymbol{\theta}_0 + \overline{\boldsymbol{\vartheta}}(\boldsymbol{\theta}_0, \mathbf{d}^*)\right] \\
&= \frac{T+1}{1-Ts} \left[\begin{array}{c} m_{ii}\left[\theta_{i0} + \overline{\vartheta}_i(\theta_{i0}, d_i^*) - \mathbb{E}\left(\theta_{i0} + \overline{\vartheta}_i(\theta_{i0}, d_i^*)\right)\right] \\ + e_i^\top (\mathbf{I} - \beta\mathbf{G})^{-1} \mathbb{E}\left[\boldsymbol{\theta}_0 + \overline{\vartheta}(\boldsymbol{\theta}_0, \mathbf{d}^*)\right] \end{array} \right] \\
&= \frac{T+1}{1-Ts} \left\{ m_{ii} \left[\theta_{i0} + \overline{\vartheta}_i(\theta_{i0}, d_i^*) - \left(\frac{-\varphi\overline{\theta}}{1-\varphi}\right) \right] + b_i \left(\frac{-\varphi\overline{\theta}}{1-\varphi}\right) \right\}.
\end{aligned}$$

Applying the integration by parts technique, we get the ex-ante information rent of agent i[22]:

$$\begin{aligned}
\mathscr{U}_i(\mathbf{G}) &\equiv \mathbb{E}[\mathcal{U}_{i0}(\theta_{i0})] = -\mathbb{E}\left[\overline{\vartheta}_i(\theta_{i0}, d_i^*)\mathcal{U}'_{i0}(\theta_{i0})\right] \\
&= -\frac{T+1}{1-Ts}\mathbb{E}\left\{ \left[m_{ii}\left[\theta_{i0} + \overline{\vartheta}_i(\theta_{i0}, d_i^*) + \frac{\varphi\overline{\theta}}{1-\varphi}\right] + b_i\left(\frac{-\varphi\overline{\theta}}{1-\varphi}\right)\right] \overline{\vartheta}_i(\theta_{i0}, d_i^*) \right\} \\
&= -\frac{(1+T)\varphi\overline{\theta}^2\left[(1+2\varphi)m_{ii} + 3\varphi b_i\right]}{3(1-Ts)(1-\varphi)^2}.
\end{aligned}$$

The marginal payoff accruing to agent i from a new link (i, j) is therefore

$$\begin{aligned}
\Delta\mathscr{U}_{ij}(\beta) &= \mathscr{U}_i(\mathbf{G}^{+ij}) - \mathscr{U}_i(\mathbf{G}) \\
&= -\frac{(1+T)\varphi\overline{\theta}^2}{3(1-Ts)(1-\varphi)^2} \left[\begin{array}{c} (1+2\varphi)e_i^\top\left[(\mathbf{I} - \beta\mathbf{G}^{+ij})^{-1} - (\mathbf{I} - \beta\mathbf{G})^{-1}\right]e_i \\ +3\varphi e_i^\top\left[(\mathbf{I} - \beta\mathbf{G}^{+ij})^{-1} - (\mathbf{I} - \beta\mathbf{G})^{-1}\right]\mathbb{1} \end{array} \right].
\end{aligned} \tag{31}$$

Since $1 + 2\varphi = -\frac{1+s-T+sT^2}{(1+T)(1-Ts+s)} > 0$, $3\varphi = \frac{-3(s+1)}{(1+T)(1-Ts+s)} < 0$, $\Delta\mathscr{U}_{ij}(\beta)$ may be either positive or negative, depending on the relative weights of $1+2\varphi$ and 3φ. If an individual has to build a new link with his potential neighbors, he will pick the one who benefits him the most or hurts him the least. If he can reject a detrimental edge incident to him and choose to stay unlinked, then only the beneficial neighbours falls in his choice set. Formally, we have the following definitions regarding the stability of a new link.

Definition 2. (i) A new link (i, j) is stable if $j \in \arg\max_{j' \in \mathcal{G}_0^i} \Delta\mathscr{U}_{ij'}$ and $i \in \arg\max_{i' \in \mathcal{G}_0^j} \Delta\mathscr{U}_{ji'}$, with at least one expression holds with equality. (ii) A new link (i, j) is strongly stable if it is stable and mutually beneficial, i.e., $\Delta\mathscr{U}_{ij} \geqslant 0$ and $\Delta\mathscr{U}_{ji} \geqslant 0$. Let $\mathcal{S}$ and $\overline{\mathcal{S}}$ denote, respectively, the sets of stable and strongly stable links.

A pair of agents constitutes a stable link if they mutually prefer each other the most, and at least one of them regard his partner as his unique best choice. A new link is strongly stable if it is stable and the linked agents are both beneficial to each other. A stable link would be unanimously agreed upon by both parties, and is immune to unilateral severance once formed. A strongly stable link can do so even if an individual is allowed to reject all detrimental links. The proposition below gives simple measures for the key and stable links.

[21] Similar results are obtained if we choose $d_i^* = \underline{\theta}, \forall i \in N$.

[22] See expressions (53) to (55) in Appendix B.

Proposition 8. *Based on the results of Proposition 1,*[23] *we have the following results regarding the key and stable links.*

- *A new link* $(i^*, j^*) \in \mathcal{K}$ *if it solves* $\max_{(i,j)\in\mathcal{G}_0} K_{ij}(\phi_1, \phi_2, \beta, \boldsymbol{G})$*, where*

$$K_{ij}(\phi_1, \phi_2, \beta, \boldsymbol{G}) \equiv \frac{\begin{bmatrix} \phi_1\left(2b_ib_j(1-\beta m_{ij}) + \beta m_{jj}b_i^2 + \beta m_{ii}b_j^2\right) + \\ \phi_2\left(2m_{ij}^{[2]}(1-\beta m_{ij}) + \beta m_{ii}m_{jj}^{[2]} + \beta m_{jj}m_{ii}^{[2]}\right) \end{bmatrix}}{\left(1-\beta m_{ij}\right)^2 - \beta^2 m_{ii}m_{jj}}, \tag{32}$$

  ϕ_1 *and* ϕ_2 *are defined by (23).*
- $(i^*, j^*) \in \mathcal{S}$ *if* $j^* \in \arg\max_{j\in\mathcal{G}_0^{i^*}} S_{i^*j}(\beta, \boldsymbol{G})$ *and* $i^* \in \arg\max_{i\in\mathcal{G}_0^{j^*}} S_{j^*i}(\beta, \boldsymbol{G})$*, with at least one expression holds with equality, where*

$$S_{ij}(\beta, \boldsymbol{G}) \equiv \frac{\begin{bmatrix} (1+2\varphi)\left(\beta m_{ii}(m_{ii}m_{jj} - m_{ij}^2) + 2m_{ii}m_{ij}\right) \\ +3\varphi\left(b_jm_{ii} + b_im_{ij} + \beta b_i(m_{ii}m_{jj} - m_{ij}^2)\right) \end{bmatrix}}{\left(1-\beta m_{ij}\right)^2 - \beta^2 m_{ii}m_{jj}}. \tag{33}$$

- $(i^*, j^*) \in \overline{\mathcal{S}}$ *if* $(i^*, j^*) \in \mathcal{S}$ *and* $S_{i^*j^*}(\beta, \boldsymbol{G}) \geqslant 0$, $S_{j^*i^*}(\beta, \boldsymbol{G}) \geqslant 0$.

Proof. See Appendix I. □

Expressions (32) and (33) have obvious advantages, as indexes c_i and h_i do in the identification of TI and KC. We can compute the contribution of an additional link from the current data, without having to recompute the inverse matrix $(\mathbf{I} - \beta\mathbf{G}^{+ij})^{-1}$ for each pair $(i, j) \in \mathcal{G}_0$.

In a connected network, a new link affects not only the payoffs of agents directly involved, but also that of the overall network. When deciding to build a new link, the seller would like to consider its impact on her aggregate profit obtained from the entire network. An individual buyer, however, aims only at his own payoff. So a wedge is driven between the objectives of the seller and the buyers. The key link is defined from the seller's optimality concern; a stable link is, however, from the consideration of individual incentives. Hence, they don't necessarily coincide with each other. Although their relationship is in general ambiguous, we find some clear results in an extreme case where β approaches zero.

Proposition 9. *Suppose that* $|\mathcal{G}_0^i| \geqslant 2, |\mathcal{G}_0^j| \geqslant 2, \forall(i, j) \in \mathcal{G}_0$*,* β *is sufficiently close to zero, then* $\mathcal{K} \cap \mathcal{S} = \emptyset$*, where* $|\cdot|$ *denotes cardinality of a set.*

Proof. See Appendix J. □

Condition $|\mathcal{G}_0^i| \geqslant 2, |\mathcal{G}_0^j| \geqslant 2, \forall(i, j) \in \mathcal{G}_0$ rules out the trivial case in which a pair of agents have no choice but to build a new link with each other. Consider, for example, a complete network with but one edge (i, j) void: $K_n \backslash \{(i, j)\}$. It is clear that the void edge is both a key and a stable link, i.e., $(i, j) \in \mathcal{K} \cap \mathcal{S}$.

Fig. 9 shows two possible channels (denoted by, respectively, green and red arrows) through which a new edge (i, j) affects the ex-ante rent of agent i. ① to ⑦ denote the local effects through which the final impact works successively. All these effects are evaluated from the ex-ante perspective. We denote by $\mathbb{E}_0[\cdot]$ the expectation operator over information available ex-ante, then these local effects are calculated as follows: ① $= \mathbb{E}_0[\bar{x}_{i0}(\boldsymbol{\theta}_0, \mathbf{G}^{+ij})] - \mathbb{E}_0[\bar{x}_{i0}(\boldsymbol{\theta}_0, \mathbf{G})] = \frac{-\varphi\overline{\theta}[b_i(\beta,\mathbf{G}^{ij}) - b_i(\beta,\mathbf{G})]}{(1-Ts)(1-\varphi)} > 0$, ② $= \mathbb{E}_0[\partial\theta_{i1}/\partial\bar{x}_{i0}] = s < 0$, ③ $= \mathbb{E}_0[\partial\mathcal{U}_{i1}/\partial\theta_{i1}] = \mathbb{E}_0\left[\sum_{s\geqslant 1}\bar{x}_{is}(\boldsymbol{\theta}_0, \boldsymbol{\theta}_s)\right] = \frac{-T\varphi\overline{\theta}b_i}{(1-Ts)(1-\varphi)} > 0$, ④ $= \mathbb{E}_0[\partial\theta_{j1}/\partial\bar{x}_{i0}] = -\beta s > 0$, ⑤ $= \mathbb{E}_0[\bar{x}_{j1}/\partial\theta_{j1}] = \frac{m_{ii}}{1-(T-t)s} > 0$, ⑥ $= \mathbb{E}_0[\partial\mathcal{U}_{i1}/\partial\bar{x}_{j1}] = \frac{\beta}{2}g_{ij}\mathbb{E}_0\bar{x}_{1t} = \frac{-\beta g_{ij}\varphi\overline{\theta}b_i}{2(1-Ts)(1-\varphi)} > 0$, effect ⑦ is obviously positive since $\mathcal{U}_{i0}(\theta_{i0}) = \mathbb{E}_{\boldsymbol{\theta}_{-i0}}u_{i0}(\boldsymbol{\theta}_0) + \mathbb{E}_{\boldsymbol{\theta}_{-i0},\theta_{i1}}\mathcal{U}_{i1}(\boldsymbol{\theta}_0, \theta_{i1})$.[24]

On one hand, the tolerance effect (②) weakens i's future preference, and thus a given level of quality becomes less satisfying in the future than at present. An agent hence obtains a lower continuation utility. On the other hand, the future preferences of i's neighbors are strengthened by the change of his own initial consumption due to the bandwagon effect (④). This increases i's future utility obtained from interactions via the network (effect ⑥). When $\beta \approx 0$, the strength of bandwagon and network effects are both negligibly small, therefore the first channel dominates the second one. As a result, an additional link incident to an agent brings to him more losses than gains. Approximating ① around zero yields ①

[23] Suppose that all assumptions in Proposition 1 are met, except that we now require a smaller β ($\beta < 1/\rho(\mathbf{G}^{+ij})$) to guarantee $(\mathbf{I} - \beta\mathbf{G}^{+ij})^{-1} = \sum_{k=0}^{\infty}\beta^k(\mathbf{G}^{+ij})^k$.

[24] Here we use $d_i^* = \overline{\theta}$, analogous results can be obtained for the case $d_i^* = \underline{\theta}$.

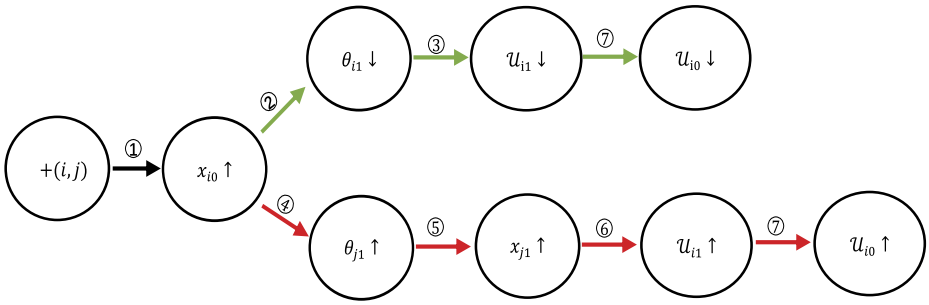
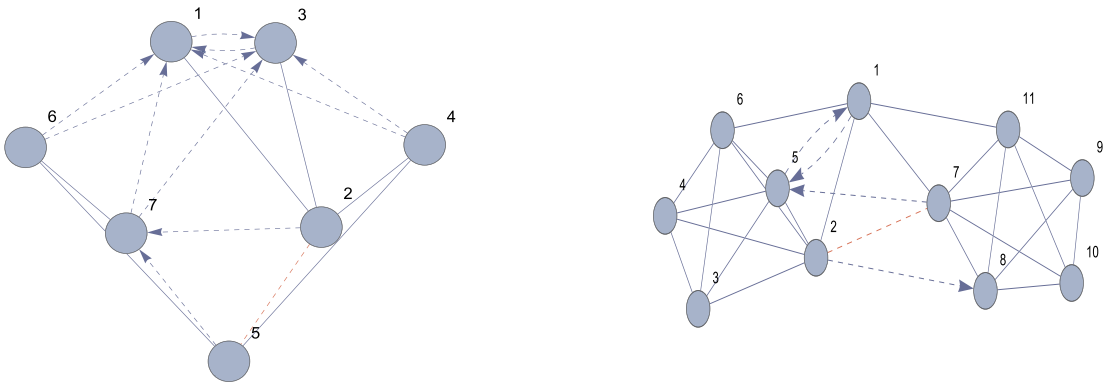


Fig. 9. The impact of new link (i, j) on rent of agent i.



(a) $\beta = 0.001$ (b) $\beta = 0.0001$

Fig. 10. $\mathcal{K} \cap \mathcal{S} = \emptyset$ for small β.

$= -\frac{\varphi\overline{\theta}\beta}{(1-Ts)(1-\varphi)}\left[1 + (1+\rho_j)\beta + o(\beta)\right] > 0$. It means that a new neighbor with a larger vertex degree brings higher loss. So an individual will choose among all potential neighbors the one having the lowest degree.

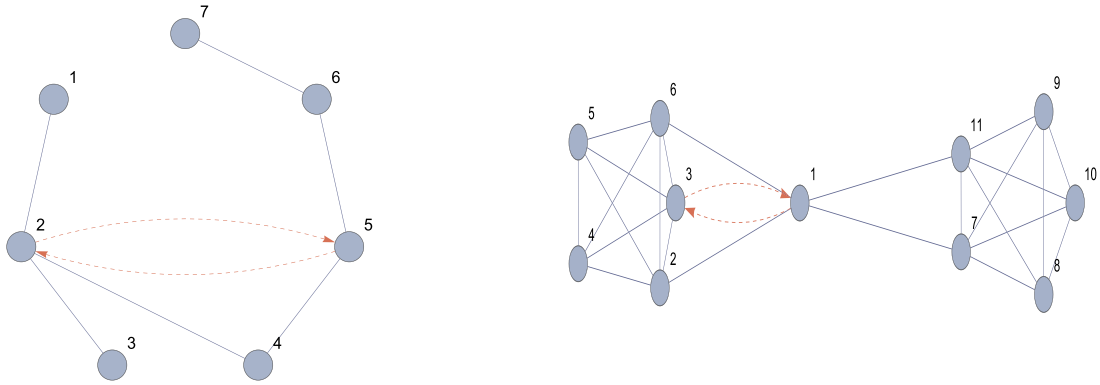
The seller's criterion for selecting a new edge is in sharp contrast to that of individual buyers. Among all void edges, the seller wish to bridge the pair of nodes with the largest vertex degrees, which hurts each other the most and thus helps the seller to save information rents by the largest margin. This leads to the incompatibility of the key and the stable links. We now illustrate the above result by some numerical examples.

Example 5. Panel (a) of Fig. 10 gives a seven-nodes network. Suppose $T = 10, \beta = 0.001, s = -0.5, \overline{\theta} = 10, \sigma_t = 1, \forall t \in \mathcal{T}\backslash\{T\}$. Solid lines denote the existing edges, dashed arrows indicate the direction of one's favorite neighbor, the red dashed line denotes the *key link*. The seller would like to build a link between a pair of highest-degree nodes 2 and 5 ($\rho_2 = 3, \rho_5 = 2$).[25] However, both 2 and 5 prefer to connect to the low-degree node 7.[26] Panel (b) depicts the bridge network given in Fig. 6, with $T = 10, \beta = 0.0001, s = -0.5, \overline{\theta} = 10, \sigma_t = 1, \forall t \in \mathcal{T}\backslash\{T\}$. It is optimal for the seller to build a new link between a pair of type II nodes, because they have the largest degree 5. That is, $\mathcal{K} = \{(2, 7), (2, 11), (6, 7), (6, 11)\}$. But all the key links are unstable, because the agents involved both prefer to connect to a low-degree type III node, rather than with each other. Notice that this result depends crucially on β being close enough to zero. We may have $\mathcal{K} \cap \mathcal{S} \neq \emptyset$ when β is not too small. Fig. 11 depicts the compatibility of the key and stable links for the same networks as Fig. 10, and the same parameters $T, s, \overline{\theta}, \sigma_t$, but larger β ($\beta = 0.4$ for panel (a) and $\beta = 0.15$ for panel (b)). In panel (a), $\mathcal{K} \cap \mathcal{S} = \{(2, 5)\}$; in panel (b), a link between node 1 and an arbitrary type$-III$ node is both the key and stable link, i.e., $\mathcal{K} \cap \mathcal{S} = \{(1, j) | j \in \{3, 4, 5, 8, 9, 10\}\}$.

Our analysis above shows that an additional link incident to a consumer is detrimental to him when β is very small. If a consumer reserves the right to reject all harmful links and stay unlinked, we got a more negative result.

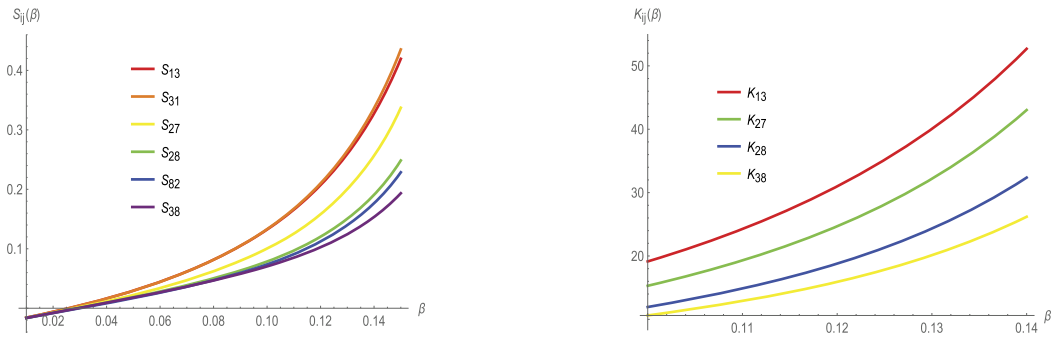
[25] Although 5 and 6 have the same degree 2, the higher order term favors 5. Note that the degree sum in 5's neighborhood is larger than that of 6, i.e., $e_5^\top \mathbf{G}^2 \mathbb{1} = \sum_{j\neq 5} g_{5j}\rho_j = 4 > e_6^\top \mathbf{G}^2 \mathbb{1} = \sum_{j\neq 6} g_{6j}\rho_j = 3$; and node 5 has a shorter two-step distance from 2 than does node 6, i.e., $g_{25}^{[2]} = 1 > g_{26}^{[2]} = 0$. So expression (101) suggests (2, 5), rather than (2, 6), is the key link.

[26] Agents 1,3 and 7 are all available to agent 5. The identical degree $\rho_1 = \rho_3 = \rho_7 = 1$ fails to provide a clear ranking among them. We have to resort to the higher-order term in $S_{ij}(\beta, \mathbf{G})$ (102) to discriminate among them. Because $3\varphi = -0.05 < 0, 7\varphi + 2 = 1.88 > 0$, $e_1^\top \mathbf{G}^2 \mathbb{1} = e_3^\top \mathbf{G}^2 \mathbb{1} = 3 > e_7^\top \mathbf{G}^2 \mathbb{1} = 2$, $g_{51}^{[2]} = g_{53}^{[2]} = 0 < g_{57}^{[2]} = 1$, agent 5 prefers 7 to 1 and 3.



(a) $\beta = 0.4$ (b) $\beta = 0.15$

Fig. 11. $\mathcal{K} \cap \mathcal{S} \neq \emptyset$ for large β.



(a) S_{ij} (b) K_{ij}

Fig. 12. $\overline{\mathcal{S}} = \emptyset$ for small β; $\overline{\mathcal{S}} \cap \mathcal{K} = \{(1, j) | j \in \{3, 4, 5, 8, 9, 10\}\}$ for large β.

Proposition 10. *If β is sufficiently small, we have $S_{ij} < 0, \forall (i, j) \in \mathcal{G}^0$, then there exists no strongly stable link, i.e., $\overline{\mathcal{S}} = \emptyset$.*

Example 6. Fig. 12 depicts $\{S_{ij}\}_{(i,j)\in\mathcal{G}^0}$ and $\{K_{ij}\}_{(i,j)\in\mathcal{G}^0}$ as functions of β for the eleven-nodes bridge network given in Fig. 6 with parameters $T = 10, s = -0.5, \overline{\theta} = 10, \sigma_t = 1, \forall t \in \mathcal{T}\backslash\{T\}$. It is obvious that $S_{ij} < 0$ for any pair $(i, j) \in \mathcal{G}^0$ as long as β is very small, so $\overline{\mathcal{S}} = \emptyset$ for small β. If β is not too small, Fig. 12(a) shows that $\overline{\mathcal{S}} = \{(i, j) | i \in \{2, 6\}, j \in \{7, 11\}\} \cup \{(1, j) | j \in \{3, 4, 5, 8, 9, 10\}\}$; Fig. 12(b) shows that $\mathcal{K} = \{(1, j) | j \in \{3, 4, 5, 8, 9, 10\}$. Therefore, $\overline{\mathcal{S}} \cap \mathcal{K} = \{(1, j) | j \in \{3, 4, 5, 8, 9, 10\}$.

5. The optimal network intervention

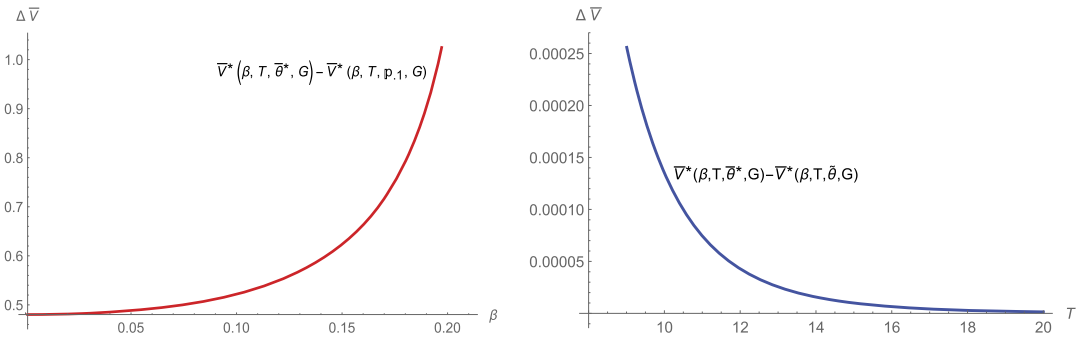
In this section, we proceed to discuss the problem of optimal network intervention based on the results given in Proposition 1. Aiming to maximize her own profit, the seller can conduct an ex-ante intervention, at a cost, on the distributions of initial types. Say, she can intervene to change $\overline{\theta}_i$s via offering free trial to, or advertising towards, the target customers. The crucial problem for the seller is to allocate optimally her fixed budget across different individuals. We assume that intervention cost paid by the seller is of quadratic form, it is separable across individuals and is increasing in the magnitude of individual uncertainty. Without loss of generality, the total budget for intervention is normalized to one. The seller's budget constraint is therefore $\|\overline{\boldsymbol{\theta}}\| \leqslant 1$, his network intervention problem $[\mathcal{NI}]$ is represented as:

$$[\mathcal{NI}] : \max_{\overline{\boldsymbol{\theta}} \in \mathbb{R}^n_+} \overline{\mathcal{V}}^*(T, \beta, \overline{\boldsymbol{\theta}}, \mathbf{G}), s.t. : \|\overline{\boldsymbol{\theta}}\| \leqslant 1.$$

We have the following results.

Proposition 11. *The optimal intervention policy is $\overline{\boldsymbol{\theta}}^* = \boldsymbol{v}_1(\boldsymbol{\Sigma})$, where $\boldsymbol{v}_1(\cdot)$ denotes the normalized principal eigenvector of a matrix, $\boldsymbol{\Sigma} \equiv \varphi^2 \boldsymbol{M} + (1 + 4\varphi^2)/3\widehat{\boldsymbol{M}}$.*

- *As $\beta \to 0, \overline{\boldsymbol{\theta}}^* \to \boldsymbol{p}_{.1} \equiv \boldsymbol{v}_1(\boldsymbol{G})$;*



(a) $T = 20, s = -0.2$ (b) $\beta = 0.2, s = -0.2$

Fig. 13. Profit loss of approximation as functions of β (a) and T (b).

- *as* $T \to \infty, \overline{\boldsymbol{\theta}}^* \to \tilde{\boldsymbol{\theta}} \equiv (\boldsymbol{v}_1(\boldsymbol{M}_{K,K}), \boldsymbol{0}_{-K}^\top)^\top$,

where $K \equiv \arg\max_{i \in N} m_{ii}(\beta, \boldsymbol{G})$ *denotes the set of nodes with the largest resolvent subgraph centrality,* $\boldsymbol{M}_{K,K}$ *is the submatrix of* $\boldsymbol{M}$ *composed of rows and columns indexed by* $k \in K$, $\boldsymbol{0}_{-K}$ *is a vector of zeros indexed by* $k \in N \backslash K$.

Proof. See Appendix K. □

The optimal intervention takes simple forms in the case of small β or large T. When β is very small, it tends to the first principal component of $\mathbf{G}$. When T is sufficiently large, all budget is allocated to the agents with the largest resolvent subgraph centrality. If there exist more than one such agents, the budget is further allocated among them according to the first principal component of $\mathbf{M}_{K,K}$. Such simple interventions are practically easy to implement because they depend only on the network structure and not on any other parameters. The final result in this section characterizes how fast the profit loss of approximation vanishes as $\beta \to 0$ or $T \to \infty$, and thus shows the accuracy of the approximations.

Proposition 12. *The converging orders of profit loss under approximations* $p_{.1}$ *and* $\tilde{\boldsymbol{\theta}}$ *are:*

$$\overline{\mathcal{V}}^*(T, \beta, \overline{\boldsymbol{\theta}}^*, \boldsymbol{G}) - \overline{\mathcal{V}}^*(T, \beta, \boldsymbol{p}_{.1}, \boldsymbol{G}) = O(\beta^2)$$

and

$$\overline{\mathcal{V}}^*(T, \beta, \overline{\boldsymbol{\theta}}^*, \boldsymbol{G}) - \overline{\mathcal{V}}^*(T, \beta, \tilde{\boldsymbol{\theta}}, \boldsymbol{G}) = O\left(\frac{1}{T^4}\right).$$

Proof. See Appendix L. □

For the bridge network given in Fig. 6 and parameters $T = 20, s = -0.2, \beta = 0.2$, we have $(p_{i1}, p_{i2}, p_{i3}) = (0.303, 0.334, 0.278)$, $\lambda_1(\mathbf{M}) = 8.390$, $K = \{2, 6, 7, 11\}$, $m^* \equiv \max_{i \in N} m_{ii} = 2.083$, $\mathbf{v}_1(\mathbf{M}_{K,K}) = (0.5, 0.5, 0.5, 0.5)^\top$, $\lambda_1(\mathbf{M}_{K,K}) = 4.167$. The following Fig. 13 depicts the profit losses as functions of β (resp. T) when we use $\mathbf{p}_{.1}$ (resp. $\tilde{\boldsymbol{\theta}}$) to approximate the optimal intervention $\overline{\boldsymbol{\theta}}^*$. It can be found that the accuracies of these approximations are very high for sufficiently small β or large T.

6. Conclusion

This paper studies a sequential nonlinear pricing problem where buyers interact on a social network and their preferences evolve stochastically over time. It bridges two bodies of research on dynamic mechanism design and social network analysis. We establish three main results.

First, we find that the traditional first-order approach fails to provide a dynamically implementable mechanism. As a remedy, we perform a slope ironing procedure on the relaxed mechanism. The optimal mechanism obtained is therefore perfectly sorting and admits no bunching. Second, we identify and compare some important nodes and edges of a network within the mechanism design framework. It is shown that the key customer, who impacts the principal's profit the most, and the top influencer, who has the greatest influence on the aggregate equilibrium activity, coincide with each other when either the strength of network complementarity β is very small or large. We also find that the key link, which is the most profit-enhancing, and the stable link, which is mutually most beneficial, are incompatible with each other when β is small. It is therefore impossible for the designer to add to the initial network a new link in an incentive compatible manner. Finally, we discuss the network intervention problem where the principal could invest to change the distribution of the

agents' initial types. We show that when the synergy parameter is sufficiently small or time horizon is sufficiently large, the optimal intervention is approximately simple in the sense that it depends only on the network structure: (i) when β is small, the optimal intervention policy is approximately proportional to the first principal component of network; (ii) when the time horizon is sufficiently large, the total intervening budget is allocated among the nodes with the largest resolvent subgraph centrality, and is further allocated according to the first principal component of a submatrix. Moreover, we show that the profit losses of these simple approximations both vanish at high orders, so that their accuracies are sufficiently high.

Appendix A. Proof of Lemma 1

- $IC_{it} \Rightarrow EN_{it} + IM_{it}$. Applying envelope theorem to

$$IC_{it} : \mathcal{U}_{it}(\widehat{\boldsymbol{\Theta}}_{t-1}, \theta_{it}) = \max_{\hat{\theta}_{it}} \widetilde{\mathcal{U}}_{it}(\widehat{\boldsymbol{\Theta}}_{t-1}, \theta_{it}, \hat{\theta}_{it})$$

yields

$$\begin{aligned}\frac{\partial \mathcal{U}_{it}(\widehat{\boldsymbol{\Theta}}_{t-1}, \theta_{it})}{\partial \theta_{it}} &= \frac{\partial \widetilde{\mathcal{U}}_{it}(\widehat{\boldsymbol{\Theta}}_{t-1}, \theta_{it}, \theta_{it})}{\partial \theta_{it}} \\ &= \left[\begin{array}{c} \mathbb{E}_{\boldsymbol{\theta}_{-it}} x_{it}(\widehat{\boldsymbol{\Theta}}_{t-1}, \boldsymbol{\theta}_t) \\ + \frac{\partial}{\partial \theta_{it}} \mathbb{E}_{\boldsymbol{\theta}_{-it}, \theta_{it+1}} \left[\mathcal{U}_{it+1}(\widehat{\boldsymbol{\Theta}}_{t-1}, \boldsymbol{\theta}_t, \theta_{it+1}) \middle| \mathcal{I}_{it} \right] \end{array} \right]. \end{aligned} \tag{34}$$

Differentiating $\mathbb{E}_{\boldsymbol{\theta}_{-it}} \mathbb{E}_{\theta_{it+1}} [\mathcal{U}_{it+1}(\widehat{\boldsymbol{\Theta}}_{t-1}, \hat{\theta}_{it}, \boldsymbol{\theta}_{-it}, \theta_{it+1}) | \widehat{\mathcal{I}}_{it}]$ with respect to the true type θ_{it} yields

$$\frac{\partial}{\partial \theta_{it}} \mathbb{E}_{\boldsymbol{\theta}_{-it}} \mathbb{E}_{\theta_{it+1}} \left[\mathcal{U}_{it+1}(\widehat{\boldsymbol{\Theta}}_{t-1}, \hat{\theta}_{it}, \boldsymbol{\theta}_{-it}, \theta_{it+1}) \middle| \widehat{\mathcal{I}}_{it} \right] \tag{35}$$

$$= \frac{\partial}{\partial \theta_{it}} \mathbb{E}_{\boldsymbol{\theta}_{-it}} \int_{-\infty}^{+\infty} \mathcal{U}_{it+1}(\widehat{\boldsymbol{\Theta}}_{t-1}, \hat{\theta}_{it}, \boldsymbol{\theta}_{-it}, \theta_{it+1}) g_t \left(\begin{array}{c} \theta_{it+1} - \theta_{it} \\ -\mathbf{a}_i^\top \mathbf{x}_t(\widehat{\boldsymbol{\Theta}}_{t-1}, \hat{\theta}_{it}, \boldsymbol{\theta}_{-it}) \end{array} \right) d\theta_{it+1} \tag{36}$$

$$= -\mathbb{E}_{\boldsymbol{\theta}_{-it}} \int_{-\infty}^{+\infty} \mathcal{U}_{it+1}(\widehat{\boldsymbol{\Theta}}_{t-1}, \hat{\theta}_{it}, \boldsymbol{\theta}_{-it}, \theta_{it+1}) g_t' \left(\begin{array}{c} \theta_{it+1} - \theta_{it} \\ -\mathbf{a}_i^\top \mathbf{x}_t(\widehat{\boldsymbol{\Theta}}_{t-1}, \hat{\theta}_{it}, \boldsymbol{\theta}_{-it}) \end{array} \right) d\theta_{it+1} \tag{37}$$

$$= \mathbb{E}_{\boldsymbol{\theta}_{-it}} \left[\begin{array}{c} \int_{-\infty}^{+\infty} g_t \left(\begin{array}{c} \theta_{it+1} - \theta_{it} \\ -\mathbf{a}_i^\top \mathbf{x}_t(\widehat{\boldsymbol{\Theta}}_{t-1}, \hat{\theta}_{it}, \boldsymbol{\theta}_{-it}) \end{array} \right) \frac{\partial \mathcal{U}_{it+1}(\widehat{\boldsymbol{\Theta}}_{t-1}, \hat{\theta}_{it}, \boldsymbol{\theta}_{-it}, \theta_{it+1})}{\partial \theta_{it+1}} d\theta_{it+1} \\ -\mathcal{U}_{it+1}(\widehat{\boldsymbol{\Theta}}_{t-1}, \hat{\theta}_{it}, \boldsymbol{\theta}_{-it}, \theta_{it+1}) g_t \left(\begin{array}{c} \theta_{it+1} - \theta_{it} \\ -\mathbf{a}_i^\top \mathbf{x}_t(\widehat{\boldsymbol{\Theta}}_{t-1}, \hat{\theta}_{it}, \boldsymbol{\theta}_{-it}) \end{array} \right) \Bigg|_{-\infty}^{+\infty} \end{array} \right] \tag{38}$$

$$= \mathbb{E}_{\boldsymbol{\theta}_{-it}} \mathbb{E}_{\theta_{it+1}} \left[\frac{\partial \mathcal{U}_{it+1}(\widehat{\boldsymbol{\Theta}}_{t-1}, \hat{\theta}_{it}, \boldsymbol{\theta}_{-it}, \theta_{it+1})}{\partial \theta_{it+1}} \middle| \widehat{\mathcal{I}}_{it} \right]. \tag{39}$$

(35) $\Rightarrow$ (36) follows from the type-evolving equation (4); (37) $\Rightarrow$ (38) follows from integration by part; (38) $\Rightarrow$ (39) follows from property of density function: $g_t(-\infty) = g_t(\infty) = 0$. Evaluating (39) at $\hat{\theta}_{it} = \theta_{it}$, then substituting it into (34) yields:

$$\frac{\partial \mathcal{U}_{it}(\widehat{\boldsymbol{\Theta}}_{t-1}, \theta_{it})}{\partial \theta_{it}} = \mathbb{E}_{\boldsymbol{\theta}_{-it}} x_{it}(\widehat{\boldsymbol{\Theta}}_{t-1}, \boldsymbol{\theta}_t) + \mathbb{E}_{\boldsymbol{\theta}_{-it}} \mathbb{E}_{\theta_{it+1}} \left[\frac{\partial \mathcal{U}_{it+1}(\widehat{\boldsymbol{\Theta}}_{t-1}, \boldsymbol{\theta}_t, \theta_{it+1})}{\partial \theta_{it+1}} \middle| \mathcal{I}_{it} \right]. \tag{40}$$

Applying (39) recursively, we get the following envelope condition EN_{it}:

$$\begin{aligned} &\frac{\partial \mathcal{U}_{it}(\widehat{\boldsymbol{\Theta}}_{t-1}, \theta_{it})}{\partial \theta_{it}} \\ &= \mathbb{E}_{\boldsymbol{\theta}_{-it}} x_{it}(\widehat{\boldsymbol{\Theta}}_{t-1}, \theta_{it}, \boldsymbol{\theta}_{-it}) \\ &\quad + \mathbb{E}_{\boldsymbol{\theta}_{-it}} \mathbb{E}_{\theta_{it+1}} \left[\begin{array}{c} \mathbb{E}_{\boldsymbol{\theta}_{-it+1}} x_{it+1}(\widehat{\boldsymbol{\Theta}}_{t-1}, \boldsymbol{\theta}_t, \boldsymbol{\theta}_{t+1}) + \\ \mathbb{E}_{\boldsymbol{\theta}_{-it+1}} \mathbb{E}_{\theta_{it+2}} \left(\frac{\partial \mathcal{U}_{it+2}(\widehat{\boldsymbol{\Theta}}_{t-1}, \boldsymbol{\theta}_t, \boldsymbol{\theta}_{t+1}, \theta_{it+2})}{\partial \theta_{it+2}} \middle| \mathcal{I}_{it+1} \right) \end{array} \middle| \mathcal{I}_{it} \right] \\ &= \mathbb{E}_{\boldsymbol{\theta}_{-it}} x_{it}(\widehat{\boldsymbol{\Theta}}_{t-1}, \theta_{it}, \boldsymbol{\theta}_{-it}) + \mathbb{E}_{\boldsymbol{\theta}_{-it}} \mathbb{E}_{\boldsymbol{\theta}_{t+1}} \left[x_{it+1}(\widehat{\boldsymbol{\Theta}}_{t-1}, \boldsymbol{\theta}_t, \boldsymbol{\theta}_{t+1}) \middle| \mathcal{I}_{it} \right] \end{aligned} \tag{41}$$

$$+\mathbb{E}_{\boldsymbol{\theta}_{-it}}\mathbb{E}_{\boldsymbol{\theta}_{t+1}}\mathbb{E}_{\theta_{it+2}}\left[\frac{\partial \mathcal{U}_{it+2}(\widehat{\boldsymbol{\Theta}}_{t-1},\boldsymbol{\theta}_t,\boldsymbol{\theta}_{t+1},\theta_{it+2})}{\partial\theta_{it+2}}\middle|\mathcal{I}_{it}\right] \tag{42}$$

$$=\cdots=\mathbb{E}_{\boldsymbol{\theta}_{-it},\{\boldsymbol{\theta}_\tau\}_{\tau\geqslant t+1}}\left[\sum_{s\geqslant t}x_{is}(\widehat{\boldsymbol{\Theta}}_{t-1},\boldsymbol{\theta}_t,\cdots,\boldsymbol{\theta}_s)\middle|\mathcal{I}_{it}\right]. \tag{43}$$

(41) $\Rightarrow$ (42) is implied by set inclusion $\mathcal{I}_{it}\subset\mathcal{I}_{it+1}$, and the law of iterated expectations.[27]
The first and second order conditions for IC_{it} are[28]:

$$\begin{aligned}\frac{\partial}{\partial\hat{\theta}_{it}}\widetilde{\mathcal{U}}_{it}(\widehat{\boldsymbol{\Theta}}_{t-1},\theta_{it},\theta_{it})=&\,\theta_{it}\frac{\partial}{\partial\hat{\theta}_{it}}\mathbb{E}_{\boldsymbol{\theta}_{-it}}x_{it}(\widehat{\boldsymbol{\Theta}}_{t-1},\theta_{it},\boldsymbol{\theta}_{-it})-\frac{1}{2}\frac{\partial}{\partial\hat{\theta}_{it}}\mathbb{E}_{\boldsymbol{\theta}_{-it}}\left[x_{it}^2(\widehat{\boldsymbol{\Theta}}_{t-1},\theta_{it},\boldsymbol{\theta}_{-it})\right]\\&+\frac{\beta}{2}\frac{\partial}{\partial\hat{\theta}_{it}}\sum_{j\neq i}g_{ij}\mathbb{E}_{\boldsymbol{\theta}_{-it}}\left[x_{it}(\widehat{\boldsymbol{\Theta}}_{t-1},\theta_{it},\boldsymbol{\theta}_{-it})x_{jt}(\widehat{\boldsymbol{\Theta}}_{t-1},\theta_{it},\boldsymbol{\theta}_{-it})\right]\\&-\frac{\partial}{\partial\hat{\theta}_{it}}\mathbb{E}_{\boldsymbol{\theta}_{-it}}\left[\tau_{it}(\widehat{\boldsymbol{\Theta}}_{t-1},\theta_{it},\boldsymbol{\theta}_{-it})\right]\\&+\frac{\partial}{\partial\hat{\theta}_{it}}\mathbb{E}_{\boldsymbol{\theta}_{-it},\theta_{it+1}}\left[\mathcal{U}_{it+1}(\widehat{\boldsymbol{\Theta}}_{t-1},\theta_{it},\boldsymbol{\theta}_{-it},\theta_{it+1})\middle|\mathcal{I}_{it}\right]=0,\end{aligned} \tag{44}$$

and

$$\begin{aligned}\frac{\partial^2}{\partial\hat{\theta}_{it}^2}\widetilde{\mathcal{U}}_{it}(\widehat{\boldsymbol{\Theta}}_{t-1},\theta_{it},\theta_{it})=&\,\theta_{it}\frac{\partial^2}{\partial\hat{\theta}_{it}^2}\mathbb{E}_{\boldsymbol{\theta}_{-it}}x_{it}(\widehat{\boldsymbol{\Theta}}_{t-1},\theta_{it},\boldsymbol{\theta}_{-it})-\frac{1}{2}\frac{\partial^2}{\partial\hat{\theta}_{it}^2}\mathbb{E}_{\boldsymbol{\theta}_{-it}}\left[x_{it}^2(\widehat{\boldsymbol{\Theta}}_{t-1},\theta_{it},\boldsymbol{\theta}_{-it})\right]\\&+\frac{\beta}{2}\frac{\partial^2}{\partial\hat{\theta}_{it}^2}\sum_{j\neq i}g_{ij}\mathbb{E}_{\boldsymbol{\theta}_{-it}}\left[x_{it}(\widehat{\boldsymbol{\Theta}}_{t-1},\theta_{it},\boldsymbol{\theta}_{-it})x_{jt}(\widehat{\boldsymbol{\Theta}}_{t-1},\theta_{it},\boldsymbol{\theta}_{-it})\right]\\&-\frac{\partial^2}{\partial\hat{\theta}_{it}^2}\mathbb{E}_{\boldsymbol{\theta}_{-it}}\left[\tau_{it}(\widehat{\boldsymbol{\Theta}}_{t-1},\theta_{it},\boldsymbol{\theta}_{-it})\right]\\&+\frac{\partial^2}{\partial\hat{\theta}_{it}^2}\mathbb{E}_{\boldsymbol{\theta}_{-it},\theta_{it+1}}\left[\mathcal{U}_{it+1}(\widehat{\boldsymbol{\Theta}}_{t-1},\theta_{it},\boldsymbol{\theta}_{-it},\theta_{it+1})\middle|\mathcal{I}_{it}\right]<0.\end{aligned} \tag{45}$$

Differentiating (44) with respect to θ_{it}, we get

$$\begin{aligned}\frac{\partial^2\widetilde{\mathcal{U}}_{it}(\widehat{\boldsymbol{\Theta}}_{t-1},\theta_{it},\theta_{it})}{\partial\hat{\theta}_{it}\partial\theta_{it}}=&\,\theta_{it}\frac{\partial^2}{\partial\hat{\theta}_{it}^2}\mathbb{E}_{\boldsymbol{\theta}_{-it}}x_{it}(\widehat{\boldsymbol{\Theta}}_{t-1},\theta_{it},\boldsymbol{\theta}_{-it})-\frac{1}{2}\frac{\partial^2}{\partial\hat{\theta}_{it}^2}\mathbb{E}_{\boldsymbol{\theta}_{-it}}\left[x_{it}^2(\widehat{\boldsymbol{\Theta}}_{t-1},\theta_{it},\boldsymbol{\theta}_{-it})\right]\\&+\frac{\beta}{2}\frac{\partial^2}{\partial\hat{\theta}_{it}^2}\sum_{j\neq i}g_{ij}\mathbb{E}_{\boldsymbol{\theta}_{-it}}\left[x_{it}(\widehat{\boldsymbol{\Theta}}_{t-1},\theta_{it},\boldsymbol{\theta}_{-it})x_{jt}(\widehat{\boldsymbol{\Theta}}_{t-1},\theta_{it},\boldsymbol{\theta}_{-it})\right]\\&+\frac{\partial}{\partial\hat{\theta}_{it}}\mathbb{E}_{\boldsymbol{\theta}_{-it}}x_{it}(\widehat{\boldsymbol{\Theta}}_{t-1},\theta_{it},\boldsymbol{\theta}_{-it})-\frac{\partial^2}{\partial\hat{\theta}_{it}^2}\mathbb{E}_{\boldsymbol{\theta}_{-it}}\left[\tau_{it}(\widehat{\boldsymbol{\Theta}}_{t-1},\theta_{it},\boldsymbol{\theta}_{-it})\right]\\&+\frac{\partial^2}{\partial\hat{\theta}_{it}^2}\mathbb{E}_{\boldsymbol{\theta}_{-it},\theta_{it+1}}\left[\mathcal{U}_{it+1}(\widehat{\boldsymbol{\Theta}}_{t-1},\theta_{it},\boldsymbol{\theta}_{-it},\theta_{it+1})\middle|\mathcal{I}_{it}\right]\\&+\frac{\partial^2}{\partial\hat{\theta}_{it}\partial\theta_{it}}\mathbb{E}_{\boldsymbol{\theta}_{-it},\theta_{it+1}}\left[\mathcal{U}_{it+1}(\widehat{\boldsymbol{\Theta}}_{t-1},\theta_{it},\boldsymbol{\theta}_{-it},\theta_{it+1})\middle|\mathcal{I}_{it}\right]=0.\end{aligned} \tag{46}$$

Expressions (45) and (46) imply:

[27] $\mathbb{E}\left[\mathbb{E}[X|\mathcal{G}_2]\middle|\mathcal{G}_1\right]=\mathbb{E}[X|\mathcal{G}_1]$ whenever $\mathcal{G}_1\subseteq\mathcal{G}_2$.

[28] Let $\Gamma(\theta_{it},\hat{\theta}_{it})\equiv\theta_{it}\mathbb{E}_{\boldsymbol{\theta}_{-it}}x_{it}(\widehat{\boldsymbol{\Theta}}_{t-1},\hat{\theta}_{it},\boldsymbol{\theta}_{-it})+\mathbb{E}_{\boldsymbol{\theta}_{-it},\theta_{it+1}}\left[\mathcal{U}_{it+1}(\widehat{\boldsymbol{\Theta}}_{t-1},\hat{\theta}_{it},\boldsymbol{\theta}_{-it},\theta_{it+1})|\theta_{it},\mathbf{x}_t(\widehat{\boldsymbol{\Theta}}_{t-1},\hat{\theta}_{it},\boldsymbol{\theta}_{-it})\right]$. We assume without loss of generality that $\hat{\theta}_{it}\geqslant\theta_{it}$. Then conditions $IC_{it}(\theta_{it}):\Gamma(\theta_{it},\theta_{it})-\tau(\theta_{it})\geqslant\Gamma(\theta_{it},\hat{\theta}_{it})-\tau(\hat{\theta}_{it})$, and $IC_{it}(\hat{\theta}_{it}):\Gamma(\hat{\theta}_{it},\hat{\theta}_{it})-\tau(\hat{\theta}_{it})\geqslant\Gamma(\hat{\theta}_{it},\theta_{it})-\tau(\theta_{it})$ imply that $\eta(x)\equiv\Gamma(x,\hat{\theta}_{it})-\Gamma(x,\theta_{it})$ is increasing in x. Following Lebesgue's Theorem, we have: $\eta(x)$ is differentiable almost everywhere. So $\eta'(x)\equiv\Gamma_x(x,\hat{\theta}_{it})-\Gamma_x(x,\theta_{it})\geqslant 0$, *a.e.*, which implies function $\Gamma_x(x,y)$ is increasing in y. Applying Lebesgue's Theorem again, $\Gamma_x(x,y)$ is differentiable in y. Hence, the local incentive-compatibility constraints implies the *a.e.* twice differentiability of function $\Gamma(\theta_{it},\hat{\theta}_{it})$. As a result, $\widetilde{\mathcal{U}}_{it}(\widehat{\boldsymbol{\Theta}},\theta_{it},\hat{\theta}_{it})=\Gamma(\theta_{it},\hat{\theta}_{it})-\tau_{it}(\hat{\theta}_{it})$ is also twice differentiable almost everywhere.

$$\frac{\partial}{\partial\hat{\theta}_{it}}\mathbb{E}_{\boldsymbol{\theta}_{-it}}x_{it}(\widehat{\boldsymbol{\Theta}}_{t-1},\theta_{it},\boldsymbol{\theta}_{-it})+\frac{\partial^2}{\partial\hat{\theta}_{it}\partial\theta_{it}}\mathbb{E}_{\boldsymbol{\theta}_{-it},\theta_{it+1}}\left[\mathcal{U}_{it+1}(\widehat{\boldsymbol{\Theta}}_{t-1},\boldsymbol{\theta}_t,\theta_{it+1})\middle|\mathcal{I}_{it}\right]>0. \tag{47}$$

Following a similar logic as (43), we have

$$\begin{aligned}&\frac{\partial}{\partial\theta_{it}}\mathbb{E}_{\boldsymbol{\theta}_{-it},\theta_{it+1}}\left[\mathcal{U}_{it+1}\left(\widehat{\boldsymbol{\Theta}}_{t-1},\boldsymbol{\theta}_t,\theta_{it+1}\right)\middle|\mathcal{I}_{it}\right]\\&=\mathbb{E}_{\boldsymbol{\theta}_{-it},\{\boldsymbol{\theta}_\tau\}_{\tau\geqslant t+1}}\left[\sum_{s\geqslant t+1}x_{is}(\widehat{\boldsymbol{\Theta}}_{t-1},\boldsymbol{\theta}_t,\cdots,\boldsymbol{\theta}_s)\middle|\mathcal{I}_{it}\right].\end{aligned} \tag{48}$$

(47) and (48) imply

$$IM_{it}:\frac{\partial}{\partial\hat{\theta}_{it}}\mathbb{E}_{\boldsymbol{\theta}_{-it},\{\boldsymbol{\theta}_\tau\}_{\tau\geqslant t+1}}\left[\sum_{s\geqslant t}x_{is}(\widehat{\boldsymbol{\Theta}}_{t-1},\boldsymbol{\theta}_t,\cdots,\boldsymbol{\theta}_s)\middle|\mathcal{I}_{it}\right]>0.$$

- $EN_{it}+IM'_{it}\Rightarrow IC_{it}$. The difference of continuation utilities upon truthful declaration and misreporting is

$$\begin{aligned}\Delta\mathcal{U}_{it}&\equiv\mathcal{U}_{it}(\widehat{\boldsymbol{\Theta}}_{t-1},\theta_{it})-\widetilde{\mathcal{U}}_{it}(\widehat{\boldsymbol{\Theta}}_{t-1},\theta_{it},\hat{\theta}_{it}) &(49)\\&=\mathcal{U}_{it}(\widehat{\boldsymbol{\Theta}}_{t-1},\theta_{it})-\mathcal{U}_{it}(\widehat{\boldsymbol{\Theta}}_{t-1},\hat{\theta}_{it})+\widetilde{\mathcal{U}}_{it}(\widehat{\boldsymbol{\Theta}}_{t-1},\hat{\theta}_{it},\hat{\theta}_{it})-\widetilde{\mathcal{U}}_{it}(\widehat{\boldsymbol{\Theta}}_{t-1},\theta_{it},\hat{\theta}_{it}) &(50)\\&=\int_{\hat{\theta}_{it}}^{\theta_{it}}\frac{\partial\mathcal{U}_{it}(\widehat{\boldsymbol{\Theta}}_{t-1},z)}{\partial\theta_{it}}dz-\int_{\hat{\theta}_{it}}^{\theta_{it}}\frac{\partial\widetilde{\mathcal{U}}_{it}(\widehat{\boldsymbol{\Theta}}_{t-1},z,\hat{\theta}_{it})}{\partial\theta_{it}}dz &(51)\\&=\int_{\hat{\theta}_{it}}^{\theta_{it}}[s(z,z)-s(z,\hat{\theta}_{it})]dz, &(52)\end{aligned}$$

where

$$s(\theta_{it},\hat{\theta}_{it})\equiv\left\{\begin{array}{c}\mathbb{E}_{\boldsymbol{\theta}_{-it}}\left[x_{it}(\widehat{\boldsymbol{\Theta}}_{t-1},\hat{\theta}_{it},\boldsymbol{\theta}_{-it})\right]+\\ \mathbb{E}_{\boldsymbol{\theta}_{-it},\{\boldsymbol{\theta}_\tau\}_{\tau\geqslant t+1}}\left[\sum_{s\geqslant t+1}x_{is}(\widehat{\boldsymbol{\Theta}}_{t-1},\hat{\theta}_{it},\boldsymbol{\theta}_{-it},\boldsymbol{\theta}_{t+1},\cdots,\boldsymbol{\theta}_s)\middle|\theta_{it},\mathbf{x}_t(\widehat{\boldsymbol{\Theta}}_{t-1},\hat{\theta}_{it},\boldsymbol{\theta}_{-it})\right]\end{array}\right\}.$$

(49) $\Rightarrow$ (50) follows from $\mathcal{U}_{it}(\widehat{\boldsymbol{\Theta}}_{t-1},\hat{\theta}_{it})=\widetilde{\mathcal{U}}_{it}(\widehat{\boldsymbol{\Theta}}_{t-1},\hat{\theta}_{it},\hat{\theta}_{it})$; (51) $\Rightarrow$ (52) follows from $EN_{it}:\partial\mathcal{U}_{it}(\widehat{\boldsymbol{\Theta}}_{t-1},\theta_{it})/\partial\theta_{it}=s(\theta_{it},\theta_{it})$, and an analogous condition $\partial\widetilde{\mathcal{U}}_{it}(\widehat{\boldsymbol{\Theta}}_{t-1},\theta_{it},\hat{\theta}_{it})/\partial\theta_{it}=s(\theta_{it},\hat{\theta}_{it})$. It follows from IM'_{it} that $s(\theta_{it},\hat{\theta}_{it})\geqslant s(\theta_{it},\theta_{it})$ whenever $\hat{\theta}_{it}\geqslant\theta_{it}$, and $s(\theta_{it},\hat{\theta}_{it})\leqslant s(\theta_{it},\theta_{it})$ whenever $\hat{\theta}_{it}\leqslant\theta_{it}$. Therefore, $\Delta\mathcal{U}_{it}\geqslant 0$, IC_{it} holds.

Appendix B. Proof of Lemma 2

Plugging EN_{i0} into the seller's objective function, then applying integration by part yields

$$\mathcal{V}=\left\{\begin{array}{c}\mathbb{E}_{\{\boldsymbol{\theta}_s\}_{s\geqslant 0}}\sum_{t=0}^{T}\left[\boldsymbol{\theta}_t^\top\mathbf{x}_t-\frac{1}{2}\mathbf{x}_t^\top(\mathbf{I}-\beta\mathbf{G})\mathbf{x}_t\right]-\sum_{i=1}^{n}\int_{\underline{\theta}_i}^{\overline{\theta}_i}\mathcal{U}_{i0}(\theta_{i0})f_i(\theta_{i0})d\theta_{i0}\\+\sum_{i=1}^{n}\int_{\underline{\theta}_i}^{\overline{\theta}_i}\lambda_i(\theta_{i0})\left[\mathbb{E}_{\boldsymbol{\theta}_{-i0},\{\boldsymbol{\theta}_s\}_{s\geqslant 1}}\sum_{t=0}^{T}x_{it}(\boldsymbol{\Theta}_t)-\mathcal{U}'_{i0}(\theta_{i0})\right]d\theta_{i0}\end{array}\right\} \tag{53}$$

$$=\left\{\begin{array}{c}\mathbb{E}_{\{\boldsymbol{\theta}_s\}_{s\geqslant 0}}\sum_{t=0}^{T}\left[\boldsymbol{\theta}_t^\top\mathbf{x}_t-\frac{1}{2}\mathbf{x}_t^\top(\mathbf{I}-\beta\mathbf{G})\mathbf{x}_t\right]+\sum_{i=1}^{n}\int_{\underline{\theta}_i}^{\overline{\theta}_i}\mathcal{U}_{i0}(\theta_{i0})\left[\lambda'_i(\theta_{i0})-f_i(\theta_{i0})\right]d\theta_{i0}+\\\sum_{i=1}^{n}\int_{\underline{\theta}_i}^{\overline{\theta}_i}\frac{\lambda_i(\theta_{i0})}{f_i(\theta_{i0})}\left[\mathbb{E}_{\boldsymbol{\theta}_{-i0},\{\boldsymbol{\theta}_s\}_{s\geqslant 1}}\sum_{t=0}^{T}x_{it}(\boldsymbol{\Theta}_t)\right]dF_i(\theta_{i0})-\sum_{i=1}^{n}\left[\lambda_i(\overline{\theta}_i)\mathcal{U}_{i0}(\overline{\theta}_i)-\lambda_i(\underline{\theta}_i)\mathcal{U}_{i0}(\underline{\theta}_i)\right]\end{array}\right\} \tag{54}$$

$$= \left\{ \begin{array}{c} \mathbb{E}_{\{\boldsymbol{\theta}_s\}_{s\geqslant 0}} \sum_{t=0}^{T} \left[(\boldsymbol{\theta}_t + \boldsymbol{\vartheta}(\boldsymbol{\theta}_0))^\top \mathbf{x}_t - \frac{1}{2}\mathbf{x}_t^\top (\mathbf{I} - \beta\mathbf{G})\mathbf{x}_t \right] + \\ \sum_{i=1}^{n} \int_{\underline{\theta}_i}^{\overline{\theta}_i} \mathcal{U}_{i0}(\theta_{i0}) \left[\lambda_i'(\theta_{i0}) - f_i(\theta_{i0}) \right] d\theta_{i0} - \sum_{i=1}^{n} \left[\lambda_i(\overline{\theta}_i)\mathcal{U}_{i0}(\overline{\theta}_i) - \lambda_i(\underline{\theta}_i)\mathcal{U}_{i0}(\underline{\theta}_i) \right] \end{array} \right\}, \tag{55}$$

where $\boldsymbol{\vartheta}(\boldsymbol{\theta}_0) \equiv (\vartheta_i(\theta_{i0}))_{i\in N}$, $\vartheta_i(\theta_{i0}) \equiv \lambda_i(\theta_{i0})/f_i(\theta_{i0})$, $\lambda_i(\theta_{i0})$ is the costate variable associated with EN_0.

We consider a parametric family of comparison controls $\mathbf{x}_t(\boldsymbol{\Theta}_t) = \mathbf{x}_t^*(\boldsymbol{\Theta}_t) + a\mathbf{m}_t(\boldsymbol{\Theta}_t)$, $\forall t \in \mathcal{T}$, where $\{\mathbf{x}_t^*(\boldsymbol{\Theta}_t)\}_{t=0}^T$ are the optimal controls, $\mathfrak{M} \equiv \{\mathbf{m}_t(\boldsymbol{\Theta}_t)\}_{t=0}^T$ denote a class of perturbation functions, a is a parameter. With a slight abuse of notation, we write $\mathcal{U}_{i0}(\theta_{i0}, a), \forall i \in N$ as the corresponding state variables. The seller's objective can thus be represented as a function of a:

$$\mathcal{V}(a) \equiv \left\{ \begin{array}{c} \phi(a, \mathfrak{M}) + \sum_{i=1}^{n} \int_{\underline{\theta}_i}^{\overline{\theta}_i} \mathcal{U}_{i0}(\theta_{i0}, a) \left[\lambda_i'(\theta_{i0}) - h(\theta_{i0}) \right] d\theta_{i0} \\ - \sum_{i=1}^{n} \left[\lambda_i(\overline{\theta}_i)\mathcal{U}_{i0}(\overline{\theta}_i, a) - \lambda_i(\underline{\theta}_i)\mathcal{U}_{i0}(\underline{\theta}_i, a) \right] \end{array} \right\},$$

where

$$\phi(a, \mathfrak{M}) \equiv \mathbb{E}_{\{\boldsymbol{\theta}_s\}_{s\geqslant 0}} \sum_{t=0}^{T} \left\{ [\boldsymbol{\theta}_t + \boldsymbol{\vartheta}(\boldsymbol{\theta}_0)]^\top \left(\mathbf{x}_t^* + a\mathbf{m}_t\right) - \frac{1}{2}\left(\mathbf{x}_t^* + a\mathbf{m}_t\right)^\top (\mathbf{I} - \beta\mathbf{G}) \left(\mathbf{x}_t^* + a\mathbf{m}_t\right) \right\}.$$

Since $\{\mathbf{x}_t^*(\boldsymbol{\Theta}_t)\}_{t=0}^T$ are optimal controls, the function $\mathcal{V}(a)$ assumes its maximum at $a = 0$. Differentiating with respect to a and evaluating at $a = 0$ gives

$$\mathcal{V}'(0) = \left\{ \begin{array}{c} \frac{\partial \phi}{\partial a}(0, \mathfrak{M}) + \sum_{i=1}^{n} \int_{\underline{\theta}_i}^{\overline{\theta}_i} \frac{\partial \mathcal{U}_{i0}}{\partial a}(\theta_{i0}, 0) \left[\lambda_i'(\theta_{i0}) - h(\theta_{i0}) \right] d\theta_{i0} \\ - \sum_{i=1}^{n} \left[\lambda_i(\overline{\theta}_i) \frac{\partial \mathcal{U}_{i0}}{\partial a}(\overline{\theta}_i, 0) - \lambda_i(\underline{\theta}_i) \frac{\partial \mathcal{U}_{i0}}{\partial a}(\underline{\theta}_i, 0) \right] \end{array} \right\} = 0.$$

Since functions $\{\mathbf{m}_t(\boldsymbol{\Theta}_t)\}_{t=0}^T$ and $\frac{\partial \mathcal{U}_{i0}}{\partial a}(\theta, a)$ are chosen arbitrarily, we have

$$\frac{\partial \phi}{\partial a}(0, \mathfrak{M}) = 0, \forall \mathfrak{M}, \tag{56}$$

$$\lambda_{i0}'(\theta) = f_i(\theta), \forall \theta \in [\underline{\theta}_i, \overline{\theta}_i], \forall i \in N, \tag{57}$$

$$\lambda_{i0}(\underline{\theta}_i) = \lambda_{i0}(\overline{\theta}_i) = 0, \forall i \in N. \tag{58}$$

(57) is the costate equation, (58) is the transversality condition with free endpoints $\underline{\theta}_i$ and $\overline{\theta}_i$.

The seller's relaxed problem simplifies to the following dynamic program

$$[\mathscr{P}_r] : \left\{ \begin{array}{c} \max_{\{\mathbf{x}_t(\boldsymbol{\Theta}_t)\}_{t=0}^T} \mathbb{E}_{\{\boldsymbol{\theta}_s\}_{s\geqslant 0}} \sum_{t=0}^{T} \left[(\boldsymbol{\theta}_t + \boldsymbol{\vartheta}(\boldsymbol{\theta}_0))^\top \mathbf{x}_t(\boldsymbol{\Theta}_t) - \frac{1}{2}\mathbf{x}_t^\top(\boldsymbol{\Theta}_t)(\mathbf{I} - \beta\mathbf{G})\mathbf{x}_t(\boldsymbol{\Theta}_t) \right] \\ s.t.: \boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}_t + \mathbf{A}\mathbf{x}_t + \boldsymbol{\epsilon}_t, (57), (58). \end{array} \right\}.$$

Representing the pointwise maximum value of date t as $V_t(\boldsymbol{\theta}_0, \boldsymbol{\theta}_t)$, we obtain the optimal contract $\mathbf{x}_t$ from the following Bellman equation

$$V_t(\boldsymbol{\theta}_0, \boldsymbol{\theta}_t) = \max_{\mathbf{x}_t \in \mathbb{R}^N} \left\{ [\boldsymbol{\theta}_t + \boldsymbol{\vartheta}(\boldsymbol{\theta}_0)]^\top \mathbf{x}_t - \frac{1}{2}\mathbf{x}_t^\top (\mathbf{I} - \beta\mathbf{G})\mathbf{x}_t + \mathbb{E}_{\boldsymbol{\theta}_{t+1}} [V_{t+1}(\boldsymbol{\theta}_0, \boldsymbol{\theta}_{t+1}) | \boldsymbol{\theta}_t, \mathbf{x}_t] \right\}.$$

To solve this problem, we adopt a guess-and-verify approach. That is, we start with an initial guess for the value function, obtain the optimal policy function, then compute the new value function. A recursive formula is thus obtained when comparing the initially guessed and the realized value functions. We make an initial guess that $V_t(\cdot, \cdot)$ is of the following quadratic form

$$V_t(\boldsymbol{\theta}_0, \boldsymbol{\theta}_t) = \frac{1}{2} [\boldsymbol{\theta}_t + \boldsymbol{\vartheta}(\boldsymbol{\theta}_0)]^\top \mathbf{M}_t [\boldsymbol{\theta}_t + \boldsymbol{\vartheta}(\boldsymbol{\theta}_0)] + \frac{a_t}{2}. \tag{59}$$

By the formula of quadratic form expectation,[29] the type-evolving equation (4) and $\mathbb{E}(\boldsymbol{\epsilon}_t)=0$, $\mathbb{V}ar(\boldsymbol{\epsilon}_t)=\sigma_t^2\mathbf{I}$, we have

$$V_t(\boldsymbol{\theta}_0,\boldsymbol{\theta}_t)=\max_{\mathbf{x}_t}\left\{\begin{array}{l}[(\mathbf{I}+\mathbf{AM}_{t+1})(\boldsymbol{\theta}_t+\boldsymbol{\vartheta}(\boldsymbol{\theta}_0))]^\top\mathbf{x}_t-\frac{1}{2}\mathbf{x}_t^\top(\mathbf{I}-\beta\mathbf{G}-\mathbf{AM}_{t+1}\mathbf{A})\mathbf{x}_t\\+\frac{1}{2}[\boldsymbol{\theta}_t+\boldsymbol{\vartheta}(\boldsymbol{\theta}_0)]^\top\mathbf{M}_{t+1}[\boldsymbol{\theta}_t+\boldsymbol{\vartheta}(\boldsymbol{\theta}_0)]+\frac{\sigma_t^2}{2}Tr(\mathbf{M}_{t+1})+\frac{1}{2}a_{t+1}\end{array}\right\}. \tag{60}$$

We therefore get the optimal quality from the first-order condition[30]

$$\mathbf{x}_t^*(\boldsymbol{\theta}_0,\boldsymbol{\theta}_t)=(\mathbf{I}-\beta\mathbf{G}-\mathbf{AM}_{t+1}\mathbf{A})^{-1}(\mathbf{I}+\mathbf{AM}_{t+1})[\boldsymbol{\theta}_t+\boldsymbol{\vartheta}(\boldsymbol{\theta}_0)], \tag{61}$$

and the value function

$$V_t(\boldsymbol{\theta}_0,\boldsymbol{\theta}_t)=\left\{\begin{array}{c}\frac{1}{2}[\boldsymbol{\theta}_t+\boldsymbol{\vartheta}(\boldsymbol{\theta}_0)]^\top\left[(\mathbf{I}+\mathbf{M}_{t+1}\mathbf{A})(\mathbf{I}-\beta\mathbf{G}-\mathbf{AM}_{t+1}\mathbf{A})^{-1}(\mathbf{I}+\mathbf{AM}_{t+1})\right][\boldsymbol{\theta}_t+\boldsymbol{\vartheta}(\boldsymbol{\theta}_0)]\\+\frac{1}{2}[\boldsymbol{\theta}_t+\boldsymbol{\vartheta}(\boldsymbol{\theta}_0)]^\top\mathbf{M}_{t+1}[\boldsymbol{\theta}_t+\boldsymbol{\vartheta}(\boldsymbol{\theta}_0)]+\frac{1}{2}\sigma_t^2Tr(\mathbf{M}_{t+1})+\frac{1}{2}a_{t+1}\end{array}\right\}. \tag{62}$$

Comparing (59) and (62) gives the following recursive expressions:

$$\mathbf{M}_t=(\mathbf{I}+\mathbf{M}_{t+1}\mathbf{A})[\mathbf{I}-\beta\mathbf{G}-\mathbf{AM}_{t+1}\mathbf{A}]^{-1}(\mathbf{I}+\mathbf{AM}_{t+1})+\mathbf{M}_{t+1}, \tag{63}$$

$$a_t=\sigma_t^2Tr(\mathbf{M}_{t+1})+a_{t+1}, \tag{64}$$

with end values $\mathbf{M}_T=\mathbf{M}\equiv(\mathbf{I}-\beta\mathbf{G})^{-1}$ and $a_T=0$. An inductive calculation yields

$$\mathbf{M}_t=\frac{T-t+1}{1-(T-t)s}\mathbf{M}, \tag{65}$$

$$a_t=\left\{\begin{array}{ll}\left(\sum_{\tau=t}^{T-1}\frac{\sigma_\tau^2(T-\tau)}{1-(T-\tau-1)s}\right)Tr\mathbf{M} & \text{if } t\leqslant T-1\\ 0 & \text{if } \quad t=T\end{array}\right. . \tag{66}$$

Substituting (65) into (61), we get

$$\mathbf{x}_t^*(\boldsymbol{\theta}_0,\boldsymbol{\theta}_t)=\frac{1}{1-(T-t)s}(\mathbf{I}-\beta\mathbf{G})^{-1}[\boldsymbol{\theta}_t+\vartheta(\boldsymbol{\theta}_0)]. \tag{67}$$

The expected payoff attained is therefore

$$\mathcal{V}^*\equiv\frac{1}{2}\mathbb{E}[\boldsymbol{\theta}_0+\boldsymbol{\vartheta}(\boldsymbol{\theta}_0)]^\top\mathbf{M}_0[\boldsymbol{\theta}_0+\boldsymbol{\vartheta}(\boldsymbol{\theta}_0)]+\frac{a_0}{2}. \tag{68}$$

The only work left is to determine $\boldsymbol{\vartheta}(\boldsymbol{\theta}_0)$. Since $\underline{\theta}_i$ and $\overline{\theta}_i$ are of different signs, the sign of $\mathcal{U}_{i0}'(\theta_{i0})=\mathbb{E}_{\boldsymbol{\theta}_{-i0},\{\boldsymbol{\theta}_s\}_{s\geqslant1}}[\sum_{t\geqslant0}x_{it}^*(\boldsymbol{\theta}_0,\boldsymbol{\theta}_t)|\mathcal{I}_{i0}]$ changes from negative to positive over interval $[\underline{\theta}_i,\overline{\theta}_i]$. Existing studies on countervailing incentives (e.g., Lewis and Sappington (1989), Maggi and Rodriguez-Clare (1995)) suggest that $\mathcal{U}_{i0}(\theta_{i0})$ is $U-$shaped with a flat bottom where IR_{i0} binds. That is, there exist cutoffs θ_i^*,θ_i^{**} such that $\mathcal{U}_{i0}(\theta_{i0})$ decreases (resp., remains zero; increases) for $\theta_{i0}\in[\underline{\theta}_i,\theta_i^*)$ (resp., $\theta_{i0}\in[\theta_i^*,\theta_i^{**}]$; $\theta_{i0}\in(\theta_i^{**},\overline{\theta}_i]$).

Using type-evolving equation (4) and optimal quality (67), we have

$$\begin{aligned}&\mathbb{E}_{\boldsymbol{\theta}_t}\left[\mathbf{x}_t^*(\boldsymbol{\theta}_0,\boldsymbol{\theta}_t)\middle|\mathcal{I}_{it-1}\right]\\&=\frac{(\mathbf{I}-\beta\mathbf{G})^{-1}}{1-(T-t)s}\left[\boldsymbol{\theta}_{t-1}+\mathbf{A}\mathbf{x}_{t-1}^*(\boldsymbol{\theta}_0,\boldsymbol{\theta}_{t-1})+\boldsymbol{\vartheta}(\boldsymbol{\theta}_0)\right]\\&=\frac{(\mathbf{I}-\beta\mathbf{G})^{-1}}{1-(T-t)s}\left[\mathbf{I}+\frac{\mathbf{A}(\mathbf{I}-\beta\mathbf{G})^{-1}}{1-(T-t+1)s}\right][\boldsymbol{\theta}_{t-1}+\boldsymbol{\vartheta}(\boldsymbol{\theta}_0)]\\&=\frac{(\mathbf{I}-\beta\mathbf{G})^{-1}}{1-(T-t+1)s}[\boldsymbol{\theta}_{t-1}+\boldsymbol{\vartheta}(\boldsymbol{\theta}_0)]=\mathbf{x}_{t-1}^*(\boldsymbol{\theta}_0,\boldsymbol{\theta}_{t-1}),\forall t\geqslant1.\end{aligned} \tag{69}$$

Repeated application of (69) and the iterated expectation formula yields

[29] Let $\mathbf{z}$ be an $n-$dimensional random vector with $\mathbb{E}(\mathbf{z})=\boldsymbol{\mu}$, and $\mathbb{V}ar(\mathbf{z})=\boldsymbol{\Sigma}$, $\boldsymbol{\Lambda}$ is a symmetric matrix, then $\mathbb{E}\left(\mathbf{z}^\top\boldsymbol{\Lambda}\mathbf{z}\right)=\boldsymbol{\mu}^\top\Lambda\boldsymbol{\mu}+Tr(\boldsymbol{\Lambda}\boldsymbol{\Sigma})$, where $Tr(\cdot)$ denotes the trace of a square matrix.

[30] The positive definiteness of matrix $\mathbf{I}-\beta\mathbf{G}-\mathbf{AM}_{t+1}\mathbf{A}$ will be checked ex post once we have the expression of $\mathbf{M}_t$.

$$\begin{aligned}\mathbb{E}_{\boldsymbol{\theta}_s}\left[\mathbf{x}_s^*(\boldsymbol{\theta}_0,\boldsymbol{\theta}_s)\middle|\mathcal{I}_{i0}\right] &= \mathbb{E}_{\boldsymbol{\theta}_1}\left[\cdots\mathbb{E}_{\boldsymbol{\theta}_{s-1}}\left[\mathbb{E}_{\boldsymbol{\theta}_s}\left[\mathbf{x}_s^*(\boldsymbol{\theta}_0,\boldsymbol{\theta}_s)|\mathcal{I}_{is-1}\right]\middle|\mathcal{I}_{is-2}\right]\cdots\middle|\mathcal{I}_{i0}\right]\\ &= \mathbb{E}\left[\cdots\mathbb{E}_{\boldsymbol{\theta}_{s-1}}\left[\mathbf{x}_{s-1}^*(\boldsymbol{\theta}_0,\boldsymbol{\theta}_{s-1})\middle|\mathcal{I}_{is-2}\right]\cdots\middle|\mathcal{I}_{i0}\right]\\ &= \cdots = \mathbb{E}_{\boldsymbol{\theta}_1}\left[\mathbf{x}_1^*(\boldsymbol{\theta}_0,\boldsymbol{\theta}_1)\middle|\mathcal{I}_{i0}\right] = \mathbf{x}_0^*(\boldsymbol{\theta}_0,\boldsymbol{\theta}_0),\ \forall s\geqslant 0.\end{aligned}$$

It follows that for $\theta_{i0}\in[\theta_i^*,\theta_i^{**}]$,

$$\begin{aligned}\mathcal{U}_{i0}'(\theta_{i0}) &= \mathbb{E}_{\boldsymbol{\theta}_{-i0}}\mathbb{E}_{\{\boldsymbol{\theta_s}\}_{s\geqslant 1}}\left[\sum_{s=0}^{T}x_{is}^*\left(\boldsymbol{\theta}_0,\boldsymbol{\theta}_s\right)\middle|\mathcal{I}_{i0}\right],\\ &= \frac{T+1}{1-Ts}e_i^\top\left(\mathbf{I}-\beta\mathbf{G}\right)^{-1}\begin{bmatrix}\theta_{i0}+\vartheta_i(\theta_{i0})\\ \mathbb{E}_{\boldsymbol{\theta}_{-i0}}\left[\boldsymbol{\theta}_{-i0}+\boldsymbol{\vartheta}_{-i}(\boldsymbol{\theta}_{-i0})\right]\end{bmatrix}\equiv 0,\end{aligned}\tag{70}$$

$$\mathcal{U}_{i0}''(\theta_{i0}) = \frac{T+1}{1-Ts}e_i^\top\left(\mathbf{I}-\beta\mathbf{G}\right)^{-1}e_i[1+\vartheta_i'(\theta_{i0})]\equiv 0,\tag{71}$$

where e_i is the ith standard base vector of space $\mathbb{R}^n$. Using conditions (57), (58) and (71), we can represent ϑ_i as a function of θ_{i0} and a constant c_i:

$$\vartheta_i(\theta_{i0},c_i)=\begin{cases}\frac{F_i(\theta_{i0})}{f_i(\theta_{i0})}=\theta_{i0}-\underline{\theta}_i & if\quad \theta_{i0}\in[\underline{\theta}_i,\theta_i^*)\\ -\theta_{i0}+c_i & if\quad \theta_{i0}\in[\theta_i^*,\theta_i^{**}]\\ \frac{F_i(\theta_{i0})-1}{f_i(\theta_{i0})}=\theta_{i0}-\overline{\theta}_i & if\quad \theta_{i0}\in(\theta_i^{**},\overline{\theta}_i]\end{cases}.\tag{72}$$

By continuity of $\vartheta_i(\cdot,\cdot)$ in θ_{i0}, we have

$$\theta_i^*(c_i)=\frac{c_i+\underline{\theta}_i}{2}\quad\text{and}\quad\theta_i^{**}(c_i)=\frac{c_i+\overline{\theta}_i}{2}.\tag{73}$$

It follows from (70) and (72) that for $\forall\theta_{i0}\in[\theta_i^*,\theta_i^{**}]$,

$$m_{ii}[\theta_{i0}+\vartheta_i(\theta_{i0})]+\sum_{j\neq i}m_{ij}\mathbb{E}[\theta_{j0}+\vartheta_j(\theta_{j0})]=m_{ii}c_i+\sum_{i\neq j}m_{ij}\mu_j(c_j)=0,\forall i\in N,\tag{74}$$

where m_{ij} is the (i,j) entry of matrix $\mathbf{M}\equiv(\mathbf{I}-\beta\mathbf{G})^{-1}$,

$$\begin{aligned}\mu_i(c_i) &\equiv \mathbb{E}\left[\theta_{i0}+\vartheta_i(\theta_{i0})\right]\\ &= \begin{bmatrix}\int_{\underline{\theta}_i}^{\theta_i^*(c_i)}\left(s+\frac{F_i(s)}{f_i(s)}\right)dF_i(s)+\int_{\theta_i^*(c_i)}^{\theta_i^{**}(c_i)}c_i dF_i(s)\\ +\int_{\theta_i^{**}(c_i)}^{\overline{\theta}_i}\left(s-\frac{1-F_i(s)}{f_i(s)}\right)dF_i(s)\end{bmatrix}=\frac{c_i}{2}.\end{aligned}$$

System (74) can be rewritten as $(\mathbf{M}+\widehat{\mathbf{M}})\mathbf{c}=\mathbf{0}$, where $\widehat{\mathbf{M}}\equiv diag\{m_{11},\cdots m_{nn}\}$, $\mathbf{0}$ denotes a vector of zeros. Since the coefficient matrix $\mathbf{M}+\widehat{\mathbf{M}}$ is invertible, the intercepts vector $\mathbf{c}^*=\mathbf{0}$ is determined uniquely.

Hence, the virtual valuation $\vartheta_i(\theta_{i0})$ in expression (67) is

$$\vartheta_i(\theta_{i0})=\begin{cases}\theta_{i0}-\underline{\theta}_i & if\quad \theta_{i0}\in\left[\underline{\theta}_i,\frac{\underline{\theta}_i}{2}\right)\\ -\theta_{i0} & if\quad \theta_{i0}\in\left[\frac{\underline{\theta}_i}{2},\frac{\overline{\theta}_i}{2}\right]\\ \theta_{i0}-\overline{\theta}_i & if\quad \theta_{i0}\in\left(\frac{\overline{\theta}_i}{2},\overline{\theta}_i\right]\end{cases}.$$

Inserting $\mu_i(c_i^*)=0$ and

$$\nu_i(c_i^*)\equiv\mathbb{V}ar\left[\theta_{i0}+\vartheta_i(\theta_{i0},c_i^*)\right]=\int_{\underline{\theta}_i}^{\frac{\underline{\theta}_i}{2}}\frac{(2s-\underline{\theta}_i)^2}{\overline{\theta}_i-\underline{\theta}_i}ds+\int_{\frac{\overline{\theta}_i}{2}}^{\overline{\theta}_i}\frac{(2s-\overline{\theta}_i)^2}{\overline{\theta}_i-\underline{\theta}_i}ds=\frac{\overline{\theta}_i^2}{6},$$

into (68), we get the seller's expected payoff:

$$\mathcal{V}^*=\frac{1}{12}\frac{T+1}{1-Ts}\bar{\boldsymbol{\theta}}^\top\widehat{\mathbf{M}}\bar{\boldsymbol{\theta}}+\frac{1}{2}\sum_{t=0}^{T-1}\frac{\sigma_t^2(T-t)}{1-(T-t-1)s}Tr\mathbf{M},$$

where $\bar{\boldsymbol{\theta}}\equiv(\overline{\theta}_1,\cdots,\overline{\theta}_n)^\top$.

Appendix C. Proof of Proposition 1

We now verify whether or not $\mathbf{x}_t^*(\boldsymbol{\theta}_0, \boldsymbol{\theta}_t)$ obtained from $[\mathscr{P}_r]$ satisfies the ignored conditions IM'_{it}. In case of violation, we resort to an ironing procedure to derive the optimal incentive compatible contract.

- **Step 1**: Checking the implementability conditions $IM'_{it}, \forall i \in N, t \in \mathcal{T}$. For $\forall t \geqslant 1$ and $\tau \geqslant t+1$, we have

$$\mathbb{E}_{\boldsymbol{\theta}_{-it}}\left[\mathbf{x}_t^*(\widehat{\boldsymbol{\Theta}}_{t-1}, \hat{\theta}_{it}, \boldsymbol{\theta}_{-it})\right] = \mathbb{E}_{\boldsymbol{\theta}_{-it}}\left[\mathbf{x}_t^*(\hat{\boldsymbol{\theta}}_0, \hat{\theta}_{it}, \boldsymbol{\theta}_{-it})\right]$$
$$= \frac{(\mathbf{I}-\beta\mathbf{G})^{-1}}{1-(T-t)s}\begin{bmatrix}\hat{\theta}_{it}+\vartheta_i(\hat{\theta}_{i0})\\ \mathbb{E}[\boldsymbol{\theta}_{-it}]+\boldsymbol{\vartheta}_{-i}(\hat{\boldsymbol{\theta}}_{-i0})\end{bmatrix}, \tag{75}$$

$$\mathbb{E}_{\boldsymbol{\theta}_{-it}}\mathbb{E}_{\{\boldsymbol{\theta}_s\}_{s=t+1}^{\tau}}\left[\mathbf{x}_\tau^*(\widehat{\boldsymbol{\Theta}}_{t-1}, \hat{\theta}_{it}, \boldsymbol{\theta}_{-it}, (\boldsymbol{\theta}_s)_{s=t+1}^{\tau})|\widehat{\mathcal{I}}_{it}\right]$$
$$= \mathbb{E}_{\boldsymbol{\theta}_{-it}}\mathbb{E}_{\boldsymbol{\theta}_\tau}\left[\mathbf{x}_\tau^*(\hat{\boldsymbol{\theta}}_0, \boldsymbol{\theta}_\tau)\middle|\widehat{\mathcal{I}}_{it}\right] \tag{76}$$
$$= \mathbb{E}_{\boldsymbol{\theta}_{-it}}\mathbb{E}_{\boldsymbol{\theta}_{t+1}}\left[\cdots\mathbb{E}_{\boldsymbol{\theta}_{\tau-1}}\left[\mathbb{E}_{\boldsymbol{\theta}_\tau}\left[\mathbf{x}_\tau^*(\hat{\boldsymbol{\theta}}_0, \boldsymbol{\theta}_\tau)|\widehat{\mathcal{I}}_{i\tau-1}^t\right]\middle|\widehat{\mathcal{I}}_{i\tau-2}^t\right]\cdots\middle|\widehat{\mathcal{I}}_{it}\right] \tag{77}$$
$$= \cdots = \mathbb{E}_{\boldsymbol{\theta}_{-it}}\mathbb{E}_{\boldsymbol{\theta}_{t+1}}\left[\mathbf{x}_{t+1}^*(\hat{\boldsymbol{\theta}}_0, \boldsymbol{\theta}_{t+1})\middle|\widehat{\mathcal{I}}_{it}\right] \tag{78}$$
$$= \frac{1}{1-(T-t-1)s}(\mathbf{I}-\beta\mathbf{G})^{-1}\,\mathbb{E}_{\boldsymbol{\theta}_{-it}}\left[\boldsymbol{\theta}_t+\boldsymbol{\vartheta}(\hat{\boldsymbol{\theta}}_0)+\mathbf{A}\mathbf{x}_t^*(\hat{\boldsymbol{\theta}}_0, \hat{\theta}_{it}, \boldsymbol{\theta}_{-it})\right] \tag{79}$$
$$= \frac{(\mathbf{I}-\beta\mathbf{G})^{-1}}{1-(T-t)s}\begin{bmatrix}\hat{\theta}_{it}+\vartheta_i(\hat{\theta}_{i0})\\ \mathbb{E}[\boldsymbol{\theta}_{-it}]+\boldsymbol{\vartheta}_{-i}(\hat{\boldsymbol{\theta}}_{-i0})\end{bmatrix} + \frac{(\mathbf{I}-\beta\mathbf{G})^{-1}\,e_i}{1-(T-t-1)s}(\theta_{it}-\hat{\theta}_{it}), \tag{80}$$

where $\widehat{\mathcal{I}}_{it} \equiv (\widehat{\boldsymbol{\Theta}}_{t-1}, \theta_{it}, \hat{\theta}_{it})$ denotes the information set available to i at date$-t$ upon a common public history $\widehat{\boldsymbol{\Theta}}_{t-1}$, when his true type is θ_{it} but he misreports it as $\hat{\theta}_{it}$. $\widehat{\mathcal{I}}_{i\tau}^t \equiv \left(\widehat{\boldsymbol{\Theta}}_{t-1}, \theta_{it}, \hat{\theta}_{it}, \boldsymbol{\theta}_{-it}, (\boldsymbol{\theta}_s)_{s\geqslant t+1}^{\tau}\right)$ denotes the date$-\tau$ ($\tau \geqslant t+1$) information set following $\widehat{\mathcal{I}}_{it}$ till time t. (76) $\Rightarrow$ (77) follows from recursive applications of the law of iterated expectations; (77) $\Rightarrow$ (78) follows from repeated applications of $\mathbb{E}[\mathbf{x}_{p+1}^*(\hat{\boldsymbol{\theta}}_0, \boldsymbol{\theta}_{p+1})|\widehat{\mathcal{I}}_{ip}^t] = \mathbf{x}_p^*(\hat{\boldsymbol{\theta}}_0, \boldsymbol{\theta}_p), \forall p \in \{t+1, \cdots, \tau-1\}$; (79) is obtained by substituting $\mathbf{x}_{t+1}^*(\hat{\boldsymbol{\theta}}_0, \boldsymbol{\theta}_{t+1}) = \frac{1}{1-(T-t-1)s}(\mathbf{I}-\beta\mathbf{G})^{-1}[\boldsymbol{\theta}_{t+1}+\boldsymbol{\vartheta}(\hat{\boldsymbol{\theta}}_0)]$ and $\mathbb{E}[\boldsymbol{\theta}_{t+1}|\widehat{\mathcal{I}}_{it}] = \boldsymbol{\theta}_t + \mathbf{A}\mathbf{x}_t^*(\hat{\boldsymbol{\theta}}_0, \hat{\theta}_{it}, \boldsymbol{\theta}_{-it})$ into (78); (80) is obtained by inserting

$$\mathbf{x}_t^*(\hat{\boldsymbol{\theta}}_0, \hat{\theta}_{it}, \boldsymbol{\theta}_{-it}) = \frac{1}{1-(T-t)s}(\mathbf{I}-\beta\mathbf{G})^{-1}\begin{bmatrix}\hat{\theta}_{it}+\vartheta_i(\hat{\theta}_{i0})\\ \boldsymbol{\theta}_{-it}+\boldsymbol{\vartheta}_{-i}(\hat{\boldsymbol{\theta}}_{-i0})\end{bmatrix}$$

into (79).

Differentiating (75) and (80), we get

$$\frac{\partial}{\partial\hat{\theta}_{it}}\mathbb{E}_{\boldsymbol{\theta}_{-it}}\left[\mathbf{x}_t^*(\widehat{\boldsymbol{\Theta}}_{t-1}, \hat{\theta}_{it}, \boldsymbol{\theta}_{-it})\right] = \frac{m_{ii}}{1-(T-t)s}, \forall\boldsymbol{\Theta}_{t-1}, \tag{81}$$

$$\frac{\partial}{\partial\hat{\theta}_{it}}\mathbb{E}_{\boldsymbol{\theta}_{-it}}\mathbb{E}_{\{\boldsymbol{\theta}_s\}_{s=t+1}^{\tau}}\left[\mathbf{x}_\tau^*(\widehat{\boldsymbol{\Theta}}_{t-1}, \hat{\theta}_{it}, \boldsymbol{\theta}_{-it}, (\boldsymbol{\theta}_s)_{s=t+1}^{\tau})\middle|\widehat{\mathcal{I}}_{it}\right]$$
$$= \left[\frac{1}{1-(T-t)s} - \frac{1}{1-(T-t-1)s}\right]m_{ii}, \forall\boldsymbol{\Theta}_{t-1}, \forall\tau \geqslant t+1, \tag{82}$$

where m_{ii} is the ith diagonal entry of matrix $\mathbf{M} \equiv (\mathbf{I}-\beta\mathbf{G})^{-1}$. We can therefore verify conditions IM'_{it} hold for $\forall t \in \mathcal{T}\backslash\{0\}, \forall i \in N$ and $\forall\boldsymbol{\Theta}_{t-1}$:

$$\frac{\partial}{\partial\hat{\theta}_{it}}\mathbb{E}_{\boldsymbol{\theta}_{-it},\{\boldsymbol{\theta}_\tau\}_{\tau=t+1}^{T}}\left[\sum_{s=t}^{T}x_{is}^*(\widehat{\boldsymbol{\Theta}}_{t-1}, \hat{\theta}_{it}, \boldsymbol{\theta}_{-it}, \cdots, \boldsymbol{\theta}_s)\middle|\widehat{\mathcal{I}}_{it}\right]$$
$$= \left\{\frac{1}{1-(T-t)s} + (T-t)\left[\frac{1}{1-(T-t)s} - \frac{1}{1-(T-t-1)}\right]\right\}m_{ii}$$
$$= \frac{(1+s)m_{ii}}{[1-(T-t)s][1-(T-t-1)s]} \geqslant 0.$$

This inequality follows from $m_{ii} > 0$ and assumption $s \in (-1, 0)$.

The only remaining work is to verify the implementability conditions of the initial period $IM'_{i0}, \forall i \in N$. For period $t = 0$ we have

$$\mathbb{E}_{\boldsymbol{\theta}_{-i0}}\mathbf{x}_0^*(\hat{\theta}_{i0},\boldsymbol{\theta}_{-i0}) = \frac{1}{1-Ts}(\mathbf{I}-\beta\mathbf{G})^{-1}\begin{bmatrix}\hat{\theta}_{i0}+\vartheta_i(\hat{\theta}_{i0})\\ \mathbb{E}[\boldsymbol{\theta}_{-i0}+\boldsymbol{\vartheta}_{-i}(\boldsymbol{\theta}_{-i0})]\end{bmatrix}. \tag{83}$$

By the same logic from (76) to (80), we have: for every $\tau \geqslant 1$

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{\theta}_{-i0}}\mathbb{E}_{\boldsymbol{\theta}_\tau}\left[\mathbf{x}_\tau^*(\hat{\theta}_{i0},\boldsymbol{\theta}_{-i0},\boldsymbol{\theta}_\tau)|\widehat{\mathcal{I}}_{i0}\right]\\
&=\mathbb{E}_{\boldsymbol{\theta}_{-i0}}\mathbb{E}_{\boldsymbol{\theta}_1}\left[\cdots\mathbb{E}_{\boldsymbol{\theta}_{\tau-1}}\left[\mathbb{E}_{\boldsymbol{\theta}_\tau}\left[\mathbf{x}_\tau^*(\hat{\theta}_{i0},\boldsymbol{\theta}_{-i0},\boldsymbol{\theta}_\tau)|\widehat{\mathcal{I}}_{i\tau-1}^0\right]\middle|\widehat{\mathcal{I}}_{i\tau-2}^0\right]\cdots\middle|\widehat{\mathcal{I}}_{i0}\right]\\
&=\cdots=\mathbb{E}_{\boldsymbol{\theta}_{-i0}}\mathbb{E}_{\boldsymbol{\theta}_1}\left[\mathbf{x}_1^*(\hat{\theta}_{i0},\boldsymbol{\theta}_{-i0},\boldsymbol{\theta}_1)\middle|\widehat{\mathcal{I}}_{i0}\right]\\
&=\frac{1}{1-(T-1)s}(\mathbf{I}-\beta\mathbf{G})^{-1}\mathbb{E}_{\boldsymbol{\theta}_{-i0}}\left[\boldsymbol{\theta}_0+\boldsymbol{\vartheta}(\hat{\theta}_{i0},\boldsymbol{\theta}_{-i0})+\mathbf{A}\mathbf{x}_0^*(\hat{\theta}_{i0},\boldsymbol{\theta}_{-i0})\right]\\
&=\frac{(\mathbf{I}-\beta\mathbf{G})^{-1}}{1-Ts}\begin{bmatrix}\hat{\theta}_{i0}+\vartheta_i(\hat{\theta}_{i0})\\ \mathbb{E}[\boldsymbol{\theta}_{-i0}+\boldsymbol{\vartheta}_{-i}(\boldsymbol{\theta}_{-i0})]\end{bmatrix}+\frac{(\mathbf{I}-\beta\mathbf{G})^{-1}e_i\left(\theta_{i0}-\hat{\theta}_{i0}\right)}{1-(T-1)s}.
\end{aligned} \tag{84}$$

Differentiating (83) and (84) with respect to $\hat{\theta}_{i0}$, then summing over time yields

$$\begin{aligned}
&\frac{\partial}{\partial\hat{\theta}_{i0}}\mathbb{E}_{\boldsymbol{\theta}_{-i0},\{\boldsymbol{\theta}_\tau\}_{\tau=1}^T}\left[\sum_{s=0}^T x_{is}^*(\hat{\theta}_{i0},\boldsymbol{\theta}_{-i0},\boldsymbol{\theta}_s)\middle|\widehat{\mathcal{I}}_{i0}\right]\\
&=\frac{m_{ii}[1+\vartheta_i'(\hat{\theta}_{i0})]}{1-Ts}+T\left\{\frac{m_{ii}[1+\vartheta_i'(\hat{\theta}_{i0})]}{1-Ts}-\frac{m_{ii}}{1-(T-1)s}\right\}\\
&=\frac{1+T}{1-Ts}m_{ii}[1+\vartheta_i'(\hat{\theta}_{i0})]-\frac{T}{1-(T-1)s}m_{ii}.
\end{aligned} \tag{85}$$

IM_{i0}' therefore requires

$$\vartheta_i'(\hat{\theta}_{i0}) \geqslant \frac{T(1-Ts)}{(T+1)(1-Ts+s)}-1\equiv\varphi. \tag{86}$$

Assumption $s\in(-1,0)$ implies $\varphi\in(-1,0)$. It can be seen easily from (14) and (15) that IM_{i0}' holds for either $\theta_{i0}\in[\underline{\theta}_i,\underline{\theta}_i/2)$ or $\theta_{i0}\in(\overline{\theta}_i/2,\overline{\theta}_i]$. But it fails for $\theta_{i0}\in[\underline{\theta}_i/2,\overline{\theta}_i/2]$. It can be found easily from expressions (81) to (84) that: for $\forall t\geqslant 1,\forall\theta_{it},\hat{\theta}_{it},\forall\theta_{i0},\theta_{i0},\forall\boldsymbol{\Theta}_{t-1}$, we have

$$\begin{aligned}
&\frac{\partial}{\partial\hat{\theta}_{it}}\mathbb{E}_{\boldsymbol{\theta}_{-it},\{\boldsymbol{\theta}_\tau\}_{\tau=t+1}^T}\left[\sum_{s=t}^T x_{is}^*(\widehat{\boldsymbol{\Theta}}_{t-1},\hat{\theta}_{it},\boldsymbol{\theta}_{-it},\cdots,\boldsymbol{\theta}_s)\middle|\widehat{\mathcal{I}}_{it}\right]=\frac{(1+s)m_{ii}}{[1-(T-t)s][1-(T-t-1)s]},\\
&\frac{\partial}{\partial\hat{\theta}_{it}}\mathbb{E}_{\boldsymbol{\theta}_{-it},\{\boldsymbol{\theta}_\tau\}_{\tau=t+1}^T}\left[\sum_{s=t}^T x_{is}^*(\widehat{\boldsymbol{\Theta}}_{t-1},\theta_{it},\boldsymbol{\theta}_{-it},\cdots,\boldsymbol{\theta}_s)\middle|\mathcal{I}_{it}\right]=\frac{(1+s)m_{ii}}{[1-(T-t)s][1-(T-t-1)s]},\\
&\frac{\partial}{\partial\hat{\theta}_{i0}}\mathbb{E}_{\boldsymbol{\theta}_{-i0},\{\boldsymbol{\theta}_\tau\}_{\tau=1}^T}\left[\sum_{s=0}^T x_{is}^*(\hat{\theta}_{i0},\boldsymbol{\theta}_{-i0},\cdots,\boldsymbol{\theta}_s)\middle|\widehat{\mathcal{I}}_{i0}\right]=\frac{1+T}{1-Ts}m_{ii}[1+\vartheta_i'(\hat{\theta}_{i0})]-\frac{T}{1-(T-1)s}m_{ii},\\
&\frac{\partial}{\partial\hat{\theta}_{i0}}\mathbb{E}_{\boldsymbol{\theta}_{-i0},\{\boldsymbol{\theta}_\tau\}_{\tau=1}^T}\left[\sum_{s=0}^T x_{is}^*(\theta_{i0},\boldsymbol{\theta}_{-i0},\cdots,\boldsymbol{\theta}_s)\middle|\mathcal{I}_{i0}\right]=\frac{1+T}{1-Ts}m_{ii}[1+\vartheta_i'(\theta_{i0})]-\frac{T}{1-(T-1)s}m_{ii}.
\end{aligned}$$

IM_{it} and IM_{it}' are therefore equivalent, so IC_{it} holds if and only if EN_{it} and IM_{it} are met. Therefore, the relaxed contract $\mathbf{x}^*(\boldsymbol{\theta}_0,\boldsymbol{\theta}_t)$ fails to meet IC_{i0} since IM_{i0} is violated.

- **Step 2**: Optimal ironing of virtual valuations. To meet IM_{i0}, we need to perform an ironing procedure on $\boldsymbol{\vartheta}(\cdot)$. Here we can't adopt the traditional horizontal ironing technique, but rather to use a negatively-sloped ironing line $\varphi(\theta_{i0}-d_i)$. The critical work is to determine the optimal intercepts $\mathbf{d}^*\equiv(d_i^*)_{i=1}^n$.
  We write the ironed virtual function as

$$\bar{\vartheta}_i(\theta_{i0},d_i)=\begin{cases}\frac{F_i(\theta_{i0})}{f_i(\theta_{i0})}=\theta_{i0}-\underline{\theta}_i & if\ \ \theta_{i0}\in[\underline{\theta}_i,\check{\theta}_i)\\ \varphi(\theta_{i0}-d_i) & if\ \ \theta_{i0}\in[\check{\theta}_i,\hat{\theta}_i]\\ \frac{F_i(\theta_{i0})-1}{f_i(\theta_{i0})}=\theta_{i0}-\overline{\theta}_i & if\ \ \theta_{i0}\in(\hat{\theta}_i,\overline{\theta}_i]\end{cases},$$

for cutoffs $\check{\theta}_i$ and $\hat{\theta}_i$. The continuity of $\bar{\vartheta}_i(\cdot,d_i)$ implies

$$\check{\theta}_i = \frac{\underline{\theta}_i - \varphi d_i}{1-\varphi} \text{ and } \hat{\theta}_i = \frac{\overline{\theta}_i - \varphi d_i}{1-\varphi}.$$

The mean and variance of virtual valuation are:

$$\begin{aligned}
\bar{\mu}_i(d_i) &\equiv \mathbb{E}\left[\theta_{i0} + \overline{\vartheta}_i(\theta_{i0}, d_i)\right] \\
&= \frac{1}{\overline{\theta}_i - \underline{\theta}_i}\left\{\int_{\underline{\theta}_i}^{\check{\theta}_i}(2s-\underline{\theta}_i)ds + \int_{\check{\theta}_i}^{\hat{\theta}_i}[s+\varphi(s-d_i)]ds + \int_{\hat{\theta}_i}^{\overline{\theta}_i}(2s-\overline{\theta}_i)ds\right\} \\
&= \frac{-\varphi d_i}{1-\varphi}, \\
\bar{\nu}_i(d_i) &\equiv \mathbb{V}ar[\theta_{i0} + \overline{\vartheta}_i(\theta_{i0}, d_i)] \\
&= \frac{1}{\overline{\theta}_i - \underline{\theta}_i}\left\{\begin{array}{c}\int_{\underline{\theta}_i}^{\check{\theta}_i}(2s-\underline{\theta}_i)^2 ds + \int_{\check{\theta}_i}^{\hat{\theta}_i}[s+\varphi(s-d_i)]^2 ds \\ + \int_{\hat{\theta}_i}^{\overline{\theta}_i}(2s-\overline{\theta}_i)^2 ds\end{array}\right\} - \mu_i^2(d_i) \\
&= \frac{\left(1+\varphi^2\right)\overline{\theta}_i^2 + 3\varphi^2 d_i^2}{3\left(1-\varphi\right)^2}.
\end{aligned}$$

The optimal ironed quality is therefore

$$\overline{\mathbf{x}}_t^*(\boldsymbol{\theta}_0, \boldsymbol{\theta}_t) = \frac{1}{1-(T-t)s}\left(\mathbf{I} - \beta\mathbf{G}\right)^{-1}\left[\boldsymbol{\theta}_t + \overline{\boldsymbol{\vartheta}}(\boldsymbol{\theta}_0, \mathbf{d})\right], \tag{87}$$

where $\overline{\boldsymbol{\vartheta}}(\boldsymbol{\theta}_0, \mathbf{d}) \equiv (\overline{\vartheta}_i(\boldsymbol{\theta}_0, d_i))_{i=1}^n$. Let $\bar{\boldsymbol{\mu}}(\mathbf{d}) \equiv (\bar{\mu}_i(d_i))_{i=1}^n$, $\bar{\boldsymbol{\Sigma}}(\mathbf{d}) \equiv diag\{(\bar{\nu}_i(d_i))_{i=1}^n\}$. Then, the seller's expected payoff associated with ironed scheme (87) is

$$\begin{aligned}
\overline{\mathcal{V}}^*(\mathbf{d}) &= \frac{1}{2}\bar{\boldsymbol{\mu}}^\top(\mathbf{d})\mathbf{M}_0\bar{\boldsymbol{\mu}}(\mathbf{d}) + \frac{1}{2}\left(Tr\bar{\boldsymbol{\Sigma}}(\mathbf{d})\mathbf{M}_0 + \sum_{t=0}^{T-1}\sigma_t^2 Tr\mathbf{M}_{t+1}\right) \\
&= \frac{1}{2}\frac{T+1}{1-Ts}\left(\frac{\varphi}{1-\varphi}\right)^2 \mathbf{d}^\top\left(\mathbf{M} + \widehat{\mathbf{M}}\right)\mathbf{d} + \left[\frac{(T+1)(1+\varphi^2)}{6(1-Ts)(1-\varphi)^2}\right]\overline{\boldsymbol{\theta}}^\top\widehat{\mathbf{M}}\overline{\boldsymbol{\theta}} \\
&\quad + \frac{1}{2}\sum_{t=0}^{T-1}\frac{(T-t)\sigma_t^2}{1-(T-t-1)s}Tr\mathbf{M}.
\end{aligned} \tag{88}$$

We define the following functions:

$$\begin{aligned}
\bar{x}_{i0}^+(\theta_{i0}, \mathbf{d}_{-i}) &\equiv \tfrac{m_{ii}}{1-Ts}\left[\theta_{i0} + \tfrac{F_i(\theta_{i0})}{f_i(\theta_{i0})}\right] + \tfrac{1}{1-Ts}\textstyle\sum_{j\neq i} m_{ij}\bar{\mu}_j(d_j), \\
\bar{x}_{i0}^0(\theta_{i0}, \mathbf{d}) &\equiv \tfrac{m_{ii}}{1-Ts}\left[\theta_{i0} + \varphi(\theta_{i0} - d_i)\right] + \tfrac{1}{1-Ts}\textstyle\sum_{j\neq i} m_{ij}\bar{\mu}_j(d_j), \\
\bar{x}_{i0}^-(\theta_{i0}, \mathbf{d}_{-i}) &\equiv \tfrac{m_{ii}}{1-Ts}\left[\theta_{i0} - \tfrac{1-F_i(\theta_{i0})}{f_i(\theta_{i0})}\right] + \tfrac{1}{1-Ts}\textstyle\sum_{j\neq i} m_{ij}\bar{\mu}_j(d_j), \\
\bar{x}_{i0}(\theta_{i0}, \mathbf{d}) &\equiv \begin{cases} \bar{x}_{i0}^+(\theta_{i0}, \mathbf{d}_{-i}) & \text{if } \theta_{i0} \in [\underline{\theta}_i, \check{\theta}_i(d_i)] \\ \bar{x}_{i0}^0(\theta_{i0}, \mathbf{d}) & \text{if } \theta_{i0} \in (\check{\theta}_i(d_i), \hat{\theta}_i(d_i)) \\ \bar{x}_{i0}^-(\theta_{i0}, \mathbf{d}_{-i}) & \text{if } \theta_{i0} \in [\hat{\theta}_i(d_i), \bar{\theta}_i] \end{cases}. \\
\theta_i^l(\mathbf{d}_{-i}) &= \frac{1}{2}\left(\frac{\varphi\sum_{j\neq i} m_{ij}d_j}{(1-\varphi)m_{ii}} + \underline{\theta}_i\right) \text{ and } \theta_i^r(\mathbf{d}_{-i}) = \frac{1}{2}\left(\frac{\varphi\sum_{j\neq i} m_{ij}d_j}{(1-\varphi)m_{ii}} + \overline{\theta}_i\right),
\end{aligned} \tag{89}$$

are critical values given implicitly by $x_{i0}^+(\theta_i^l, \mathbf{d}_{-i}) = 0$ and $x_{i0}^-(\theta_i^r, \mathbf{d}_{-i}) = 0$, respectively. Fig. 14 depicts piecewisely functions $\bar{x}_{i0}(\theta_{i0}, \mathbf{d})$ and $\bar{\vartheta}_i(\theta_{i0}, d_i)$ by red curves. It shows that the ironing line BC with slope $\varphi + 1 \in (0, 1)$ (resp. $B'C'$ with slope $\varphi \in (-1, 0)$) is under points D and E (resp. points D' and E'), and is above points A and F (resp. points A' and F'). d_i is therefore bounded by

$$\underline{d}_i(\mathbf{d}_{-i}) \equiv \max\left\{\theta_i^r + \frac{1-F_i(\theta_i^r)}{\varphi f_i(\theta_i^r)}, \underline{\theta}_i\right\} = \max\left\{\left(1-\frac{1}{\varphi}\right)\theta_i^r + \frac{1}{\varphi}\overline{\theta}_i, \underline{\theta}_i\right\} \tag{90}$$

and

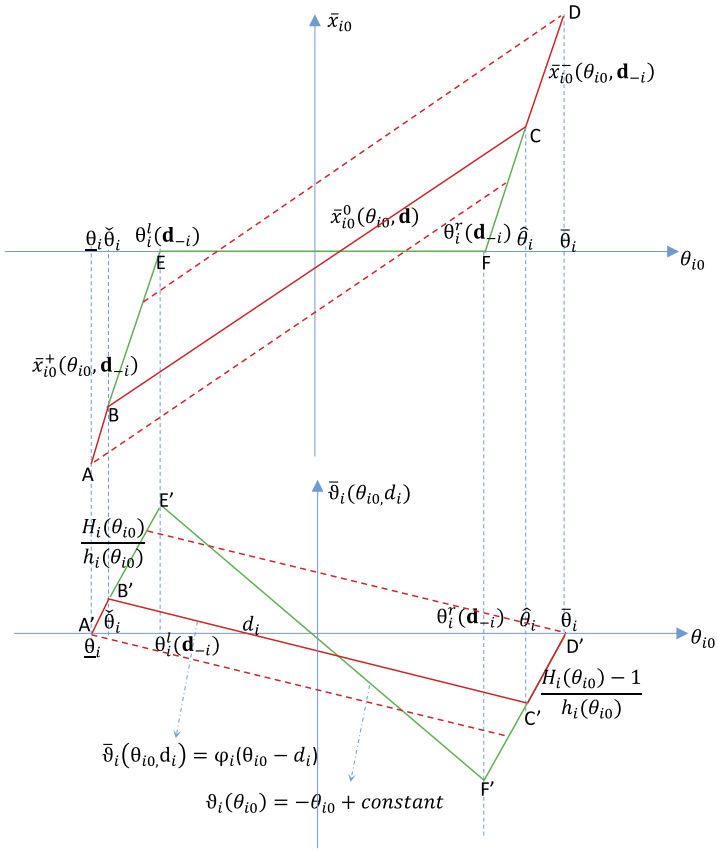


Fig. 14. Ironed quality.

$$\bar{d}_i(\mathbf{d}_{-i}) \equiv \min\left\{\theta_i^l - \frac{F_i(\theta_i^l)}{\varphi f_i(\theta_i^l)}, \bar{\theta}_i\right\} = \min\left\{\left(1 - \frac{1}{\varphi}\right)\theta_i^\ell + \frac{1}{\varphi}\underline{\theta}_i, \bar{\theta}_i\right\} \tag{91}$$

for given $\mathbf{d}_{-i} \equiv (d_j)_{j \neq i}$.
Condition

$$T \geqslant \frac{1}{2s}\left[1 - \sqrt{1 - 4s(1+s)\left(2 + \frac{\sum_{j\neq i} m_{ij}}{m_{ii}}\right)}\right]$$

is equivalent to $(1+3\varphi)m_{ii} + \varphi \sum_{j\neq i} m_{ij} \geqslant 0$, which guarantees

$$\theta_i^\ell \geqslant \frac{\frac{\varphi}{1-\varphi}\sum_{j\neq i} m_{ij}\bar{\theta}_i + m_{ii}\underline{\theta}_i}{2m_{ii}} \geqslant \frac{1+\varphi}{1-\varphi}\underline{\theta}_i,$$

$$\theta_i^r \leqslant \frac{\frac{\varphi}{1-\varphi}\sum_{j\neq i} m_{ij}\underline{\theta}_i + m_{ii}\bar{\theta}_i}{2m_{ii}} \leqslant \frac{1+\varphi}{1-\varphi}\bar{\theta}_i.$$

This in turn implies $\bar{d}_i = \bar{\theta}_i$ and $\underline{d}_i = \underline{\theta}_i$. The domain of $\mathbf{d}$ is thus $\mathcal{D} \equiv \prod_{i=1}^n [\underline{\theta}_i, \bar{\theta}_i]$. Given $0 < \beta < 1/\rho(\mathbf{G})$, $\mathbf{M} + \widehat{\mathbf{M}}$ is entry-wise positive, we thus have $\mathbf{d}^* \in \arg\max_{\mathbf{d}\in\mathcal{D}} \overline{\mathcal{V}}^*(\mathbf{d}) = \{\pm\bar{\boldsymbol{\theta}}\}$. The optimal ironed quality is therefore

$$\bar{\mathbf{x}}_t^*(\boldsymbol{\theta}_0, \boldsymbol{\theta}_t) = \frac{1}{1-(T-t)s}(\mathbf{I} - \beta\mathbf{G})^{-1}[\boldsymbol{\theta}_0 + \bar{\boldsymbol{\vartheta}}(\boldsymbol{\theta}_0)],$$

where $\bar{\boldsymbol{\vartheta}}(\boldsymbol{\theta}) \equiv (\bar{\vartheta}(\theta_{i0}))_{i\in N}$,

$$\bar{\boldsymbol{\vartheta}}(\theta_{i0}) = \begin{cases} \theta_{i0} - \underline{\theta}_i & \text{if } \theta_{i0} \in \left[\underline{\theta}_i, \frac{(1+\varphi)\underline{\theta}_i}{1-\varphi}\right) \\ \varphi(\theta_{i0} - \bar{\theta}_i) & \text{if } \theta_{i0} \in \left[\frac{(1+\varphi)\underline{\theta}_i}{1-\varphi}, \bar{\theta}_i\right] \end{cases} \text{ or } \begin{cases} \varphi(\theta_{i0} - \underline{\theta}_i) & \text{if } \theta_{i0} \in \left[\underline{\theta}_i, \frac{(1+\varphi)\bar{\theta}_i}{1-\varphi}\right) \\ \theta_{i0} - \bar{\theta}_i & \text{if } \theta_{i0} \in \left[\frac{(1+\varphi)\bar{\theta}_i}{1-\varphi}, \bar{\theta}_i\right] \end{cases}.$$

The corresponding optimal payoff is

$$\overline{\mathcal{V}}^*(\overline{\boldsymbol{\theta}}) = \frac{1}{2}\frac{T+1}{(1-Ts)(1-\varphi)^2}\overline{\boldsymbol{\theta}}^\top\left(\varphi^2\mathbf{M} + \frac{1+4\varphi^2}{3}\widehat{\mathbf{M}}\right)\overline{\boldsymbol{\theta}} + \frac{1}{2}\sum_{t=0}^{T}\frac{\sigma_t^2(T-t)}{1-(T-t-1)s}Tr(\mathbf{M}).$$

Appendix D. Proof of Proposition 2

Since

$$\frac{\partial}{\partial s}\left(\frac{-\varphi}{1-\varphi}\right) = \frac{T(1+T)}{\left[2+T-s(T^2-2)\right]^2} > 0,\ \frac{\partial}{\partial s}\left(\frac{1}{1-(T-t)s}\right) = \frac{T-t}{[1-(T-t)s]^2} \geqslant 0,$$

$$\frac{\partial}{\partial T}\left(\frac{-\varphi}{1-\varphi}\right) = \frac{(1+s)(2sT-1)}{\left[2+T-s(T^2-2)\right]^2} < 0,\ \frac{\partial}{\partial T}\left(\frac{1}{1-(T-t)s}\right) = \frac{s}{[1-(T-t)s]^2} < 0,$$

we have $\partial d_{it}/\partial s > 0$, $\partial d_{it}/\partial T < 0$, and

$$\frac{\partial d_{it}}{\partial t} = \frac{s(s+1)b_i}{\left(s\left(T^2-2\right)-T-2\right)(s(t-T)+1)^2} > 0.$$

Applying the matrix differentiation formula $\frac{\partial \mathbf{A}^{-1}}{\partial x} = -\mathbf{A}^{-1}\frac{\partial \mathbf{A}}{\partial x}\mathbf{A}^{-1}$, we get

$$\frac{\partial d_{it}}{\partial \beta} = \frac{-\varphi}{[1-(T-t)s](1-\varphi)}e_i^\top(\mathbf{I}-\beta\mathbf{G})^{-1}\mathbf{G}(\mathbf{I}-\beta\mathbf{G})^{-1}\mathbb{1}.$$

Since $0 < \beta < 1/\rho(\mathbf{G}) = 1/\lambda_1(\mathbf{G})$,[31] $\mathbf{M} \equiv (\mathbf{I}-\beta\mathbf{G})^{-1}$ can be expanded as the Neuman series: $(\mathbf{I}-\beta\mathbf{G})^{-1} = \sum_{k=0}^{\infty}\beta^k\mathbf{G}^k$. It follows that $\mathbf{M}$ is entry-wise positive,[32] and thus $\partial d_{it}/\partial\beta > 0$.

- *The case of small* β. Following the Neuman series expansion, we have $b_i(\beta,\overline{\boldsymbol{\theta}},\mathbf{G}) = \overline{\theta}_i + \sum_{k=1}^{\infty}\beta^k e_i^\top\mathbf{G}^k\overline{\boldsymbol{\theta}}$. Therefore, $\arg\max_{i\in N} d_{it}(\beta,\overline{\boldsymbol{\theta}},T,s,\mathbf{G}) = \arg\max_{i\in N}\overline{\theta}_i$ whenever $\beta \approx 0$ and $\overline{\theta}_i$s are heterogeneous. If $\overline{\theta}_i = \overline{\theta}, \forall i \in N$, $b_i(\beta,\mathbf{G}) = \overline{\theta}(1+\beta\rho_i+o(\beta))$, then $\arg\max_{i\in N} d_{it}(\beta,\overline{\boldsymbol{\theta}},T,s,\mathbf{G}) = \arg\max_{i\in N}\rho_i$.
- *The case of large* β. Let λ_is be the eigenvalues of $\mathbf{G}$ labelled in decreasing order: $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$. $\mathbf{P}$ is an orthonormal matrix with its ith column $\mathbf{p}_{\cdot i}$ the orthonormal eigenvector associated to λ_i. Since $\mathbf{G}$ is an irreducible/connected graph, by the Perron-Frobenius theorem, $\lambda_1 = \rho(\mathbf{G})$ is a simple eigenvalue, i.e., $\lambda_1 > \lambda_2$, and the principal eigenvector $\mathbf{p}_{\cdot 1}$ assigns the same sign to all nodes in the network. We assume, without loss of generality, that $p_{j1} > 0, \forall j \in N$. Then, the spectral decomposition $\mathbf{G} = \sum_{j=1}^{n}\lambda_j\mathbf{p}_{\cdot j}\mathbf{p}_{\cdot j}^\top$ implies

$$b_i(\beta,\overline{\boldsymbol{\theta}},\mathbf{G}) = \frac{e_i^\top\mathbf{p}_{\cdot 1}\mathbf{p}_{\cdot 1}^\top\overline{\boldsymbol{\theta}}}{1-\beta\lambda_1} + \sum_{j=2}^{n}\frac{e_i^\top\mathbf{p}_{\cdot j}\mathbf{p}_{\cdot j}^\top\overline{\boldsymbol{\theta}}}{1-\beta\lambda_j}.$$

As $\beta \approx 1/\rho(\mathbf{G}) = 1/\lambda_1$, $b_i(\beta,\overline{\boldsymbol{\theta}},\mathbf{G}) \approx \frac{e_i^\top\mathbf{p}_{\cdot 1}\mathbf{p}_{\cdot 1}^\top\overline{\boldsymbol{\theta}}}{1-\beta\lambda_1} = \frac{p_{i1}\mathbf{p}_{\cdot 1}^\top\overline{\boldsymbol{\theta}}}{1-\beta\lambda_1}$, therefore $\arg\max_{i\in N} d_{it}(\beta,\overline{\boldsymbol{\theta}},T,s,\mathbf{G}) = \arg\max_{i\in N} b_i(\beta,\overline{\boldsymbol{\theta}},\mathbf{G}) = \arg\max_{i\in N} p_{i1}$.

Appendix E. Proof of Proposition 3

Since $\Delta\overline{\theta} > 0, s < 0, \varphi < 0$, and $m_{ij} > 0, \forall i, j$, we have

$$\Delta d_{ijt}(\beta,\Delta\overline{\theta},T,s,\mathbf{G}) = -\frac{\varphi\Delta\overline{\theta}}{[1-(T-t)s](1-\varphi)}m_{ij}(\beta,\mathbf{G}) > 0.$$

- *Case* 1: $\beta \to 0$.

$$\begin{aligned}\Delta d_{ijt}(\beta,\Delta\overline{\theta},T,s,\mathbf{G}) &= -\frac{\varphi\Delta\overline{\theta}}{[1-(T-t)s](1-\varphi)}\sum_{s=0}^{\infty}\beta^s g_{ij}^{[s]}\\ &= -\frac{\varphi\Delta\overline{\theta}}{[1-(T-t)s](1-\varphi)}\sum_{s=k(i,j)}^{\infty}\beta^s g_{ij}^{[s]},\end{aligned}$$

[31] $\lambda_1(\mathbf{G})$ represents the first/largest eigenvalue of $\mathbf{G}$. By Perron-Frobenius Theorem, the irreducibility (connectivity) of $\mathbf{G}$ implies $\rho(\mathbf{G}) = \lambda_1(\mathbf{G})$.

[32] A positive definite matrix $\mathbf{A}$ with nonpositive off-diagonal elements, i.e., $a_{ij} \leqslant 0, \forall i \neq j$, is called an $M-$matrix in linear algebra literature. An $M-$matrix has an entry-wise positive inverse, i.e., $(\mathbf{A}^{-1})_{ij} > 0, \forall i, j$.

where $k(i,j) \equiv \min\{s | g_{ij}^{[s]} > 0\}$ is the geodesic distance, i.e., the count of edges along the shortest path, between nodes i and j. Therefore, $d_{it}(\beta, \overline{\boldsymbol{\theta}} + \Delta\overline{\boldsymbol{\theta}}_j, T, s, \mathbf{G}) - d_{it}(\beta, \overline{\boldsymbol{\theta}}, T, s, \mathbf{G}) = O(\beta^k)$ as $\beta \to 0$.

- *Case* 2: $\beta \to 1/\rho(\mathbf{G})$. It follows from

$$\Delta d_{ijt}(\beta, \Delta\overline{\theta}, T, s, \mathbf{G}) = -\frac{\varphi\Delta\overline{\theta}}{[1-(T-t)s](1-\varphi)}\left(\frac{e_i^\top \mathbf{p}_{\cdot 1}\mathbf{p}_{\cdot 1}^\top e_j}{1-\beta\lambda_1} + \sum_{s=2}^{n}\frac{e_i^\top \mathbf{p}_{\cdot s}\mathbf{p}_{\cdot s}^\top e_j}{1-\beta\lambda_s}\right),$$

that $\Delta d_{ijt}(\beta, \Delta\overline{\theta}, T, s, \mathbf{G}) \approx -\frac{\varphi\Delta\overline{\theta}}{[1-(T-t)s](1-\varphi)}\frac{p_{i1}p_{j1}}{1-\beta\lambda_1}$. Therefore,

$$\arg\max_{i\in N} \Delta d_{ijt}(\beta, \Delta\overline{\theta}, T, s, \mathbf{G}) = \arg\max_{j\in N} \Delta d_{ijt}(\beta, \Delta\overline{\theta}, T, s, \mathbf{G}) = \arg\max_{s\in N} p_{s1}$$

as $\beta \to 1/\rho(\mathbf{G})$.

Appendix F. Proof of Proposition 5

The change of expected absolute aggregate quality upon removal of agent i is

$$\left|\mathcal{Q}(\beta, \mathbf{G}) - \mathcal{Q}(\beta, \mathbf{G}^{-i})\right| \tag{92}$$

$$= -\frac{T+1}{1-Ts}\frac{\varphi\overline{\theta}}{1-\varphi}\left[\mathbb{1}^\top(\mathbf{I}-\beta\mathbf{G})^{-1}\mathbb{1} - \mathbb{1}^\top\left(\mathbf{I}-\beta\mathbf{G}^{-i}\right)^{-1}\mathbb{1}\right] \tag{93}$$

$$= \frac{(T+1)(-\varphi)\overline{\theta}\mathbb{1}^\top\begin{bmatrix} \frac{1}{1-\beta^2\mathbf{g}_i^\top(\mathbf{I}_{n-1}-\beta\mathbf{G}_{-i,-i})^{-1}\mathbf{g}_i} - 1 & \frac{\beta\mathbf{g}_i^\top(\mathbf{I}_{n-1}-\beta\mathbf{G}_{-i,-i})^{-1}}{1-\beta^2\mathbf{g}_i^\top(\mathbf{I}_{n-1}-\beta\mathbf{G}_{-i,-i})^{-1}\mathbf{g}_i} \\ \frac{\beta(\mathbf{I}_{n-1}-\beta\mathbf{G}_{-i,-i})^{-1}\mathbf{g}_i}{1-\beta^2\mathbf{g}_i^\top(\mathbf{I}_{n-1}-\beta\mathbf{G}_{-i,-i})^{-1}\mathbf{g}_i} & \frac{\beta^2(\mathbf{I}_{n-1}-\beta\mathbf{G}_{-i,-i})^{-1}\mathbf{g}_i\mathbf{g}_i^\top\mathbf{I}_{n-1}-\beta\mathbf{G}_{-i,-i})^{-1}}{1-\beta^2\mathbf{g}_i^\top(\mathbf{I}_{n-1}-\beta\mathbf{G}_{-i,-i})^{-1}\mathbf{g}_i} \end{bmatrix}\mathbb{1}}{(1-Ts)(1-\varphi)} \tag{94}$$

$$= -\frac{(T+1)\varphi\overline{\theta}}{(1-Ts)(1-\varphi)}\left\{\frac{\left[\frac{\beta\mathbf{g}_i^\top(\mathbf{I}_{n-1}-\beta\mathbf{G}_{-i,-i})^{-1}\mathbb{1}_{n-1}+1}{1-\beta^2\mathbf{g}_i^\top(\mathbf{I}_{n-1}-\beta\mathbf{G}_{-i,-i})^{-1}\mathbf{g}_i}\right]^2}{\left[\frac{1}{1-\beta^2\mathbf{g}_i^\top(\mathbf{I}_{n-1}-\beta\mathbf{G}_{-i,-i})^{-1}\mathbf{g}_i}\right]} - 1\right\} \tag{95}$$

$$= -\frac{(T+1)\varphi\overline{\theta}}{(1-Ts)(1-\varphi)}\left(\frac{b_i^2}{m_{ii}} - 1\right) = -\frac{(T+1)\varphi\overline{\theta}(c_i-1)}{(1-Ts)(1-\varphi)}, \tag{96}$$

where $m_{ii} = 1/[1-\beta^2\mathbf{g}_i^\top(\mathbf{I}_{n-1}-\beta\mathbf{G}_{-i,-i})^{-1}\mathbf{g}_i]$. (93) $\Rightarrow$ (94) follows from the partitioned inverse formula of matrix

$$(\mathbf{I}-\beta\mathbf{G})^{-1} = \begin{bmatrix} \frac{1}{1-\beta^2\mathbf{g}_i^\top(\mathbf{I}_{n-1}-\beta\mathbf{G}_{-i,-i})^{-1}\mathbf{g}_i} & \frac{\beta\mathbf{g}_i^\top(\mathbf{I}_{n-1}-\beta\mathbf{G}_{-i,-i})^{-1}}{1-\beta^2\mathbf{g}_i^\top(\mathbf{I}_{n-1}-\beta\mathbf{G}_{-i,-i})^{-1}\mathbf{g}_i} \\ \frac{\beta(\mathbf{I}_{n-1}-\beta\mathbf{G}_{-i,-i})^{-1}\mathbf{g}_i}{1-\beta^2\mathbf{g}_i^\top(\mathbf{I}_{n-1}-\beta\mathbf{G}_{-i,-i})^{-1}\mathbf{g}_i} & \begin{pmatrix} (\mathbf{I}_{n-1}-\beta\mathbf{G}_{-i,-i})^{-1} \\ +\frac{\beta^2(\mathbf{I}_{n-1}-\beta\mathbf{G}_{-i,-i})^{-1}\mathbf{g}_i\mathbf{g}_i^\top\mathbf{I}_{n-1}-\beta\mathbf{G}_{-i,-i})^{-1}}{1-\beta^2\mathbf{g}_i^\top(\mathbf{I}_{n-1}-\beta\mathbf{G}_{-i,-i})^{-1}\mathbf{g}_i} \end{pmatrix} \end{bmatrix}.$$

Therefore, the top influencer is $i^* \in \arg\max_{j\in N} c_j$.

In analogy, we have

$$\begin{aligned} &Tr(\mathbf{I}-\beta\mathbf{G})^{-1} - Tr(\mathbf{I}-\beta\mathbf{G}^{-i})^{-1} \\ &= Tr\begin{bmatrix} \frac{1}{1-\beta^2\mathbf{g}_i^\top(\mathbf{I}_{n-1}-\beta\mathbf{G}_{-i,-i})^{-1}\mathbf{g}_i} - 1 & \frac{\beta\mathbf{g}_i^\top(\mathbf{I}_{n-1}-\beta\mathbf{G}_{-i,-i})^{-1}}{1-\beta^2\mathbf{g}_i^\top(\mathbf{I}_{n-1}-\beta\mathbf{G}_{-i,-i})^{-1}\mathbf{g}_i} \\ \frac{\beta(\mathbf{I}_{n-1}-\beta\mathbf{G}_{-i,-i})^{-1}\mathbf{g}_i}{1-\beta^2\mathbf{g}_i^\top(\mathbf{I}_{n-1}-\beta\mathbf{G}_{-i,-i})^{-1}\mathbf{g}_i} & \frac{\beta^2(\mathbf{I}_{n-1}-\beta\mathbf{G}_{-i,-i})^{-1}\mathbf{g}_i\mathbf{g}_i^\top(\mathbf{I}_{n-1}-\beta\mathbf{G}_{-i,-i})^{-1}}{1-\beta^2\mathbf{g}_i^\top(\mathbf{I}_{n-1}-\beta\mathbf{G}_{-i,-i})^{-1}\mathbf{g}_i} \end{bmatrix} \\ &= \frac{\beta^2\mathbf{g}_i^\top(\mathbf{I}_{n-1}-\beta\mathbf{G}_{-i,-i})^{-2}\mathbf{g}_i + 1}{1-\beta^2\mathbf{g}_i^\top(\mathbf{I}_{n-1}-\beta\mathbf{G}_{-i,-i})^{-1}\mathbf{g}_i} - 1 = \frac{m_{ii}^{[2]}}{m_{ii}} - 1, \end{aligned}$$

where $m_{ii}^{[2]}$ is the diagonal element of matrix $\mathbf{M}^2 = (\mathbf{I}-\beta\mathbf{G})^{-2}$, i.e.,

$$m_{ii}^{[2]} \equiv e_i^\top(\mathbf{I}-\beta\mathbf{G})^{-2}e_i = \frac{1+\beta^2\mathbf{g}_i^\top(\mathbf{I}_{n-1}-\beta\mathbf{G}_{-i,-i})^{-2}\mathbf{g}_i}{\left[1-\beta^2\mathbf{g}_i^\top(\mathbf{I}_{n-1}-\beta\mathbf{G}_{-i,-i})^{-1}\mathbf{g}_i\right]^2}.$$

It follows that

$$\begin{aligned}
&\overline{\mathcal{V}}^*(\mathbf{G}) - \overline{\mathcal{V}}^*\left(\mathbf{G}^{-i}\right) \\
&= \phi_1 \mathbb{1}^\top \left[(\mathbf{I}-\beta\mathbf{G})^{-1} - \left(\mathbf{I}-\beta\mathbf{G}^{-i}\right)^{-1} \right] \mathbb{1} + \phi_2 Tr \left[(\mathbf{I}-\beta\mathbf{G})^{-1} - \left(\mathbf{I}-\beta\mathbf{G}^{-i}\right)^{-1} \right] \\
&= \phi_1 (c_i - 1) + \phi_2 \left(\frac{m_{ii}^{[2]}}{m_{ii}} - 1 \right).
\end{aligned}$$

Therefore, the key customer satisfies $j^* \in \arg\max_{i\in N} h_i$, where $h_i \equiv \phi_1 c_i + \phi_2 (m_{ii}^{[2]}/m_{ii})$.

Appendix G. Proof of Proposition 6

Using $b_i(\beta,\mathbf{G}) = e_i^\top \sum_{k=0}^{\infty} \beta^k \mathbf{G}^k \mathbb{1}$, $m_{ii}^{[2]}(\beta,\mathbf{G}) = \sum_{k=0}^{\infty}(k+1)\beta^k g_{ii}^{[k]}$, and $m_{ii}(\beta,\mathbf{G}) = \sum_{k=0}^{\infty}\beta^k g_{ii}^{[k]}$, we approximate $c_i(\beta)$ and $h_i(\beta)$ around zero by their Taylor expansions as

$$c_i(\beta) = 1 + 2\rho_i\beta + \left(2A_{2i} + \rho_i^2 - \rho_i\right)\beta^2 + \left(2A_{2i}\rho_i + 2A_{3i} - g_{ii}^{[3]} - 2\rho_i^2\right)\beta^3 + o(\beta^3), \tag{97}$$

$$h_i(\beta) = \begin{bmatrix} (\phi_1+\phi_2) + 2\phi_1\rho_i\beta + \left(2\phi_1 A_{2i} + \phi_1\rho_i^2 - \phi_1\rho_i + 2\phi_2\rho_i\right)\beta^2 + \\ \left(2\phi_1 A_{2i}\rho_i + 2\phi_1 A_{3i} + (3\phi_2 - \phi_1) g_{ii}^{[3]} - 2\phi_1\rho_i^2\right)\beta^3 + o(\beta^3) \end{bmatrix}, \tag{98}$$

where $g_{ii}^{[k]}$ is the ith diagonal element of matrix $\mathbf{G}^k$, $\rho_i = g_{ii}^{[2]} = e_i^\top \mathbf{G}\mathbb{1}$ is the degree of a vertex i, $A_{ki} \equiv e_i^\top \mathbf{G}^k \mathbb{1}$, $o(\cdot)$ denotes a higher order infinitesimal.

- If $\mathbf{G}$ is irregular and β is close to zero, the ranking in terms of c_is is as same as the ranking in terms of $2\rho_i\beta$; while the ranking in term of h_is aligns with the ranking in terms of $2\phi_1\rho_i\beta$. Since $\phi_1 > 0$, both TI and KC coincide with the node(s) with the largest vertex degree: $\arg\max_i c_i = \arg\max_i h_i = \arg\max_i \rho_i$.
- If $\mathbf{G}$ is a regular network, $\rho_i \equiv \rho$, $A_{ki} = e_i^\top \mathbf{G}^k \mathbb{1} = \rho^k, \forall i \in N, \forall k \geqslant 1$. The linear and quadratic terms of c_i in expressions (97) takes the form $\left(2A_{2i} + \rho_i^2 - \rho_i\right)\beta^2 + 2\rho_i\beta + 1 = \left(3\rho^2 - \rho\right)\beta^2 + 2\rho\beta + 1$. It is identical across i and is unable to discriminate among nodes. The corresponding terms of h_i in (98) are independent of i as well: $\left(2\phi_1 A_{2i} + \phi_1\rho_i^2 - \phi_1\rho_i + 2\phi_2\rho_i\right)\beta^2 + 2\phi_1\rho_i\beta + (\phi_1+\phi_2) = \left[3\phi_1\rho^2 + (2\phi_2 - \phi_1)\rho\right]\beta^2 + 2\phi_1\rho\beta + (\phi_1+\phi_2)$. So when β is close to zero, nodes are ranked by, respectively, the higher order terms $(2A_{2i}\rho_i + 2A_{3i} - g_{ii}^{[3]} - 2\rho_i^2)\beta^3 = (4\rho^3 - 2\rho^2 - g_{ii}^{[3]})\beta^3$ and $[2\phi_1 A_{2i}\rho_i + 2\phi_1 A_{3i} + (3\phi_2 - \phi_1)g_{ii}^{[3]} - 2\phi_1\rho_i^2]\beta^3 = [4\phi_1\rho^3 - 2\phi_1\rho^2 + (3\phi_2 - \phi_1)g_{ii}^{[3]}]\beta^3$. Since $g_{ii}^{[3]}$ is not identical across i, and $3\phi_2 - \phi_1 = \frac{(T+1)\overline{\theta}^2(1+3\varphi^2)}{2(1-Ts)(1-\varphi)^2} + \frac{3}{2}\sum_{t=0}^{T-1}\frac{(T-t)\sigma_t^2}{1-(T-t-1)s} > 0$, TI is the one with the smallest $g_{ii}^{[3]}$, but KC is the one having the highest $g_{ii}^{[3]}$.

Appendix H. Proof of Proposition 7

- Using again the spectral decomposition of $\mathbf{G}$, we get

$$\begin{aligned}
b_i &= e_i^\top (\mathbf{I}-\beta\mathbf{G})^{-1}\mathbb{1} = e_i^\top \sum_j \frac{\mathbf{p}_{\cdot j}\mathbf{p}_{\cdot j}^\top}{1-\beta\lambda_j}\mathbb{1} = \frac{p_{i1}\mathbf{p}_{\cdot 1}^\top\mathbb{1}}{1-\beta\lambda_1} + \sum_{j\neq 1}\frac{p_{ij}\mathbf{p}_{\cdot j}^\top\mathbb{1}}{1-\beta\lambda_j}, \\
m_{ii} &= e_i^\top (\mathbf{I}-\beta\mathbf{G})^{-1} e_i = e_i^\top \sum_j \frac{\mathbf{p}_{\cdot j}\mathbf{p}_{\cdot j}^\top}{1-\beta\lambda_j} e_i = \frac{p_{i1}^2}{1-\beta\lambda_1} + \sum_{j\neq 1}\frac{p_{ij}^2}{1-\beta\lambda_j}, \\
m_{ii}^{[2]} &= e_i^\top (\mathbf{I}-\beta\mathbf{G})^{-2} e_i = e_i^\top \sum_j \frac{\mathbf{p}_{\cdot j}\mathbf{p}_{\cdot j}^\top}{(1-\beta\lambda_j)^2} e_i = \frac{p_{i1}^2}{(1-\beta\lambda_1)^2} + \sum_{j\neq 1}\frac{p_{ij}^2}{(1-\beta\lambda_j)^2}.
\end{aligned}$$

As $\beta \to 1/\rho(\mathbf{G}) = 1/\lambda_1$, the Taylor expansions of $c_i(t)$ and $m_{ii}^{[2]}(t)/m_{ii}(t)$ with respect to $t \equiv 1-\beta\lambda_1$ are

$$\begin{aligned}
c_i(t) &= \frac{(\mathbf{p}_{\cdot 1}^\top\mathbb{1})^2}{t} + \lambda_1(\mathbf{p}_{\cdot 1}^\top\mathbb{1})^2(\Delta_i + \delta_i) + o(1), \\
\frac{m_{ii}^{[2]}(t)}{m_{ii}(t)} &= \frac{1}{t} + \lambda_1\Delta_i + o(1),
\end{aligned}$$

where $\Delta_i \equiv -\frac{1}{p_{i1}^2}\sum_{j\neq 1}\frac{p_{ij}^2}{\lambda_1-\lambda_j}$, $\delta_i \equiv \frac{2}{p_{i1}}\sum_{j\neq 1}\frac{p_{ij}(\mathbf{p}_{\cdot j}^\top\mathbb{1})}{(\lambda_1-\lambda_j)(\mathbf{p}_{\cdot 1}^\top\mathbb{1})}$. Condition (29) implies that $\Delta_i + \delta_i \geqslant \Delta_i + \underline{\delta} \geqslant \Delta_j + \overline{\delta} \geqslant \Delta_k + \overline{\delta} \geqslant \Delta_k + \delta_k$, $\forall i \in \mathbb{P}_1 \equiv \arg\max_{i\in N}\Delta_i$, $\forall j \in \mathbb{P}_2 \equiv \arg\max_{i\in N\backslash\mathbb{P}_1}\Delta_i$, $\forall k \in N\backslash\mathbb{P}_1$. So $\arg\max_{i\in N} c_i(t) = \arg\max_{i\in N} m_{ii}^{[2]}(t)/m_{ii}(t) = \mathbb{P}_1$. Therefore, both TI and KC are within $\mathbb{P}_1$.

- Δ_i has bounds $\underline{\Delta}_i$ and $\overline{\Delta}_i$:

$$\underline{\Delta}_i \equiv \frac{1}{\lambda_1-\lambda_2}\left(-\frac{1}{p_{i1}^2}\sum_{j=2}^{n} p_{ij}^2\right) \leqslant \Delta_i \leqslant \overline{\Delta}_i \equiv \frac{1}{\lambda_1-\lambda_n}\left(-\frac{1}{p_{i1}^2}\sum_{j=2}^{n} p_{ij}^2\right).$$

Since matrix $\mathbf{P}$ is orthogonal, $\sum_{j=1}^{n} p_{ij}^2 = 1$, we thus have

$$\underline{\Delta}_i = \frac{1}{\lambda_1-\lambda_2}\left(1-\frac{1}{p_{i1}^2}\right), \overline{\Delta}_i = \frac{1}{\lambda_1-\lambda_n}\left(1-\frac{1}{p_{i1}^2}\right).$$

Let $\mathbb{Q}_1 \equiv \arg\max_{i\in N} p_{i1}$, $\mathbb{Q}_2 \equiv \arg\max_{i\in N\setminus\mathbb{Q}_1} p_{i1}$, condition (30) implies $\Delta_i \geqslant \underline{\Delta}_i > \overline{\Delta}_j \geqslant \Delta_j \geqslant \Delta_k, \forall i \in \mathbb{Q}_1, j \in \mathbb{Q}_2, k \in N\setminus\mathbb{Q}_1$, so $\arg\max_{i\in N} \Delta_i = \mathbb{Q}_1$. Therefore both $TI(i^*)$ and $KC(j^*)$ are the nodes with the largest eigenvector centrality: $\{i^*, j^*\} \subseteq \mathbb{Q}_1$.

Appendix I. Proof of Proposition 8

Applying the Sherman-Morrison-Woodbury formula, we get

$$(\mathbf{I}-\beta\mathbf{G}^{+ij})^{-1} = \left(\mathbf{I}-\beta\mathbf{G}-\beta\mathbf{U}\mathbf{C}\mathbf{U}^\top\right)^{-1} = \mathbf{M}+\beta\mathbf{M}\mathbf{\Omega}^{ij}\mathbf{M}, \tag{99}$$

where

$$\mathbf{U} = \left[e_i, e_j\right], \mathbf{C} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

$$\mathbf{\Omega}^{ij} \equiv \frac{\left[e_i, e_j\right]\begin{bmatrix} \beta m_{jj} & 1-\beta m_{ij} \\ 1-\beta m_{ij} & \beta m_{ii} \end{bmatrix}\begin{bmatrix} e_i^\top \\ e_j^\top \end{bmatrix}}{\left(1-\beta m_{ij}\right)^2-\beta^2 m_{ii}m_{jj}}.$$

- *Key link.* $\Delta\mathcal{V}_{ij}(\beta) \equiv \overline{\mathcal{V}}^*(\mathbf{G}^{+ij}) - \overline{\mathcal{V}}^*(\mathbf{G}) = \phi_1\Delta\Pi_{ij}(\beta) + \phi_2\Delta\Xi_{ij}(\beta)$, where

$$\begin{aligned}\Delta\Pi_{ij}(\beta) &\equiv \mathbb{1}^\top(\mathbf{I}-\beta\mathbf{G}^{+ij})^{-1}\mathbb{1} - \mathbb{1}^\top(\mathbf{I}-\beta\mathbf{G})^{-1}\mathbb{1} = \beta\mathbb{1}^\top\mathbf{M}\Omega^{ij}\mathbf{M}\mathbb{1} \\ &= \frac{\beta\left[2b_ib_j(1-\beta m_{ij})+\beta m_{jj}b_i^2+\beta m_{ii}b_j^2\right]}{\left(1-\beta m_{ij}\right)^2-\beta^2 m_{ii}m_{jj}}, \\ \Delta\Xi_{ij}(\beta) &\equiv Tr\left(\mathbf{I}-\beta\mathbf{G}^{+ij}\right)^{-1} - Tr\left(\mathbf{I}-\beta\mathbf{G}\right)^{-1} = \beta Tr\mathbf{M}\Omega^{ij}\mathbf{M} \\ &= \frac{\beta\left[2m_{ij}^{[2]}(1-\beta m_{ij})+\beta m_{ii}m_{jj}^{[2]}+\beta m_{jj}m_{ii}^{[2]}\right]}{\left(1-\beta m_{ij}\right)^2-\beta^2 m_{ii}m_{jj}}.\end{aligned}$$

Therefore, the seller will choose a pair (i, j) among $\mathcal{G}_0$ to maximize

$$K_{ij}(\phi_1,\phi_2,\beta,\mathbf{G}) \equiv \frac{\begin{bmatrix} \phi_1\left[2b_ib_j(1-\beta m_{ij})+\beta m_{jj}b_i^2+\beta m_{ii}b_j^2\right] \\ +\phi_2\left[2m_{ij}^{[2]}(1-\beta m_{ij})+\beta m_{ii}m_{jj}^{[2]}+\beta m_{jj}m_{ii}^{[2]}\right]\end{bmatrix}}{\left(1-\beta m_{ij}\right)^2-\beta^2 m_{ii}m_{jj}}.$$

- *Stable link.* Agent i's gain from a new link (i, j) is

$$\Delta\mathscr{U}_{ij} = -\frac{(1+T)\varphi\overline{\theta}^2\beta}{3(1-\varphi)^2}\left[(1+2\varphi)e_i^\top\mathbf{M}\mathbf{\Omega}^{ij}\mathbf{M}e_i + 3\varphi e_i^\top\mathbf{M}\mathbf{\Omega}^{ij}\mathbf{M}\mathbb{1}\right].$$

Agent i will choose his favorite partner j among $\mathcal{G}_0^i$ to maximize

$$\begin{aligned}S_{ij}(\beta,\mathbf{G}) &= (1+2\varphi)e_i^\top\mathbf{M}\mathbf{\Omega}^{ij}\mathbf{M}e_i + 3\varphi e_i^\top\mathbf{M}\mathbf{\Omega}^{ij}\mathbf{M}\mathbb{1} \\ &= \frac{\left\{\begin{array}{l}(1+2\varphi)\left[\beta m_{ii}(m_{ii}m_{jj}-m_{ij}^2)+2m_{ii}m_{ij}\right] \\ +3\varphi\left[b_jm_{ii}+b_im_{ij}+\beta b_i(m_{ii}m_{jj}-m_{ij}^2)\right]\end{array}\right\}}{\left(1-\beta m_{ij}\right)^2-\beta^2 m_{ii}m_{jj}}.\end{aligned}$$

A link (i, j) is stable if

$$j \in \arg\max_{j' \in \mathcal{G}_0^i} S_{ij'}(\beta, \mathbf{G}) \text{ and } i \in \arg\max_{i' \in \mathcal{G}_0^j} S_{ji'}(\beta, \mathbf{G}) \tag{100}$$

with at least one expression holds with "=". (i, j) is strongly stable if (100) holds and $S_{ij}(\beta, \mathbf{G}) \geqslant 0, S_{ji}(\beta, \mathbf{G}) \geqslant 0$.

Appendix J. Proof of Proposition 9

From $(\mathbf{I} - \beta\mathbf{G})^{-1} = \sum_{k=0}^{\infty} \beta^k \mathbf{G}^k$ and $(\mathbf{I} - \beta\mathbf{G})^{-2} = \sum_{k=0}^{\infty}(k+1)\beta^k\mathbf{G}^k$, we have: for any pair $(i, j) \in \mathcal{G}_0$,

$$m_{ii}(\beta) = 1 + \rho_i\beta^2 + \sum_{k=3}^{\infty} \beta^k g_{ii}^{[k]},$$

$$m_{ij}(\beta) = \sum_{k=2}^{\infty} \beta^k g_{ij}^{[k]},$$

$$b_i(\beta) = 1 + \rho_i\beta + \sum_{k=2}^{\infty} \beta^k e_i^\top \mathbf{G}^k \mathbb{1},$$

$$m_{ii}^{[2]}(\beta) = 1 + 3\rho_i\beta^2 + \sum_{k=3}^{\infty}(k+1)\beta^k g_{ii}^{[k]},$$

$$m_{ij}^{[2]}(\beta) = \sum_{k=2}^{\infty}(k+1)\beta^k g_{ij}^{[k]}.$$

The Taylor expansions of $K_{ij}(\beta, \mathbf{G})$ and $S_{ij}(\beta, \mathbf{G})$ around zero are:

$$K_{ij}(\beta, \mathbf{G}) = \begin{bmatrix} 2\phi_1 + \left[2(1+\rho_i+\rho_j)\phi_1 + 2\phi_2\right]\beta + o(\beta^2) + \\ \left[2\left(1 + e_i^\top\mathbf{G}^2\mathbb{1} + e_j^\top\mathbf{G}^2\mathbb{1} + \rho_i + \rho_j + \rho_i\rho_j\right)\phi_1 + 6g_{ij}^{[2]}\phi_2\right]\beta^2 \end{bmatrix}, \tag{101}$$

$$S_{ij}(\beta, \mathbf{G}) = \begin{bmatrix} 3\varphi + (1 + 5\varphi + 3\varphi\rho_j)\beta + \\ \left(3\varphi + 6\varphi\rho_i + 3\varphi e_j^\top\mathbf{G}^2\mathbb{1} + (7\varphi+2)g_{ij}^{[2]}\right)\beta^2 + o(\beta^2) \end{bmatrix}. \tag{102}$$

When $\beta \approx 0$, the key link (i^*, j^*) maximizes $\rho_i + \rho_j$ within $\mathcal{G}_0$. Notice that $\frac{\partial S_{ij}}{\partial \rho_j} = 3\beta\varphi < 0$, so agent i prefers a new neighbor with the lowest degree.[33] Suppose that $\exists(i^*, j^*) \in \mathcal{K} \cap \mathcal{S}$. Then $\rho_{i^*} + \rho_{j^*} \geqslant \rho_i + \rho_j, \forall(i, j) \in \mathcal{G}_0$. Also, $\rho_{j^*} \leqslant \rho_j, \forall j \in \mathcal{G}_0^{i^*} \backslash \{j^*\}$, and $\rho_{i^*} \leqslant \rho_i, \forall i \in \mathcal{G}_0^{j^*} \backslash \{i^*\}$, with at least one inequality holding strictly.[34] So $\exists(i, j) \in (\mathcal{G}_{i^*}^0 \backslash \{j^*\}) \times (\mathcal{G}_{j^*}^0 \backslash \{i^*\})$ such that either $\rho_{i^*} + \rho_{j^*} < \rho_{i^*} + \rho_j$ or $\rho_{i^*} + \rho_{j^*} < \rho_i + \rho_{j^*}$ holds. This contradicts our assumption that (i^*, j^*) maximizes $\rho_i + \rho_j$ among $\mathcal{G}_0$.

Appendix K. Proof of Proposition 11

The optimal network intervention policy solves

$$\max_{\overline{\boldsymbol{\theta}} \in \mathbb{R}_+^n} \overline{\boldsymbol{\theta}}^\top \boldsymbol{\Sigma} \overline{\boldsymbol{\theta}}, s.t.: \|\overline{\boldsymbol{\theta}}\| \leqslant 1.$$

This is a canonical Rayleigh-Ritz eigenvalue problem except that the optimal solution is within the positive orthant $\mathbb{R}_+^n$. Nevertheless, the Perron-Frobenius Theorem guarantees $\mathbf{v}_1(\boldsymbol{\Sigma}) \in \mathbb{R}_+^n$ since $\boldsymbol{\Sigma}$ is entry-wise positive. So the optimal intervention policy is $\overline{\boldsymbol{\theta}}^* = \mathbf{v}_1(\boldsymbol{\Sigma})$.

- It follows from Neumann series expansion $\mathbf{M} \equiv (\mathbf{I} - \beta\mathbf{G})^{-1} = \sum_{k=0}^{\infty} \beta^k \mathbf{G}^k$ that

$$\boldsymbol{\Sigma} \equiv \varphi^2\mathbf{M} + \frac{1+4\varphi^2}{3}\widehat{\mathbf{M}} = \frac{1+7\varphi^2}{3}\mathbf{I} + \beta\varphi^2\mathbf{G} + o(\beta).$$

  So $\boldsymbol{\Sigma} \approx \frac{1+7\varphi^2}{3}\mathbf{I} + \beta\varphi^2\mathbf{G}$ when $\beta \approx 0$. Therefore, $\overline{\boldsymbol{\theta}}^*$ approaches the first principal component $\mathbf{p}_{\cdot 1} \equiv \mathbf{v}_1(\mathbf{G})$.

[33] Higher order term $g_{ij}^{[2]}$ helps to discriminate among nodes when ρ_js are identical and fails to provide a clear ranking.

[34] Condition $|\mathcal{G}_0^i| \geqslant 1, |\mathcal{G}_0^j| \geqslant 1, \forall(i, j) \in \mathcal{G}_0$ guarantees that $\mathcal{G}_0^{j^*} \backslash \{i^*\} \neq \emptyset$ and $\mathcal{G}_0^{i^*} \backslash \{j^*\} \neq \emptyset$.

- When $T \to \infty$, we have $\varphi \to 0$, then a higher priority is assigned to the maximization of $\frac{1+4\varphi^2}{3}\overline{\boldsymbol{\theta}}^\top \widehat{\mathbf{M}}\overline{\boldsymbol{\theta}}$ than to that of $\varphi^2\overline{\boldsymbol{\theta}}^\top \mathbf{M}\overline{\boldsymbol{\theta}}^\top$. So we need to find a solution within space $S_K \equiv \{\overline{\theta} \in \mathbb{R}^n_+ | \sum_{i\in K}\overline{\theta}_i^2 = 1, \overline{\theta}_j = 0, \forall j \in N\backslash K\}$, and attain a value $\frac{1+4\varphi^2}{3}m^*$ for the second term $\frac{1+4\varphi^2}{3}\overline{\boldsymbol{\theta}}^\top \widehat{\mathbf{M}}\overline{\boldsymbol{\theta}}$, where $m^* = m_{ii}(\beta, \mathbf{G}), \forall i \in K$. Then, we implement the maximization of term $\varphi^2\overline{\boldsymbol{\theta}}^\top \mathbf{M}\overline{\boldsymbol{\theta}}$ within S_K. It amounts to a lower dimensional problem

$$\max_{\overline{\boldsymbol{\theta}}_K \in \mathbb{R}^{|K|}_+} (\overline{\boldsymbol{\theta}}_K)^\top \mathbf{M}_{K,K}\overline{\boldsymbol{\theta}}_K, s.t.: \|\overline{\boldsymbol{\theta}}_K\| \leqslant 1,$$

which admits a solution $\mathbf{v}_1(\mathbf{M}_{K,K})$.[35] So the optimal solution to the overall problem is

$$\overline{\boldsymbol{\theta}}^* = \arg\max_{\overline{\boldsymbol{\theta}} \in S_K} \overline{\boldsymbol{\theta}}^\top \mathbf{M}\overline{\boldsymbol{\theta}} = \begin{bmatrix} \mathbf{v}_1(\mathbf{M}_{K,K}) \\ \mathbf{0}_{-K} \end{bmatrix},$$

and the maximized value attained is $\varphi^2\lambda_1(\mathbf{M}_{K,K}) + \frac{1+4\varphi^2}{3}m^*$.

Appendix L. Proof of Proposition 12

When approximating $\overline{\boldsymbol{\theta}}^*$ by $\mathbf{p}_{\cdot 1}$, the seller suffers a profit loss

$$0 \leqslant \overline{\mathcal{V}}^*(T, \beta, \overline{\boldsymbol{\theta}}^*, \mathbf{G}) - \overline{\mathcal{V}}^*(T, \beta, \mathbf{p}_{\cdot 1}, \mathbf{G}) \tag{103}$$

$$= \frac{(T+1)\left[\lambda_1\left(\varphi^2\mathbf{M} + \frac{1+4\varphi^2}{3}\widehat{\mathbf{M}}\right) - \left(\varphi^2\lambda_1(\mathbf{M}) + \frac{1+4\varphi^2}{3}\sum_{i=1}^n p_{i1}^2 m_{ii}\right)\right]}{2(1-Ts)(1-\varphi)^2} \tag{104}$$

$$\leqslant \frac{T+1}{2(1-Ts)(1-\varphi)^2}\frac{1+4\varphi^2}{3}\left(m^* - \sum_{i=1}^n p_{i1}^2 m_{ii}\right) \tag{105}$$

$$= \frac{T+1}{2(1-Ts)(1-\varphi)^2}\frac{1+4\varphi^2}{3}\beta^2\left(\rho^* - \sum_{i=1}^n p_{i1}^2\rho_i\right) + o(\beta^2), \tag{106}$$

where $\rho^* = \max_{i\in N}\rho_i$ and $m^* = \max_{i\in N} m_{ii}$ denote, respectively, the largest vertex degree and the largest resolvent subgraph centralities. (103) $\Rightarrow$ (104) follows from

$$\overline{\mathcal{V}}^*(T, \beta, \mathbf{p}_{\cdot 1}, \mathbf{G}) = \varphi^2\mathbf{p}_{\cdot 1}^\top\mathbf{M}\mathbf{p}_{\cdot 1} + \frac{1+4\varphi^2}{3}\mathbf{p}_{\cdot 1}^\top\widehat{\mathbf{M}}\mathbf{p}_{\cdot 1} = \varphi^2\lambda_1(\mathbf{M}) + \frac{1+4\varphi^2}{3}\sum_{i=1}^n p_{i1}^2 m_{ii};$$

(104) $\Rightarrow$ (105) follows from inequality $\lambda_1\left(\varphi^2\mathbf{M} + \frac{1+4\varphi^2}{3}\widehat{\mathbf{M}}\right) \leqslant \varphi^2\lambda_1(\mathbf{M}) + \frac{1+4\varphi^2}{3}\lambda_1(\widehat{\mathbf{M}}) = \varphi^2\lambda_1(\mathbf{M}) + \frac{1+4\varphi^2}{3}m^*$; (105) $\Rightarrow$ (106) is implied by $m_{ii} = 1 + \beta^2\rho_i + o(\beta^2)$, and $m^* = 1 + \beta^2\rho^* + o(\beta^2)$ around zero. Therefore, $\overline{\mathcal{V}}^*(T, \beta, \overline{\boldsymbol{\theta}}^*, \mathbf{G}) - \overline{\mathcal{V}}^*(T, \beta, \mathbf{p}_{\cdot 1}, \mathbf{G}) = O(\beta^2)$.

Similarly, the profit loss under approximation $\tilde{\boldsymbol{\theta}}$ is

$$\begin{aligned} 0 &\leqslant \overline{\mathcal{V}}^*(T, \beta, \overline{\boldsymbol{\theta}}^*, \mathbf{G}) - \overline{\mathcal{V}}^*(T, \beta, \tilde{\boldsymbol{\theta}}, \mathbf{G}) \\ &= \frac{T+1}{2(1-Ts)(1-\varphi)^2}\left[\lambda_1\left(\varphi^2\mathbf{M} + \frac{1+4\varphi^2}{3}\widehat{\mathbf{M}}\right) - \left(\varphi^2\lambda_1(\mathbf{M}_{K,K}) + \frac{1+4\varphi^2}{3}m^*\right)\right] \\ &\leqslant \frac{(T+1)\varphi^2\left[\lambda_1(\mathbf{M}) - \lambda_1(\mathbf{M}_{K,K})\right]}{2(1-Ts)(1-\varphi)^2} \\ &= \frac{(1+s)^2(T+1)\left[\lambda_1(\mathbf{M}) - \lambda_1(\mathbf{M}_{K,K})\right]}{2(1-Ts)[2+T-s(T^2-2)]^2}. \end{aligned}$$

Therefore, $\overline{\mathcal{V}}^*(T, \beta, \overline{\boldsymbol{\theta}}^*, \mathbf{G}) - \overline{\mathcal{V}}^*(T, \beta, \tilde{\boldsymbol{\theta}}, \mathbf{G}) = O(1/T^4)$.

[35] Note that $\mathbf{M}$, and thus its principal submatrix $\mathbf{M}_{K,K}$, is entry-wise positive. Again, the Perron-Frobenius theorem implies $\mathbf{v}_1(\mathbf{M}_{K,K}) \in \mathbb{R}^{|K|}_+$.

References

Armstrong, M., Zhou, J., 2016. Search deterrence. Rev. Econ. Stud. 83, 26–57.
Ballester, C., Calvó-Armengol, A., Zenou, Y., 2006. Who's who in networks, wanted: the key player. Econometrica 74, 1403–1417.
Ballester, C., Calvó-Armengol, A., Zenou, Y., 2010. Delinquent networks. J. Eur. Econ. Assoc. 8, 34–61.
Baron, D., Besanko, D., 1984. Regulation and information in a continuing relationship. Inf. Econ. Policy 1, 267–302.
Battaglini, M., 2005. Long-term contracting with markovian consumers. Am. Econ. Rev. 95, 637–658.
Battaglini, M., Lamba, R., 2019. Optimal dynamic contracting: the first-order approach and beyond. Theor. Econ. 14, 1435–1482.
Bergemann, D., Castro, F., Weintraub, G., 2020. The scope of sequential screening with ex-post participation constraints. J. Econ. Theory 188, 105055.
Bergemann, D., Strack, P., 2015. Dynamic revenue maximization: a continuous time approach. J. Econ. Theory 159 (B), 819–853.
Courty, P., Li, H., 2000. Sequential screening. Rev. Econ. Stud. 67, 697–717.
Crémer, J., Mclean, R., 1985. Optimal selling strategies under uncertainty for a discriminating monopolist when demands are interdependent. Econometrica 53, 345–361.
Crémer, J., Mclean, R., 1988. Full extraction of the surplus in bayesian and dominant strategy auction. Econometrica 56, 1247–1258.
Darby, M., Karni, E., 1973. Free competition and the optimal amount of fraud. J. Law Econ. 16, 67–88.
Denbee, E., Julliard, C., Li, Y., Yuan, K., 2021. Network risk and key players: a structural analysis of interbank liquidity. J. Financ. Econ. 141, 831–859.
Esö, P., Szentes, B., 2007. Optimal information disclosure in auctions and the handicap auction. Rev. Econ. Stud. 74, 705–731.
Galeotti, A., Golub, B., Goyal, S., 2020. Targeting interventions in networks. Econometrica 88, 2445–2471.
Garrett, D., Pavan, A., 2012. Managerial turnover in a changing world. J. Polit. Econ. 120, 879–925.
Halac, M., Yared, P., 2014. Fiscal rules and discretion under persistent shocks. Econometrica 82, 1557–1614.
Jackson, M., 2008. Social and Economic Networks. Princeton University Press, Princeton.
Jadbabaie, A., Kakhbod, A., 2019. Optimal contracting in networks. J. Econ. Theory 183, 1094–1153.
Kakade, S., Lobel, I., Nazerzadeh, H., 2013. Optimal dynamic mechanism design and the virtual-pivot mechanism. Oper. Res. 61, 837–854.
König, M., Liu, X., Zenou, Y., 2019. R&D networks: theory, empirics and policy implications. Rev. Econ. Stat. 101, 476–491.
Krähmer, D., Strausz, R., 2011. Optimal procurement contracts with pre-project planning. Rev. Econ. Stud. 78, 1015–1041.
Laffont, J., Tirole, J., 1990. Adverse selection and renegotiation in procurement. Rev. Econ. Stud. 57, 597–625.
Laffont, J., Tirole, J., 1996. Pollution permits and compliance strategies. J. Public Econ. 62, 85–125.
Lee, L., Liu, X., Patacchini, E., Zenou, Y., 2021. Who is the key player? A network analysis of juvenile delinquency. J. Bus. Econ. Stat. 39, 849–857.
Lewis, T., Sappington, D., 1989. Countervailing incentives in agency problems. J. Econ. Theory 49, 294–313.
Lindquist, M., Zenou, Y., 2019. Crime and networks: 10 policy lessons. Oxf. Rev. Econ. Policy 35, 746–771.
Liu, B., Liu, D., Lu, J., 2020. Shifting supports in Esö and Szentes (2007). Econ. Lett. 193, 109251.
Liu, B., Lu, J., 2018. Pairing provision price and default remedy: Optimal two-stage procurement with hidden R&D efficiency. Rand J. Econ. 49, 619–655.
Lu, J., Wang, Z., 2021. Optimal selling mechanisms with buyer price search. J. Econ. Theory 196, 105307.
Maggi, G., Rodriguez-Clare, A., 1995. On countervailing incentives. J. Econ. Theory 66, 238–263.
Maskin, E., Riley, J., 1984. Monopoly with incomplete information. Rand J. Econ. 15, 171–196.
Meng, D., Tian, G., 2021. The competitive and welfare effects of long-term contracts with network externalities and bounded rationality. Econ. Theory 72, 337–375.
Mierendorff, K., 2016. Optimal dynamic mechanism design with deadlines. J. Econ. Theory 161, 190–222.
Mussa, M., Rosen, S., 1978. Monopoly and product quality. J. Econ. Theory 18, 301–317.
Nelson, P., 1970. Information and consumer behavior. J. Polit. Econ. 78, 311–329.
Pavan, A., Segal, I., Toikka, J., 2014. Dynamic mechanism design: a myersonian approach. Econometrica 82, 601–653.
Ushchev, P., Zenou, Y., 2020. Social norms in networks. J. Econ. Theory 185, 104969.
Wasserman, S., Faust, K., 1994. Social Network Analysis. Methods and Applications. Cambridge University Press, Cambridge.