Influencer Networks*

Yanlin Chen[†] Xueheng Li[‡] Tianle Song[§]

Draft Version: April 10, 2024

Abstract

This paper examines the formation of influencer networks in user-generated content markets. Players decide whether to provide content, its provision level, and whom to follow. Each strict equilibrium network is a nested upward-linking network where different levels of influencers can co-exist, with those at higher tiers providing higher levels of content and being followed by players in all lower tiers. Under a wide range of parameters, all payoff-dominant strict equilibria conform to the law of the vital few: a small yet significant proportion of players provide all the content. However, unlike previous models where the number of influencers is limited and their proportion diminishes rapidly to zero, the complementarity between influence and content provision causes the number of influencers to grow indefinitely with the population. Further, for sufficiently large populations, a single nested upward-linking network connecting all players can emerge, even when players have heterogeneous preferences over content categories.

JEL classifications: D83, D85, Z13.

Keywords: User-generated content networks, network formation, influencers, the law of the vital few.

---

*The three authors contributed equally to this study. Correspondence can be directed to Xueheng Li at Xingang West Road, 135, Sun Yat-sen University, Lingnan College, Guangzhou, China, or emailed to lixueheng@mail.sysu.edu.cn.

[†]Institute for Social and Economic Research, Nanjing Audit University, Nanjing, China. Email: yanlin.chen2019@outlook.com. Yanlin Chen acknowledges financial support from the National Natural Science Foundation of China (No. 72303106).

[‡]Lingnan College, Sun Yat-sen University, Guangzhou, China. Email: lixueheng@mail.sysu.edu.cn. Xueheng Li acknowledges financial support from the National Natural Science Foundation of China (No. 72203105).

[§]Wenlan School of Business, Zhongnan University of Economics and Law, Wuhan, China. Email: tsong@connect.ust.hk.

1 Introduction

Digital and virtual technologies have developed rapidly and relentlessly in recent years. One major product of this development is the rise of digital content and knowledge sharing platforms such as YouTube, TikTok, Quora, and Stack Overflow. These platforms have become deeply embedded in modern life worldwide. For instance, YouTube is estimated to have over 2.5 billion monthly active users in 2022 (Statista, 2023), implying that more than one in four people worldwide actively engage with the platform as either a content provider, a consumer, or both.

This paper examines the choices and network structures arising in these large-scale, decentralized markets for user-generated content. Besides low entry barriers (Burgess & Green, 2018; Cunningham et al., 2016; Mohan, 2022) and the non-rivalry of digital content sharing (Iyer & Katona, 2016), our model captures two distinct features of these social media markets ignored by previous work. First, our model captures the complementarity between influence status and content provision: players with a larger follower base have stronger motivations to provide higher quality of content (Zhang & Zhu, 2011), which in turn attracts more followers (Pagan et al., 2021). This complementarity arises from one of the defining features of the influencer economy (Cong & Li, 2023), namely, providing high level of content to achieve superior influence status can bring significant benefits to content providers both psychologically (Marwick, 2015; Lampel & Bhalla, 2007; Toubia & Stephen, 2013) and financially (Bojkov, 2023; Duffy, 2020; Brown & Freeman, 2022).[1] These strong incentives motivate influencers with a large follower base to provide high quality of content.[2]

Second, our model accommodates the possibility of non-reciprocal relationships. It is found that following and content consumption relationships are often *non-reciprocal* in online sharing networks. For instance, most following relationships on Twitter are non-reciprocal, and the most-followed users typically do not follow many others (Wu et al., 2011). Similarly, influencers on Instagram are unlikely to follow many others, whereas general users on average have a large number of followees (Kim et al., 2017). This low level of reciprocity in online sharing networks contrasts sharply with traditional social networks or trading networks which comprise primarily reciprocal links, and has not been fully accounted for by previous research.

We show that introducing the complementarity between influence and content provision

[1]The top 10 most paid YouTubers in 2021 were estimated to have earned a total of 304.5 million US dollars (Brown & Freeman, 2022).

[2]Throughout the paper, we refer to the quality of content as subjective, judged based on consumers' revealed preferences. We set aside the problem of whether a piece of "high-quality" content ultimately benefits individuals or necessarily obtains the truth.

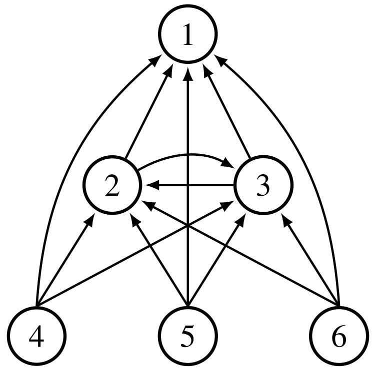

Figure 1: A nested upward-linking network with multiple levels of influencers. Player 1 is the top-level influencer followed by all others and providing the highest level of content. Players 2 and 3 are low-level influencers followed only by the three pure consumers at the bottom tier.

and the possibility of non-reciprocal relationships has significant impacts on network structure, the distribution of influencers, and the influencers' payoffs relative to pure consumers. Moreover, the results provide explanations for some regularities of real-world social media networks that are difficult to be explained by previous models.

The closest work to ours is Galeotti & Goyal (2010). As in Galeotti & Goyal (2010), our model is a game of complete information among n players who simultaneously choose whether to provide content, determine its provision level, and decide whose content to follow.[3] However, we depart from Galeotti & Goyal (2010) in a key aspect: we incorporate the complementarity between influence status and content provision. In our model, players gain benefits from having a larger number of followers. These benefits, in reality, may represent followers' voluntary monetary transfers. Further, as in reality, the transfers increase with the extent to which the provided content fulfills the followers' needs, the benefits from followers are assumed to increase with the content provision level.[4] Consequently, influential players–those with more followers–are motivated to provide a higher level of content, which attracts more followers in equilibrium.

First, the complementarity leads to a multi-level, nested hierarchy: every strict equilibrium network is a *nested upward-linking network* (Proposition 1), such that players endogenously sort into multiple tiers, with those at higher tiers providing higher levels of content and followed by all those at lower tiers. Figure 1 displays a nested upward-linking network. In this

[3]By following a player, we mean actual consumption of the player's provided content.

[4]In the main model, for tractability, we incorporate the benefits obtained from followers into the utility function as a multiplication of the number of followers and the content provision level. In Section 4, we endogenize this complementarity and explicitly model the transfers between players in a network formation model with transfers, with the cost of increased complexity.

network, multiple levels of influencers co-exist. While top-level influencers have a significant number of followers and are incentivized to provide the highest level of content, low-level influencers in the middle tier provide a necessary level of content to attract only the pure consumers at the bottom tier. As players in each tier are followed by all players in each lower tier, players' follower sets are nested and ordered by their content provision levels. Thus, our result provides an explanation for one of the most basic observations regarding social media networks, namely, different levels of influencers can co-exist even when searching for the most prominent influencers involves little difficulty (Bakshy et al., 2011; Cha et al., 2010; Duffy, 2020; Bärtl, 2018). Furthermore, our finding of upward-linking means that players never follow a player in a lower tier. Thus, most links are non-reciprocal, which explains the low reciprocity of online sharing networks (Kim et al., 2017; Wu et al., 2011).

Second, we examine whether the number of influencers can increase steadily with the population or if their proportion inevitably declines. This question is important as the concentration of influence could impact the distribution of benefits derived from social media and affect the concentration and diffusion of (mis-)information (Golub & Jackson, 2010a,b; Yanagizawa-Drott, 2014; Muller & Peres, 2019; Becker et al., 2017; Bakshy et al., 2011; Jackson et al., 2016). Further, the inquiry pertains to the capacity of the influencer market, which is of interest to both individual users aspiring to become influencers and platform providers planning to manipulate the share of influencers to maximize profits.[5]

Unlike the result of previous work, we find that the complementarity between influence and content provision causes the number of influencers to increase systematically with the population. Specifically, we identify two bounds. The first is an *upper bound* on the proportion of influencers over all strict equilibria. As the exogenous marginal benefit parameter of content consumption decreases, this bound decreases and can be arbitrarily small (Proposition 2). However, there also exists a *lower bound* on the proportion of influencers over all payoff-dominant strict equilibria. This lower bound is increasing in the marginal benefit of content consumption and can be strictly positive, *regardless of the population size* (Proposition 3). Thus, as the population grows, the number of influencers also increases indefinitely. This finding contrasts sharply with Galeotti & Goyal (2010). In their model, the number of influencers cannot grow with the population size, leading the share of influencers to rapidly diminish to zero as the population grows regardless of parameters. In contrast, our findings indicate that (i) although influencers may be small in proportion, their absolute number can

[5] Indeed, a set of normative questions from platform providers' point of view and regulators alike can be asked and examined based on our model. However, as a first step, we focus on the set of positive questions in this paper and leave the many interesting normative questions for future research.

increase steadily as the user pool expands, and (ii) the proportion of influencers is likely to vary across platforms and domains in reality, as the marginal benefits from content consumption vary.

Third, we examine how payoffs relate to positions in the network hierarchy. If the marginal benefit of content consumption is small, then payoffs are increasing in positions in the influencer hierarchy: higher-level influencers obtain more than the lower-level ones, and all influencers earn more than pure consumers (Proposition 4). However, if the marginal benefit of content consumption is sufficiently large relative to the marginal benefit of content production, then this pattern is reversed: pure consumers obtain more benefits than influencers, and influencers at lower levels earn more than those at higher levels (Proposition 5).

Last, we consider two extensions to check the robustness of our results. One extension allows for heterogeneous preferences for content categories, akin to education versus gaming channels on YouTube. There we show that each of the following structures can arise in equilibrium: a single nested upward-linking network with some marginal players left out and not served; or a clustering of separated communities, each of which is an independent nested upward-linking network; or, a clustering of nested upward-linking networks with overlapping follower bases. In all cases, nested upward-linking networks serve as the basic building blocks. Moreover, despite the preference heterogeneity, a single nested upward-linking network that connects all players can emerge in equilibrium for large populations (Proposition 6).

The other extension endogenizes the complementarity between influence and content provision by explicitly allowing players to offer and demand monetary transfers to form links. In this extension, a nested upward-linking network with a positive mass of influencers can arise for each sufficiently large population (Proposition 7), confirming our main results.

Related literature This study develops a model to examine the formation and characteristics of user-generated content networks. The model captures three distinct aspects of the markets for user-generated content in reality: low entry barriers, non-rivalry of content sharing, and the complementarity between influence and content provision. In addition, we consider directed networks to allow for non-reciprocal relationships.

The most closely related work to ours is Galeotti & Goyal (2010), which also captures low entry barriers and non-rivalry of content sharing.[6] However, their model ignores the complementarity between influence and content provision. In their model, content is not provided for sharing with others or attracting followers, but to satisfy personal needs. As a

[6]Kinateder & Merlino (2017) extend Galeotti & Goyal (2010)'s model to allow for heterogeneous players.

result, free-riding arises: rather than encouraging influencers to provide better content, having more connections may even discourage content contribution. This leads to their result that the number of influencer is limited independent of the population size, and their proportion diminishes quickly to zero. Furthermore, in their model, every strict equilibrium network is a two-tier, core-periphery network, unless indirect information flow is introduced. In contrast, we show that the complementarity between influence and content provision (i) leads to to a multi-level, nested upward-linking network and (ii) causes influencers to grow systematically with the population.

Our model endogenizes both network links and content provision actions. Besides Galeotti & Goyal (2010), studies that share this feature include Cabrales et al. (2011), König et al. (2014), Baetz (2015), Hiller (2017), Belhaj et al. (2016), Kinateder & Merlino (2017), Ding (2022), Li (2023) and the recent study by Sadler & Golub (2022).[7] However, all these studies, as well as Galeotti & Goyal (2010), restrict to undirected networks, assuming reciprocal bilateral relationships from setup. Thus, they provide no explanation to the low reciprocity in many online sharing networks (Wu et al., 2011, Kim et al., 2017). In a different context, Herskovic & Ramos (2020) consider the formation of information acquisition networks in a beauty contest setting under incomplete information. They identify a similar hierarchical structure similar to our nested upward-linking networks, but differ from ours in that players can sometimes follow one in a lower tier in equilibrium.[8]

Several recent studies (König et al., 2014; Belhaj et al., 2016; Hiller, 2017; Li, 2023; Sun et al., 2023) show that network formation or optimization processes under complementarities result in nested split graphs, which can also be interpreted as multi-level hierarchies. However, in the environments considered by these studies, the incentives to take higher actions and to form links toward others are complements. Consequently, even if players are allowed to differentiate between their inward and outward links, the more central players would end up having more inward links *and* more outward links (Li, 2023). This differs from the results of the current paper, where players endogenously sort into influencers, who possess many inward links but few outward links, and pure consumers, who have many outward links but no inward links.

[7]There is another, less related strand of literature examining stochastically stable equilibrium (Young, 1993; Kandori et al., 1993) in learning processes where agents revise both actions and connections over time in a coordination game with binary actions. Contributors in this literature include Jackson & Watts (2002), Goyal & Vega-Redondo (2005), Staudigl & Weidenholzer (2014), and Cui & Weidenholzer (2021), among others. The primary question in this context is the conditions under which the payoff-dominant equilibrium or the risk dominant equilibrium would be selected in the long run.

[8]See Herskovic & Ramos (2020, page 2148) , Figure 4, for examples.

More broadly, this study contributes to the burgeoning research on the influencer economy and social media platforms in the digital age. On this subject, Iyer & Katona (2016) study influencers' entry decisions in social communication markets and how they compete for receivers' attention in an *exogenous* network. Pagan et al. (2021) introduce a dynamic network formation model in which agents form new links over time based on the principle that agents producing higher levels of content are more likely to be linked. They provide empirical evidence for the complementarity between influence and content provision. However, they assume that each provider's content provision level is *exogenously endowed*. Cong & Li (2023) examine the industrial organization of influencer economy without considering network structure.

Our paper proceeds as follows. Section 2 introduces the model with homogeneous players. Section 3 derives that each strict equilibrium network is a nested upward-linking network and examines the upper and lower bounds on the proportion of influencers. Section 4 examines the two extensions. Section 5 concludes. Appendix A contains all proofs of the formal results presented in the main text. The Online Appendix considers the situation where the marginal benefit from content consumption is particularly large.

2 Model

The game Similar to Galeotti & Goyal (2010), our model is a simultaneous-move game among $n \geq 3$ players. The set of players is denoted by $N = \{1, ..., n\}$. Each $i \in N$ makes two dimensions of choices. The first is the level of content to provide, denoted by $x_i \geq 0$, for each $i \in N$. If i does not provide any content, then $x_i = 0$. The second choice is the set of players to follow, denoted by $g_{ij} \in \{0, 1\}$, where $g_{ij} = 1$ if i follows j and $g_{ij} = 0$ otherwise. By "following," we mean a player actually spending time and attention to consume another player's provided content, thereby incurring costs. We write $g_i = (g_{ij})_{j \in N}$, with $g_{ii} = 0$ for convention. Then the n-by-n matrix $g = [g_{ij}]_{i,j \in N}$ represents the following network. The level of content represents the content's subjective quality based on the crowd's revealed preferences. Other things equal, players prefer following those providing a higher level of content.

A strategy for player $i \in N$ is a pair $s_i = (x_i, g_i)$, and the strategy profile $s = (x, g)$ collects all players' strategies, with $x = (x_1, \ldots, x_n)$ and $g = (g_1, \ldots, g_n)$. The set S contains all strategy profiles. Let $N_i^{in} = \{j \in N | g_{ji} = 1\}$ denote the set of players who follow i, and $N_i^{out} = \{j \in N | g_{ij} = 1\}$ denote the set of players that i follows. Let $d_i^{in} = |N_i^{in}|$ denote the number of i's followers, and $d_i^{out} = |N_i^{out}|$ be the number of players that i follows. They are also called i's indegree and outdegree in the network.

Let $C(\cdot)$ be an increasing and strictly convex cost function. Given all players' strategies, the payoffs for $i \in N$ are given by the utility function

$$u_i(x,g) = \alpha \sum_{j \in N_i^{out}} x_j + \beta d_i^{in} x_i - C\Big(x_i + d_i^{out}\Big), \tag{1}$$

where $\alpha > 0$ and $\beta > 0$ are exogenous parameters. The first term $\alpha \sum_{j \in N_i^{out}} x_j$ reflects the benefits accrued from following other players (by consuming their content). The second term $\beta d_i^{in} x_i$ quantifies the benefits of being followed, derived from, for example, transfers offered by i's followers, which are increasing with content provision level x_i.[9]

Hence, α and β are the marginal returns to following others and being followed, respectively. The third term in the utility function encapsulates the total costs incurred from participating in the network. These costs include those for content production and content consumption. For instance, we can think of $y = x_i + d_i^{out}$ as representing the aggregate time and resources player i commits to the production and consumption of content on the platform, and $C(y)$ the opportunity costs of these resources.[10]

Following the literature, and to derive explicit bounds on the proportion of influencers, we assume a quadratic cost function:

$$C\Big(x_i + d_i^{out}\Big) = \frac{1}{2}\Big(x_i + d_i^{out}\Big)^2.$$

This leads to a linear-quadratic utility function, which is commonly used in the network game literature (e.g., Chen et al., 2018; Belhaj et al., 2016; Zhou & Chen, 2015; König et al., 2014; Calvó-Armengol et al., 2009; Ballester et al., 2006). Adopting this linear-quadratic form is primarily for clarity in exposition; our main qualitative results do not depend on this specific functional form but hold for any increasing and strictly convex cost functions.[11]

For a given strategy profile (x,g), a *provider* is one with $x_i > 0$, and an *influencer* is one with followers. In any equilibrium, a player will be a provider if and only if she has followers. Hence, the terms "provider" and "influencer" are often used interchangeably. A provider may also follow other players. Those players who neither provide content nor have followers are referred to as *pure consumers.*

[9]We endogenize these benefits by explicitly modeling monetary transfers between players in Section 4.2.

[10]While it is possible to introduce additional parameters into the cost function, such as $C(cx_i + kd_i^{out})$, to differentiate the relative costs of producing public goods and following other players, these parameters c and k are redundant given the existence of α and β which already signify the relative benefits of these activities.

[11]Our Proposition 1 is proved directly for a general increasing and strictly convex cost function.

Solution A *(Nash) equilibrium* is a strategy profile $s = (x, g) \in S$ such that for each player $i \in N$, the strategy $s_i = (x_i, g_i)$ maximizes $u_i(s_i, s_{-i})$ given the strategies s_{-i} of other players. If each player's strategy *uniquely* maximizes their payoffs, then s is a *strict equilibrium*.

The network g is an *equilibrium network* if there is $x \in \mathbb{R}^n_+$ such that $s = (x, g)$ is an equilibrium, and it is a *strict equilibrium network* if (x, g) is a strict equilibrium.

A *payoff-dominant equilibrium* $s = (x, g) \in S$ is one such that there is no equilibrium s' with $u_i(s') \geq u_i(s)$ for each $i \in N$ while $u_i(s') > u_i(s)$ for some $i \in N$. If s is both a strict equilibrium and a payoff-dominant equilibrium, then it is a *payoff-dominant strict equilibrium*.

2.1 Discussion

We discuss several features of the model in the following. First, our model exhibits *substitutability* between content production and content consumption, captured by the strictly convex cost function $C(x_i + d_i^{out}) = \frac{1}{2}(x_i + d_i^{out})^2$. This implies that as content provision x_i increases, the marginal cost of following an additional player also increases; conversely, if outdegree d_i^{out} increases, then the marginal cost of content provision rises. Hence, there is a trade-off between content production and consumption, which hints on an endogenous division of content providers and consumers. This trade-off is also present in Galeotti & Goyal (2010) and is key to their core-periphery result.[12]

Second, similar to Galeotti & Goyal (2010), our model assumes (i) free entry for content providers and (ii) non-rivalry in content provision: once provided, there is no additional cost to serve more users. The first feature is captured by that all players are ex ante equal and have the choice to become content providers. Both (i) and (ii) are prominent features of empirical social media networks (Mohan, 2022; Burgess & Green, 2018; Cunningham et al., 2016; Iyer & Katona, 2016).

Third, however, our model diverges from Galeotti & Goyal (2010) in two key respects, allowing us to better capture certain distinct features of online social media networks. One such feature is the significant economic (Bojkov, 2023; Duffy, 2020; Brown & Freeman, 2022) and psychological (Marwick, 2015; Lampel & Bhalla, 2007; Toubia & Stephen, 2013) benefits associated with prominent influencer status. These benefits incentivize influencers to provide higher content level, thereby increasing their attractiveness. This complementarity is missing in Galeotti & Goyal (2010) but captured by our model in the following way. Holding other

[12]In Galeotti & Goyal (2010), the cost term in players' utility functions is linear in content provision x_i and outdegree. This difference between their model and ours is superficial, because they have a strictly concave function encompassing the direct benefits of providing x_i and the benefits of accessing others' content, which implies a similar substitutability between content production and consumption as in our model.

terms constant, let $u_i(x_i, d_i^{in})$ denote player i's payoffs given her content provision decision x_i and indegree d_i^{in}. Then we have

$$\frac{\partial^2 u_i(x_i, d_i^{in})}{\partial x_i \partial d_i^{in}} = \beta > 0,$$

so that if a player has more followers, then she has incentives to provide higher content level. In Section 4.2, we endogenize this complementarity by explicitly modeling the bargaining and monetary transfers between followers and content providers (with the cost of considerable loss of tractability).

Another distinctive feature of online social media networks is that following and content consumption relationships are often *non-reciprocal*. For instance, most following relationships on Twitter are non-reciprocal, and the most-followed users usually do not follow many others (Wu et al., 2011). Similarly, influencers on Instagram are unlikely to follow many others, whereas general users on average have a large number of followees (Kim et al., 2017). Such low reciprocity in online social media networks contrasts sharply with traditional social networks or trading relationships that comprise primarily of reciprocal links. Our model captures the *possibility* of this low reciprocity by considering the formation of *directed* content consumption networks: player i choosing to consume player j's content does not imply that j will necessarily follow back and consume i's content – that is a choice of j. Unlike our model, Galeotti & Goyal (2010) restrict to the formation of undirected networks: if i sponsors a link to j, then automatically a mutual exchange relationship is established between them.[13]

Finally, we focus on characterizing strict equilibria in this paper. Imposing strictness to refine the set of Nash equilibria is common in the literature of noncooperative network formation models (e.g., Kinateder & Merlino, 2017; Baetz, 2015; Galeotti & Goyal, 2010; Bala & Goyal, 2000). One reason is that the set of Nash equilibria in these games often admit too many possibilities, while imposing the simple condition of strictness can sometimes yield much sharper predictions. Another reason is that if a decentralized network system involves individuals who are indifferent among multiple actions, and slight individual change from the status qua would propagate to and eventually change the entire network dramatically, then this network is unlikely to be stable. Indeed, a network is in a steady state of a best-response dynamic or is robust to small heterogeneity or perturbations if and only if it is a strict equilibrium (Kinateder & Merlino, 2017; Baetz, 2015; Bala & Goyal, 2000). When discussing

[13] We agree that this restriction may not be much loss of generality for traditional offline word-of-mouth information exchange networks that are the focus of Galeotti & Goyal (2010).

the plausible range of influencers, we further apply the criterion of payoff dominance.

3 Analysis

3.1 Equilibrium: nested upward-linking networks

If a provider is not followed by anyone, then she has no incentive to provide content, which in turn explains why she has no followers. Therefore, an empty strict equilibrium network always exists, in which no player provides content and no one follows other players. This differs from Galeotti & Goyal (2010), where the purpose of providing content (i.e., acquiring information in their context) is to satisfy private needs rather than to derive benefits from others. In our model, we first show that if the population is large enough, then a non-empty strict equilibrium network must exist, such that some players provide content and are followed by others. We then show that all strict equilibrium networks are nested upward-linking networks, possibly with multiple tiers of influencers.

Definition 1. A network g in a strategy profile (x,g) is a *nested upward-linking network* if it is

1. *nested*: for each $i, j \in N, i \neq j$, we have $N_i^{out} \backslash \{j\} \subset N_j^{out}$ or $N_j^{out} \backslash \{i\} \subset N_i^{out}$; and
2. *upward-linking*: $x_i < x_j$ implies $g_{ij} = 1$ and $g_{ji} = 0$.

In a nested upward-linking network, players are partitioned into $\bar{t} \geq 1$ tiers, $N_1, \ldots, N_{\bar{t}} \subset N$, according to their content provision level. All players with the same x_i are in the same tier. Players with a higher x_i are placed in a higher tier. The upward-linking is in a strong sense: while all players follow all of those in a higher tier, no player follows a single player in a lower tier. In particular, players in tier 1 are pure consumers who follow all players in all tiers $t \geq 2$. Players in the same tier may or may not follow each other. The class of *core-periphery networks* is a special case of nested upward-linking networks; they are nested upward-linking networks with $\bar{t} = 2$ and that players in the second tier, the core players, follow each other (Galeotti & Goyal, 2010). Further, a *periphery-sponsored star* is a core-periphery network with a single player in the second tier.

Here is our first main result:

Proposition 1. *1. A non-empty strict equilibrium network exists if and only if* $2\alpha\beta > \frac{1}{n-1}$.

2. Every strict equilibrium network g is a nested upward-linking network. Furthermore, each strict equilibrium $s = (x,g)$ *exhibits reciprocal links at the top: if* $x_i = x_j$ *and* $g_{ij} = 1$, *then* $g_{k\ell} = 1$ *for each* $k \in N$ *and* $\ell \in N$ *with* $k \neq \ell$ *and* $x_\ell = x_k \geq x_i$.

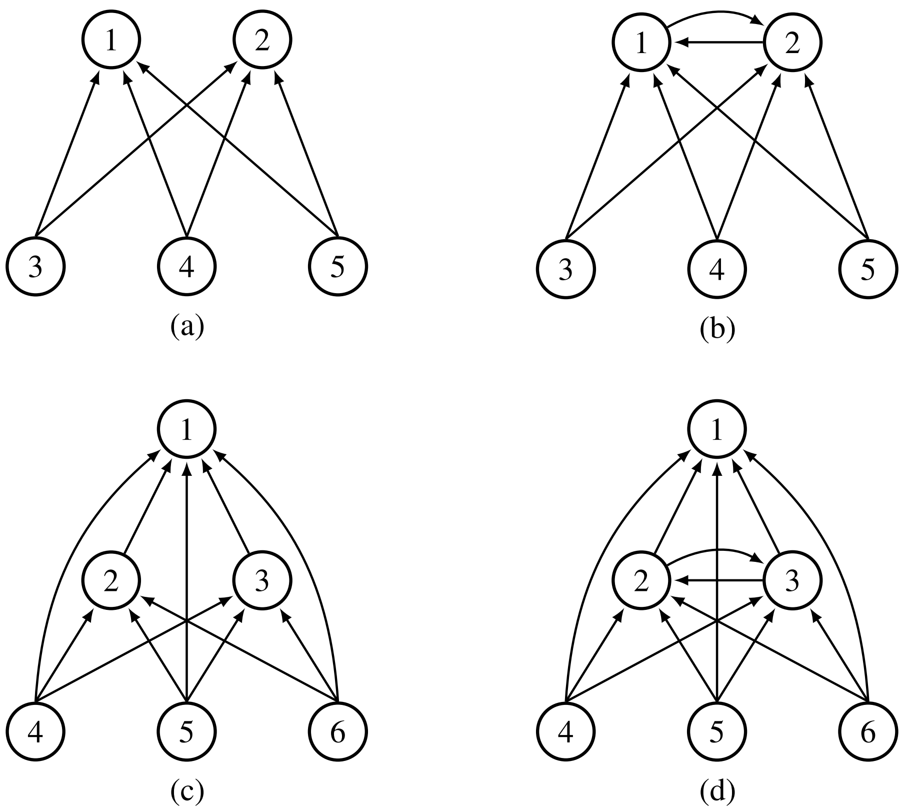

Figure 2: Examples of nested upward-linking networks.

The proposition states that, first, a non-empty strict equilibrium network exists if and only if $2\alpha\beta > \frac{1}{n-1}$. The condition implies that given $\alpha > 0$ and $\beta > 0$, a non-empty strict equilibrium network must exist for a sufficiently large population. This existence condition is obtained by considering a periphery-sponsored star where there is a single provider i followed by all the rest. Given a fixed n, the periphery-sponsored star maximizes the provider's the number of followers and motivates her to provide the highest content provision level possible. Due to the complementarity between influence and content provision, the provider's content level is increasing in n as the number of followers increases, which in turn makes it easier for the rest to follow the provider.

Second, every strict equilibrium network is a nested upward-linking network. Figure 2 provides four examples. Networks (a) and (b) are nested upward-linking networks with two content provision levels, whereas (c) and (d) have three provision levels. In all cases, players at tier 1, the bottom tier, are pure consumers who do not provide content but follow all others in the above tiers. Players in higher tiers are providers with $x_i > 0$ and are influencers with followers. Networks (a) and (b) contain a single tier of influencers who provide the same level of content and have the same number of followers. In contrast, networks (c) and (d) contain two tiers of influencers, who provide different levels of content and have different numbers of followers. Among the examples, only (b) is a strictly defined core-periphery network (Galeotti & Goyal, 2010).

Furthermore, the contrast between (a) and (b), and that between (c) and (d), in Figure 2 illustrates that influencers within a tier may or may not follow each other. However, Proposition 1 imposes a restriction on the reciprocal relationships: that influencers follow each other within a tier can only occur in the upper part of the network. This result explains why on Instagram influencers with many common followers are found to frequently follow each other, while the whole network exhibits low reciprocity (Kim et al., 2017).

In what follows, we explain how combining (i) the complementarity between influence status and content provision and (ii) the substitutability between content provision and content consumption leads to the multi-level, nested upward-linking structure.

First, to decide whom to follow and the optimal number of followees d_i^{out}, each player ranks the providers according to their content provision level and choose providers from high to low down the list. This selection principle, combined with the substitutability between content production and consumption, implies that players with a higher content provision level follow fewer providers, and players' followee sets N_i^{out} are nested.

Lemma 1. *In every strict equilibrium, if $x_i \geq x_j$, then $d_i^{out} \leq d_j^{out}$ and $N_i^{out} \backslash \{j\} \subset N_j^{out}$.*

The reason is the following. When a player follows an additional provider, she obtains the additional benefit of consuming that additional provider's content, but also bears the additional cost of spending time and attention for the consumption. As a player follows more providers down the ranking list, the additional benefit of following one more provider decreases, while the marginal cost of linking increases. The optimal number of followees d_i^{out} is then determined by the point where the marginal benefit crosses the marginal cost. Given the convexity of the cost function, as a player's content provision x_i increases, the marginal cost of following an additional player increases. As a result, a player's outdegree d_i^{out} is decreasing in her content provision x_i. Moreover, the followee set N_i^{out} includes precisely the top d_i^{out} providers among the provider ranking list. Since all players agree on the same ranking of providers, their followee sets N_i^{out} must be nested.

Second, due to the complementarity between x_i and d_i^{in}, players with more followers provide higher levels of content. Furthermore, the selection of providers from high to low implies that the converse is also true: higher-ranked providers must attract more followers. Consequently, a player has more followers if and only if she provides a higher level of content.

Lemma 2. *In every strict equilibrium, $x_i \geq x_j$ if and only if $d_i^{in} \geq d_j^{in}$.*

Thus, the complementarity between influence status and content provision leads to the emergence of providers at various levels. Top influencers, with the greatest number of followers, are incentivized to offer the highest level of content, which justifies the massive number of followers. Mid-tier influencers, having fewer but still many followers, provide a moderate level of content. Low-tier influencers provide a lower level of content to support and attract a relatively small number of followers, although it can still be a considerable number in a large network.

Therefore, Lemmas 1 and 2 together imply a nested hierarchy of influencers atop a base group of pure consumers. What remains to show is the strong upward-linking property for all strict equilibria: each player i must follow all those with $x_j > x_i$ and never follow a single player with $x_j < x_i$. This outcome does not follow immediately from the two lemmas. To explain this outcome, index players by $x_1 \leq \ldots \leq x_n$, and suppose there are some players i and j such that j follows i while $x_i < x_j$. Consider the highest-indexed player j among those ever following a lower-indexed player i with $x_i < x_j$. Then we can show that Lemmas 1 and 2 imply a contradiction. The intuition is that, on the one hand, if j follows i, then by Lemma 1, so must all the players $1, 2, \ldots, j-1$, except i herself. Therefore, player i must have at least $j-1$ followers. On the other hand, however, player j is already the highest ranked player among those ever following a lower ranked player. Therefore, no player $k > j$ follows j. Thus, player j's indegree is at most $j-1$. Thus, $d_j^{in} \leq j-1 \leq d_i^{in}$. But then, by Lemma 2, this implies $x_i \geq x_j$; thus, a contradiction is established. This shows that the effects represented by Lemmas 1 and 2 would accumulate and eventually lead to a strong restriction on strict equilibrium networks.

3.2 Proportion of influencers

Next, we examine whether influencers can consistently grow alongside the population or if their proportion inevitably declines. This question is related to the degree of concentration of influence within larger populations and has implications for the distribution of benefits from social media and the spread and containment of (mis-)information in a broader context.[14]

For a strategy profile $s = (x, g)$, we denote by $\rho(s) = \frac{1}{n}|\{i \in N | x_i > 0\}|$ the proportion of influencers, which is equivalent to $\frac{1}{n}|\{i \in N | d_i^{in} > 0\}|$ in any strict equilibrium.[15] We quantify

[14] See discussions in Golub & Jackson (2010a,b); Yanagizawa-Drott (2014); Muller & Peres (2019); Becker et al. (2017); Bakshy et al. (2011); Jackson et al. (2016).

[15] Proposition 3 remains valid if alternatively we define influencers as those with content provision levels higher

the plausible range of $\rho(s)$ by driving an upper bound of $\rho(s)$ over all strict equilibria and a lower bound of it over all payoff-dominant strict equilibria. We are interested in the proportion of influencers relative to pure consumers because this, as we show in the next subsection, affects inequality among players via their network positions.

First, observe that holding n fixed, if the marginal benefit of content consumption, α, and that of accumulating followers, β, are large enough, then there exists a complete strict equilibrium network such that all players follow each another and all of them provide content at the same sufficiently high level to support the followers. In such a network, $\rho(s) = 1$. In the Appendix, we show that if $\alpha > 1$ and β is sufficiently large, then $\rho(s) = 1$ is attainable for each $n \geq 3$. Here, we focus on the more interesting case of $\alpha \leq 1$.[16]

Proposition 2. *If* $\alpha \leq 1$, *then* $\rho(s) < \rho^H \equiv \frac{\alpha\beta}{\alpha\beta+1} + \frac{1}{(\alpha\beta+1)n}$ *for each strict equilibrium* $s \in S$.

Proposition 2 states that if $\alpha \leq 1$, then the proportion of influencers is smaller than ρ^H, which is strictly below one. Note that although the bound decreases with n, it does not diminish to zero. Instead, the limit

$$\lim_{n\to\infty} \rho^H = \frac{\alpha\beta}{\alpha\beta+1}$$

is continuous and increasing in $\alpha\beta$, and can take any value from zero to one as $\alpha\beta$ increases from indefinitely small to infinity.

Second, as noted before, an empty strict equilibrium network always exists. Thus, there is a trivial lower limit of $\rho(s)$ over all strict equilibria, which is zero. However, whenever a non-empty strict equilibrium network exists, the empty network can only arise from a coordination failure. This is because there is an alternative non-empty network, such as a periphery-sponsored star, in which everyone can be strictly better off. Thus, at least in the context of social media networks, strategy profiles that fail to be a payoff-dominant equilibrium due to insufficient activity levels like the empty network might be unstable.[17] Therefore, we instead provide a lower bound of $\rho(s)$ over the set of payoff-dominant strict equilibria, to show that the complementarity between influence and content provision may drive influencers

than a given threshold which can be even increasing in n. For instance, consider the alternative definition of influencers as those with $x_i > bn$ for a given constant $b > 0$. We provide the proof for Proposition 3 in Appendix A for this alternative, more stringent, criterion of influencers.

[16]An immediate consequence of $\alpha \leq 1$ is that players would not follow another with the same x_i; see Lemma 4 for details. This result combined with Proposition 1 implies that there are no reciprocal links.

[17]In real-world social media networks, the users who aspire to become influencers, anticipating the future benefits, would initiate content provision before they have accumulated a large number of followers. And to survive and earn profit, platform providers would take various measures to activate and promote connection and content provision activities. Indeed, it is commonly observed that the platform providers frequently use intelligent algorithms to strategically push appealing content to individual users.

to grow systematically with the population. As Proposition 2, we focus on the case of $\alpha \leq 1$. Results for $\alpha > 1$, which are similar but with additional details to consider, are discussed in our Online Appendix.

Proposition 3. *If $\alpha \leq 1$ and $\beta > \frac{1-\alpha}{\alpha^2}$, then there exists $n^* \geq 3$ such that, for each $n \geq n^*$, we have*

$$\rho(s) \geq \rho^L \equiv \frac{\alpha^2\beta^2 + \alpha\beta - \beta}{\alpha^2\beta^2 + \alpha^2\beta + 2\alpha\beta + \alpha}$$

for each payoff-dominant strict equilibrium $s \in S$.

The existence of the lower bound $\rho^L > 0$ immediately implies that for each large enough n, there is a strict equilibrium in which the proportion of influencers is at least ρ^L rather than diminishing in the limit. Furthermore, based on payoff-dominant strict equilibria, influencers not only can increase, but they *must* increase, without bound as the population grows. This finding is in stark contrast to Galeotti & Goyal (2010)'s result of "the law of the few." In their model, the absolute number of influencers is bounded above independent of population size; thus, the proportion of influencers rapidly diminishes to zero as the population grows. While their result provides a succinct explanation for the relatively small number of influencers observed in many empirical networks, we fail to find any evidence in the context of social media networks suggesting that the number of influencers is fixed and cannot grow with the population. By contrast, statistics indicate that influencers have increased significantly over the past few years as the number of users grows on platforms such as YouTube and Bilibili (Scorus, 2021; Funk, 2020). More importantly, we consider Galeotti & Goyal (2010) inadequate in understanding online social media networks due to its neglect of the complementarity between influence and content provision. This complementarity, as we explain immediately, drives influencers to increase systematically with the population.

The intuition behind Proposition 3 is the following: the complementarity between d_i^{in} and x_i allows all providers' content provision x_i to increase as the population grows, which enables all pure consumers to follow more providers. This leads to the expansion of providers, increasing the activity levels of all players and making everyone better off. Thus, as the population grows, influencers must also increase among the payoff-dominant equilibria.

More precisely, due to the complementarity, a provider's content level

$$x_i = \beta d_i^{in} - d_i^{out}$$

increases with indegree d_i^{in}. Consider a sequence of strict equilibrium networks with a fixed number of providers, k, and a growing number of population size, n. Then the number of pure consumers, $n-k$, also increases. By Proposition 1, all providers are followed by all pure consumers, leading to an increase in content provision. It follows that, from a sufficiently large n^* onward, there is room to move up a fraction of players, say, $\rho^L(n-n^*-k)$ players, among the increased pure consumers to be a new set of providers followed by the remaining pure consumers. The new providers continue to follow all providers they initially followed. In this way, the number of providers is increased to $k+\rho^L(n-n^*-k)$ in the modified network. Moreover, if β is large enough, then the modified profile remains a strict equilibrium. This is because as all providers' content provision increases, the marginal benefits of following more providers increase. Thus, the remaining pure consumers are happy with following more providers, and the new set of providers are happy with providing content while maintaining their links to the old providers. As n grows, the fraction of providers in the modified networks approaches $\lim_{n\to\infty}\frac{1}{n}(k+\rho^L(n-n^*-k)) = \rho^L$.[18]

Further, all players are better off, and some strictly so, in the modified network. Consider them one by one: (i) the remaining pure consumers in the modified network; (ii) the players who are providers in both networks, i.e., the old providers; and (iii) the new providers. The first group's revealed preferences for following more providers, combined with strictly convex linking costs, suggest that they must be strictly better off. The old providers' indegree, outdegree, and content provision levels are unchanged. Thus, they are at least as well off. Finally, the new providers could have chosen to provide zero content while enjoying the additional benefits of being followed by the remaining pure consumers, but choose to provide positive content, revealing that they must be strictly better off. Note that if $\beta = 0$, i.e., there is no complementarity between influence and content provision, then the above outcome would not occur even with private motivations for providing content (e.g., acquiring information for personal use), because in that case content provision would not be increase with more followers.

3.3 Payoffs and network positions

It may be conjectured that payoffs increase with one's position in the influencer hierarchy: influencers in higher tiers receive greater payoffs than those in lower tiers, and even the lowest-tier influencers earn more than pure consumers. This is not always true; the opposite can occur.

[18] That is, for each $n > n^*$ and the initial strict equilibrium considered given n, we obtain a modified network with $k+\rho^L(n-n^*-k)$ providers. The limit fraction of providers for the sequence of the modified networks is ρ^L.

Consider strict equilibrium networks with $\bar{t} \geq 2$ tiers. The set of players in tier $t \in \{1, \ldots, \bar{t}\}$ is denoted by $N_t \subset N$, and let $n_t = |N_t|$ denote the number of players in tier t. Also, let $\bar{n}_t = \sum_{t'=1}^{t} n_{t'}$ the total number of players up to tier t. According to Proposition 1, players in the same tier are completely symmetric; their indegree, outdegree, and content provision are all the same. Thus, their payoffs are the same. Given a strict equilibrium $s = (x, g)$, let $u_{[t]}$ denote the payoff number of players in tier t, i.e., $u_i(s) = u_{[t]}$ for each $i \in N_t$.

First, we present the condition under which payoffs increase with one's position in the hierarchy.

Proposition 4. *Consider a strict equilibrium with $\bar{t} \geq 2$. If $\alpha < 1$, then $u_{[t]} > u_{[t-1]}$ for each $t \in \{2, \ldots, \bar{t}\}$.*

Proposition 4 shows that given $\alpha < 1$, the individual payoff is strictly increasing with tier level. This is because $\alpha < 1$ implies that the marginal benefit from content consumption, α, is small relative to the marginal benefit from being followed, β. Therefore, a tier-t player who, in equilibrium, has more followers than a player in tier $t-1$, can earn greater payoffs than the latter. However, the following proposition implies that if $\alpha \geq 1$, then the monotonicity may not hold, and the higher-the-better-off outcome may even be reversed.

Proposition 5. *Consider a strict equilibrium network with $\bar{t} \geq 2$. If $\alpha \geq 1$, then for each $t \in \{2, \ldots, \bar{t}\}$, there exists a threshold $\hat{\alpha}(\beta, n, n_t, n_{t-1}, \bar{n}_{t-1}) \in \mathbb{R}$ such that $u_{[t]} \leq u_{[t-1]}$ if $\alpha \geq \hat{\alpha}(\beta, n, n_t, n_{t-1}, \bar{n}_{t-1})$, and $u_{[t]} > u_{[t-1]}$ otherwise.*

Proposition 5 implies that the payoff of a player may be non-monotone in the tier level when $\alpha \geq 1$. Specifically, whether players in a higher tier obtain more payoffs depends on the marginal benefit parameters α and β. To illustrate this outcome, consider the periphery-sponsored star where player 1 is the sole provider. In strict equilibrium, player 1's provision is $x_1 = \beta(n-1)$, and her payoff is

$$u_{[2]} = \beta(n-1)x_1 - \frac{1}{2}x_1^2 = \frac{1}{2}x_1^2.$$

In contrast, a pure consumer's payoff is

$$u_{[1]} = \alpha x_1 - \frac{1}{2}.$$

Compared with a pure consumer, the provider has more followers but no other providers to follow. Thus, as revealed by the above payoff equations, we observe that a pure consumer's payoffs increase with α, while the provider's payoffs increase with the content provision level x_1 which is in turn increasing in β. Hence, if β is held constant and α is sufficiently large, then a consumer's payoff can surpass that of the provider. The intuition is simple. We are comparing not the provider's advantage from being followed with the consumers' utility gains from consumption but rather their *net benefits*. Although the provider benefits from having more followers, they also put significant effort into content production. Consequently, if the utility gains from consuming content are substantial, then consumers may earn greater net benefits. This intuition extends to equilibrium networks with multiple tiers. Players in higher tiers have more followers but fewer providers to follow, thereby benefiting more from content production but less from content consumption. Consequently, if the enjoyment derived from content consumption is sufficiently large, players in lower tiers may obtain more net benefits.

4 Extensions

4.1 Horizontal differentiation

This subsection examines an extension of the model with heterogeneous preferences over content types, akin to YouTube channels on cooking versus gaming. The model is adapted from Chen et al. (2018); Kor & Zhou (2023).

Suppose that there are two goods, A and B, and players have different preferences over them. For example, A and B represent two categories of YouTube channels. Simultaneously, each $i \in N$ makes three choices: which good to provide, $\omega_i \in \{A, B\}$; the content provision level, $x_i \geq 0$; and the subset of players to follow, $N_i^{out} = \{j \in N | g_{ij} = 1\}$. Player i's strategy is $s_i = (\omega_i, x_i, g_i)$, and the strategy profile for all players is (Ω, x, g), where $\Omega = (\omega_1, \ldots, \omega_n)$. Given a profile (Ω, x, g), the utility for $i \in N$ is

$$u_i(\Omega, x, g) = \alpha \theta_i^A \left(\sum_{j \in N_i^{out}, \omega_j = A} x_j \right) + \alpha \theta_i^B \left(\sum_{j \in N_i^{out}, \omega_j = B} x_j \right) + \beta d_i^{in} x_i - \frac{1}{2} \left(x_i + d_i^{out} \right)^2,$$

where $\theta_i^A \geq 0$ and $\theta_i^B \geq 0$ are individual preferences parameters for the two goods.

We further assume that players are of two types, such that the set of players N is partitioned into two subsets of equal size, $N_A = \{1, \ldots, \frac{n}{2}\}$ and $N_B = \{\frac{n}{2} + 1, \ldots, n\}$, and that $n \geq 4$ is an even number. Consider a discount parameter $\delta \in (0, 1)$. We assume $\theta_i^A = 1$ and $\theta_i^B = 1 - \delta$ for each $i \in N_A$, and $\theta_i^A = 1 - \delta$ and $\theta_i^B = 1$ for each $i \in N_B$. Hence, other things equal, players in N_A prefer good A over good B, players in N_B prefer good B over good A. In a strict equilibrium, each player plays a unique best-response strategy (ω_i, x_i, g_i) given other players' strategies.

First, we present an example to show that a richer and more realistic set of networks can arise in this more complex environment. However, nested upward-linking networks remain the basic building blocks to construct the equilibrium networks.

Example 1. Suppose $n = 8$. Then networks (a) and (b) displayed in Figure 3 are strict equilibrium networks when $\alpha\beta > \frac{1}{6}$ and $0 < 1-\delta < \min\{\frac{1}{6\alpha\beta}, \frac{1}{\alpha}\}$. Network (c) is a strict equilibrium network when $\alpha\beta > \frac{1}{2}$, $\beta > \frac{1}{2}$, and $\frac{1}{4\alpha\beta} < 1-\delta < \min\{\frac{1}{2\alpha\beta}, \frac{1}{2\alpha}\}$.

Network (a) consists of a single nested upward-linking network that covers only part of the population. Player 1, followed by all players in N_A except herself, provides $x = 3\beta$ of good A, while all players in N_B are isolated and have no content of their type to consume. Network (b) contains two separated communities, each of which is a nested upward-linking network. Network (c) consists of two nested upward-linking networks with overlapping followers. Player 1, followed by only players in N_A, is a *local influencer*. In contrast, player 8 is a *global influencer* who provides a higher level of content and is followed by players in N_A as well as players in N_B.

Next, we show that despite the preference heterogeneity, if n is large enough, then a nested upward-linking network that connects all players can always be a strict equilibrium network. In such a network, every player participates as a provider or a consumer, and all consumers in the same tier follow the same set of providers, despite their heterogeneous preferences over content categories.

Proposition 6. *If $2\alpha\beta(1-\delta) > \frac{1}{n-1}$, then there exists a strict equilibrium (Ω, x, g) such that g is a nested upward-linking network without isolated agents.*

Proposition 6 provides a sufficient condition for a nested upward-linking network that connects all players to be a strict equilibrium network. It implies that given $\alpha > 0$, $\beta > 0$ and $\delta > 0$, an integrated market can always arise if the user pool is large, which aligns with Proposition 1. However, the discount parameter δ also plays a role here. If δ is sufficiently close to one, i.e., players' tastes are sufficiently polarized, then an integrated market may not exist, because players would not want to follow providers who provide a different category of content.

4.2 Monetary transfers

So far, players directly gain utility from having followers. Now, we consider an extension where players explicitly propose monetary offers or demands to each other to form a network.

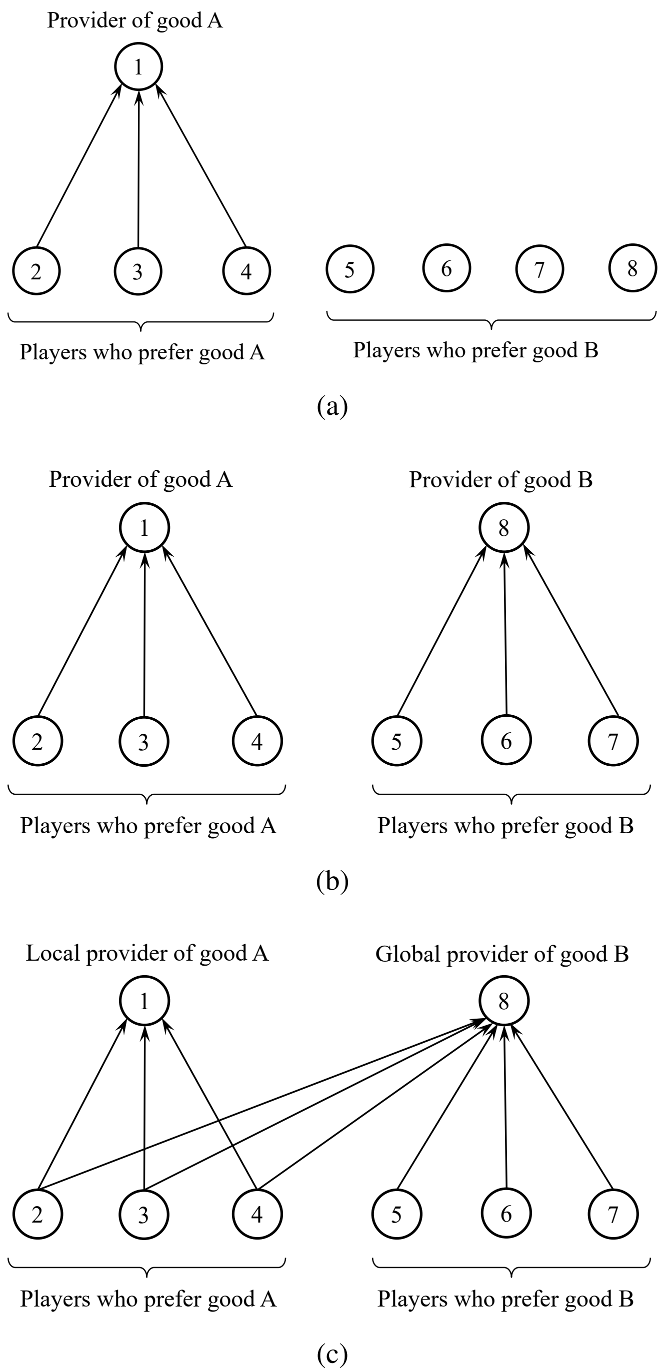

Figure 3: Networks that can emerge under heterogeneous preferences over content categories.

We show that a nested upward-linking structure and a positive mass of influencers can still arise.

This network formation model with transfers is a combination and adaptation of Galeotti & Goyal (2010) and Bloch & Jackson (2007).[19] There are two stages. In stage 1, each $i \in N$ chooses the public goods provision, $x_i \geq 0$. In stage 2, each $i \in N$ observes $x = (x_1, \ldots, x_n)$ and proposes a vector of transfers $\tau_i = (\tau_{ij})_{j \in N, j \neq i} \in \mathbb{R}^{n-1}$. The proposal τ_{ij} can be positive or negative. When $\tau_{ij} \geq 0$, player i offers transfers to j. When $\tau_{ij} < 0$, player i demands transfers from j. A directed link from i to j is formed if and only if $\tau_{ij} + \tau_{ji} \geq \eta$ and $\tau_{ij} \geq \eta$, where $\eta > 0$ is the transaction cost and can be arbitrarily small. This transaction cost may represent, for example, service and digital facility prices charged by the platform provider. In particular, if and only if $\tau_{ij} \geq \eta$ and $\tau_{ji} \geq \eta$, then $g_{ij} = g_{ji} = 1$. Let $\tau = (\tau_1, \ldots, \tau_n)$ be the transfer profile of all players and $\mathbb{T}$ be the set of all transfer profiles. Let $g(\tau)$ denote the network induced by $\tau \in \mathbb{T}$, with $g_{ii}(\tau) = 0$ for each $i \in N$. Let $\bar{g}_{ij}(\tau) = \max\{g_{ij}(\tau), g_{ji}(\tau)\}$. Let d_i^{out} be i's outdegree in $g(\tau)$. The game ends after Stage 2 and each $i \in N$ receives

$$u_i(x, \tau) = \alpha \sum_{j \in N} g_{ij}(\tau) x_j - C\left(x_i + d_i^{out}\right) - \sum_{j \in N} \bar{g}_{ij}(\tau) \tau_{ij}.$$

In this two-stage game, each $x \in \mathbb{R}^n_+$ induces a different subgame in stage 2.[20] A strategy for player i is a pair $(x_i, \hat{\tau}_i)$, where $\hat{\tau}_i(x) \in \mathbb{R}^{n-1}$ specifies i's transfers in stage 2 given $x = (x_1, \ldots, x_n)$, for each $x \in \mathbb{R}^n_+$.[21] Let $\hat{\tau} = (\hat{\tau}_1, \ldots, \hat{\tau}_n)$. Let $(x, \hat{\tau})$ be the strategy profile of all players and S be the set of all strategy profiles.

Definition 2. Given a provision profile $x \in \mathbb{R}^n_+$, a transfer profile $\tau \in \mathbb{T}$ is *pairwise stable* if for each $i \neq j$, there are no $\tau_i', \tau_j' \in \mathbb{R}^{n-1}$ such that $u_i(x, \tau_{-ij}, \tau_i', \tau_j') \geq u_i(x, \tau)$, $u_j(x, \tau_{-ij}, \tau_i', \tau_j') \geq u_j(x, \tau)$, and one of the inequalities is strict.

A strategy profile $(x, \hat{\tau}) \in S$ is a *pairwise stable SPE (subgame perfect equilibrium)* if (i) $(x, \hat{\tau})$ is a subgame perfect equilibrium, and (ii) $\hat{\tau}(x)$ is pairwise stable for each $x \in \mathbb{R}^n_+$.

For each $x_i \geq 0$ and outdegree $d \geq 1$, denote by $\Delta C(x_i, d) = C(x_i + d) - C(x_i + d - 1)$ the increase in costs of following one more player. Then for each provision profile $x \in \mathbb{R}^n_+$, we

[19] Galeotti & Goyal (2010) assume a fixed linking cost and that it is shared equally between the linked players. Bloch & Jackson (2007)'s model has no content provision.

[20] We use $\mathbb{R}^n_+$ to denote the set of nonnegative n-tuples.

[21] This two-stage game is equivalent to a one-stage game in which each player simultaneously announces the public goods provision x_i and proposes a set of *contingent* contracts about transfers $\tau_i(x) \in \mathbb{R}^{n-1}$ that depend on the profile $x = (x_1, \ldots, x_n)$.

can construct a class of networks $G^*(x)$ that feature nestedness and upward-linking, where each $g^* \in G^*(x)$ is obtained by the following procedure:

Step 1: Reindex the players so that $x_1 \geq x_2 \geq \cdots \geq x_n$.

Step 2: Define a critical integer d_i^* for each $i \in N$ as follows, which determines i's outdegree. If $\alpha x_d \geq \Delta C(x_i, d-1) + \eta$ for some $d \in \{1, \ldots, n\}$ and $d > i$, then let $d_i^* \in \{1, \ldots, n\}$ be the largest number such that the inequality holds. If $\alpha x_d < \Delta C(x_i, d-1) + \eta$ for all $d \in \{1, \ldots, n\}$, $d > i$, then let d_i^* be the largest number $d \in \{1, \ldots, n\}$ with $d < i$ such that $\alpha x_d \geq \Delta C(x_i, d) + \eta$. If $\alpha x_d < \Delta C(x_i, d) + \eta$ for each $d < i$, then let $d_i^* = 0$.

Step 3: For each $i, j \in N$, $i \neq j$, we set $g_{ij}^* = 1$ if and only if $j \leq d_i^*$, i.e., $N_i^{out} = \{j \in N | j \leq d_i^*, j \neq i\}$. Then, for players with $i > d_i^*$ we have $d_i^{out} = d_i^*$, and for those $i < d_i^*$ we have $d_i^{out} = d_i^* - 1$. This completes the procedure.

By the above procedure, we obtain at least one such g^* for each $x \in \mathbb{R}_+^n$. Since reindexing players in Step 1 is not necessarily unique, the constructed g^* admits multiple possibilities. Thus, $G^*(x)$ need not be singleton. Note that if $x_i > x_j$, then in each $g^* \in G^*(x)$ we have $d_i^{in} \geq d_j^{in}$, $d_i^{out} \leq d_j^{out}$, $N_i^{out} \backslash \{j\} \subset N_j^{out}$, and $N_j^{in} \backslash \{i\} \subset N_i^{in}$. Hence, all networks in $G^*(x)$ exhibit nestedness. Let $S_{\alpha,n}^* \subset S$ denote the set of all pairwise stable SPEs given α and n.

Proposition 7. *1. For each $x \in \mathbb{R}_+^n$ and $g^* \in G^*(x)$, there exists $\tau \in \mathbb{T}$ such that $g(\tau) = g^*$ and τ is pairwise stable and a Nash equilibrium in the stage 2 subgame.*

2. Assume $0 < \alpha \leq 1$. Then there exist $n^* \geq 3$ and a sequence of pairwise stable SPEs

$$(x, \hat{\tau})_{n^*}, (x, \hat{\tau})_{n^*+1}, (x, \hat{\tau})_{n^*+2}, \ldots,$$

with $(x, \hat{\tau})_n \in S_{\alpha,n}^*$ for each $n \geq n^*$, such that

i) $g(\hat{\tau}(x))$ is a nested upward-linking network, and

ii) $\lim_{n \to \infty} \rho_n = 1 - \sqrt{\frac{1}{1+\alpha^2}}$, where ρ_n is the proportion of influencers induced by $(x, \hat{\tau})_n$.

The proposition presents two findings. First, for any given provision profile $x \in \mathbb{R}_+^n$, there exists an equilibrium transfer profile $\tau \in \mathbb{T}$ in the second-stage subgame that leads to a network characterized by nestedness and upward-linking. Second, there exists a sequence of pairwise stable SPEs where as $n \to \infty$, the proportion of influencers approaches a limit that is strictly positive and increases with $\alpha \in (0, 1]$. This immediately implies that the upper limit of the proportion of influencers over all pairwise stable SPEs must be strictly positive.

5 Conclusion

The advancement of digital and virtual technologies has led to the continued rise of large-scale user-generated content markets on online platforms, where an ever-growing range of creation, information acquisition, and consumption activities occur. The emergence of these markets poses novel, intriguing, and significant questions for researchers and regulators. These questions include how influencers, the "vital few," emerge from the decentralized decisions of countless individual users; what network topologies to expect; and the welfare implications.

Our study contributes to the discourse on these issues in two ways. First, building on the work of Galeotti & Goyal (2010), we develop a simple model that encapsulates the three prominent features of online user-generated content markets: low entry barriers, non-rivalrous content consumption, and the complementarity between influence and content provision. Second, we use the model to demonstrate that (i) nestedness and upward linking are the primary features of user-generated content networks; (ii) a multi-level hierarchy of influencers can arise; (iii) while influencers may be small in proportion, they can increase indefinitely and proportionally with the population; and (iv) while an increase in the user base tends to benefit everyone, pure consumers can sometimes reap greater utility gains than influencers.

To conclude, our study elucidates the emergence of influencers and network structures in online user-generated content markets, setting the stage for future research into their boarder consequences and policy implications. For instance, we have not addressed the incentives and policies of platform providers, who may choose particular compensation schemes for content providers, and charge or subside participants based on their network positions. These choices of platform providers are likely to impact the network structure and the distribution of different levels of influencers, all of which warrant future research.

Appendix A: Proofs

Notations. In the proofs, when the argument is applicable to the general cost function $C(\cdot)$ with $C'(y) > 0$ and $C''(y) > 0$, we use $C(x_i + d_i^{out}) = \frac{1}{2}(x_i + d_i^{out})^2$ to denote the cost function and $D(\cdot) \equiv C'^{-1}(\cdot)$ denote the inverse of its first-order derivative. In addition, from the proofs of Proposition 2 onward, given a strict equilibrium $s = (x, g)$ – which is a nested upward-linking network according to Proposition 1, we use $x_{[t]}, d_{[t]}^{out}, d_{[t]}^{out}$, and $u_{[t]}$ to denote the content provision level x_i, indegree d_i^{in}, outdegree d_i^{out}, and payoffs u_i for each player in tier t within the network g, respectively. This is possible because, according to Proposition 1, all players in the same tier have the same x_i, d_i^{in}, d_i^{out}, and u_i in a strict equilibrium network.

Proof of Lemma 1

Consider a strict equilibrium. Without loss of generality, suppose that $x_1 \geq x_2$ and $d_1^{out} > d_2^{out}$ for contradiction. First, suppose that $N_1^{out} \backslash \{2\} \subset N_2^{out} \backslash \{1\}$ does not hold, so that there is some $i \neq 1, 2$ such that $i \in N_1^{out}$ and $i \notin N_2^{out}$. Then, holding all other things fixed, player 1's utility of linking to i, $u_1(g_{1i} = 1)$, is strictly greater than the utility of deleting the link, $u_1(g_{1i} = 0)$. We have $u_1(g_{1i} = 1) > u_1(g_{1i} = 0)$ if and only if $\alpha x_i > C(d_1^{out} + \frac{1}{2}) - C(d_1^{out} - 1 + \frac{1}{2})$. But then, by the convexity of $C(\cdot)$ and given $d_1^{out} > d_2^{out}$, we have

$$\begin{aligned}
\alpha x_i &> C(d_1^{out} + \frac{1}{2}) - C(d_1^{out} - 1 + \frac{1}{2}) \\
&\geq C(d_2^{out} + 1 + \frac{1}{2}) - C(d_2^{out} + 1 - 1 + \frac{1}{2}) \\
&= C(d_2^{out} + 1 + \frac{1}{2}) - C(d_2^{out} + \frac{1}{2}).
\end{aligned}$$

It follows that $u_2(g_{2i} = 1) > u_2(g_{2i} = 0)$. Therefore, player 2 has strict incentives to deviate to linking to i, a contradiction. Hence, it cannot be both $d_1^{out} > d_2^{out}$ and that $i \in N_1^{out}$ and $i \notin N_2^{out}$ for some $i \neq 1, 2$.

Next, suppose $d_1^{out} > d_2^{out}$ and $N_1^{out} \backslash \{2\} \subset N_2^{out} \backslash \{1\}$. Then, it must be $d_1^{out} = d_2^{out} + 1$, $g_{12} = 1$ and $g_{21} = 0$. That player 1 linking to player 2 implies $u_1(g_{12} = 1) > u_1(g_{12} = 0)$, leading to $\alpha x_2 > C(x_1 + d_1^{out}) - C(x_1 + d_1^{out} - 1)$. However, $x_1 \geq x_2$. Hence, by Property the convexity of $C(\cdot)$,

$$\begin{aligned}
\alpha x_1 \geq \alpha x_2 &> C(x_1 + d_1^{out}) - C(x_1 + d_1^{out} - 1) \\
&\geq C(x_2 + d_1^{out}) - C(x_2 + d_1^{out} - 1) \\
&= C(x_2 + d_2^{out} + 1) - C(x_2 + d_2^{out}).
\end{aligned}$$

Therefore, player 2 strictly prefers to linking to player 1: $u_2(g_{21} = 1) > u_2(g_{21} = 0)$, a contradiction with $g_{21} = 0$. Therefore, $x_1 \geq x_2$ implies both $d_1^{out} \leq d_2^{out}$ and $N_1^{out} \backslash \{2\} \subset N_2^{out} \backslash \{1\}$, establishing the lemma.

Proof of Lemma 2

Claim 1. In every strict equilibrium, if $d_i^{in} \geq d_j^{in}$, then $x_i \geq x_j$.

Suppose not, and that without loss of generality, $d_1^{in} \geq d_2^{in}$ and $x_1 < x_2$. Then, $g_{i1} \leq g_{i2}$ for each $i \neq 1,2$; otherwise, some player $i \neq 1,2$ would have incentive to delete the link to player 1 and deviate to linking with 2. Together with $d_1^{in} = \sum_{i \neq 1,2} g_{i1} + g_{21} \geq d_2^{in} = \sum_{i \neq 1,2} g_{i2} + g_{12}$, we obtain $g_{21} \geq g_{12}$. Holding all other things fixed, let $u_i(x_i, N_i^{out}, d_i^{in})$ denote the payoffs for i given her indegree and choices on x_i and N_i^{out}. We separate two cases.

Case 1: Suppose $g_{21} = g_{12}$. If $g_{21} = g_{12} = 1$, then let $\bar{N}_1^{out} = N_1^{out} \setminus \{2\} \cup \{1\}$ and $\bar{N}_2^{out} = N_2^{out} \setminus \{1\} \cup \{2\}$. If $g_{21} = g_{12} = 0$, let $\bar{N}_1^{out} = N_1^{out}$ and $\bar{N}_2^{out} = N_2^{out}$. Note $|\bar{N}_1^{out}| = |N_1^{out}| = d_1^{out}$ and $|\bar{N}_2^{out}| = |N_2^{out}| = d_2^{out}$. Then,

$$\begin{aligned} u_2(x_2, N_2^{out}, d_2^{in}) - u_2(x_1, \bar{N}_1^{out}, d_2^{in}) = {} & \alpha \left(\sum_{i \in N_2^{out} \setminus N_1^{out}, i \neq 1,2} x_i \right) - \alpha \left(\sum_{i \in N_1^{out} \setminus N_2^{out}, i \neq 1,2} x_i \right) \\ & + \beta (x_2 - x_1) d_2^{in} - C(x_2 + d_2^{out}) + C(x_1 + d_1^{out}). \end{aligned}$$

Since player 2 plays a unique best-response, $u_2(x_2, N_2^{out}, d_2^{in}) > u_2(x_1, \bar{N}_1^{out}, d_2^{in})$. Hence, given $(x_2 - x_1) d_1^{in} \geq (x_2 - x_1) d_2^{in}$,

$$\begin{aligned} u_1(x_2, \bar{N}_2^{out}, d_1^{in}) - u_1(x_1, N_1^{out}, d_1^{in}) = {} & \alpha \left(\sum_{i \in N_2^{out} \setminus N_1^{out}, i \neq 1,2} x_i \right) - \alpha \left(\sum_{i \in N_1^{out} \setminus N_2^{out}, i \neq 1,2} x_i \right) \\ & + \beta (x_2 - x_1) d_1^{in} - C(x_2 + d_2^{out}) + C(x_1 + d_1^{out}) \\ \geq {} & u_2(x_2, N_2^{out}, d_2^{in}) - u_2(x_1, \bar{N}_1^{out}, d_2^{in}) \\ > {} & 0. \end{aligned}$$

However, this implies that player 1 has incentives to deviate to the strategy $(x_2, \bar{N}_2^{out})$.

Case 2: Suppose $g_{21} = 1$ and $g_{12} = 0$. Let $\bar{N}_2^{out} = N_2^{out} \setminus \{1\} \cup \{2\}$. By

$$\begin{aligned} u_2(x_2, N_2^{out}, d_2^{in}) - u_2(x_1, N_1^{out}, d_2^{in}) = {} & \alpha \left(\sum_{i \in N_2^{out} \setminus N_1^{out}, i \neq 1,2} x_i \right) - \alpha \left(\sum_{i \in N_1^{out} \setminus N_2^{out}, i \neq 1,2} x_i \right) \\ & + \alpha x_1 + \beta (x_2 - x_1) d_2^{in} - C(x_2 + d_2^{out}) + C(x_1 + d_1^{out}) \end{aligned}$$

and

$$u_1(x_2,\bar{N}_2^{out},d_1^{in})-u_1(x_1,N_1^{out},d_1^{in})=\alpha\left(\sum_{i\in N_2^{out}\setminus N_1^{out},i\neq 1,2}x_i\right)-\alpha\left(\sum_{i\in N_1^{out}\setminus N_2^{out},i\neq 1,2}x_i\right)$$
$$+\alpha x_2+\beta(x_2-x_1)d_1^{in}-C(x_2+d_2^{out})+C(x_1+d_1^{out}),$$

we obtain

$$\begin{aligned}&\left[u_1(x_2,\bar{N}_2^{out},d_1^{in})-u_2(x_1,N_1^{out},d_1^{in})\right]-\left[u_2(x_2,N_2^{out},d_2^{in})-u_2(x_1,N_1^{out},d_2^{in})\right]\\&=\alpha(x_2-x_1)+\beta(d_1^{in}-d_2^{in})(x_2-x_1)\\&>0.\end{aligned}$$

Hence, by $u_2(x_2,N_2^{out},d_2^{in})>u_2(x_1,N_1^{out},d_2^{in})$, we obtain $u_1(x_2,\bar{N}_2^{out},d_1^{in})>u_1(x_1,N_1^{out},d_1^{in})$. But then, player 1 has incentives to deviate to the strategy $(x_2,\bar{N}_2^{out})$, contradicting with our equilibrium supposition. Therefore, $d_1^{in}\geq d_2^{in}$ implies $x_1\geq x_2$.

Claim 2. In every strict equilibrium, if $x_i\geq x_j$, then $d_i^{in}\geq d_j^{in}$.

Suppose $x_i\geq x_j$ but $d_i^{in}<d_j^{in}$. By Claim 1, $d_j^{in}>d_i^{in}\geq 0$ implies $x_j\geq x_i$. Thus, $x_i=x_j>0$. Hence, by Lemma 1, $d_i^{out}=d_j^{out}$. A contradiction then follows: $x_i=D(\beta d_i^{in})-d_i^{out}<D(\beta d_j^{in})-d_j^{out}=x_j$.

Proof of Proposition 1

The condition for existence of a non-empty network strict equilibrium is established by the following lemma.

Lemma 3. *A nonempty equilibrium network exists if and only if $\beta(n-1)\geq C'\left(\frac{C(1)}{\alpha}\right)$. A nonempty strict equilibrium network exists if and only if the inequality is strict.*

Proof of Lemma 3. We first prove the sufficiency part. Consider the periphery-sponsored star where player 1 is the sole provider with $x_1=D(\beta(n-1))$. Since $\beta(n-1)\geq C'\left(\frac{C(1)}{\alpha}\right)>C'(0)$, $C''(\cdot)>0$, and $C''(\cdot)>0$, there is a unique $x_1>0$ such that $\beta(n-1)=C'(x_1)$ which yields $x_1=D(\beta(n-1))$. Given that other players provide zero content, player 1 will form no link and then $x_1=D(\beta(n-1))$ is the unique optimal content provision. A follower has no incentive to deviate if and only if $\alpha x_1-C(1)\geq 0$, i.e., $\beta(n-1)\geq C'\left(\frac{C(1)}{\alpha}\right)$. Thus, this periphery-sponsored star is an equilibrium network.

Now suppose that a nonempty equilibrium network exists. Then there exist players i and j such that $g_{ij}=1$, which implies

$$\alpha x_j\geq C\left(x_i+d_i^{out}\right)-C\left(x_i+d_i^{out}-1\right)\geq C(1)>0.$$

Then we have $x_j = D(\beta d_j^{in}) - d_j^{out} \le D(\beta(n-1))$ and thus $\alpha D(\beta(n-1)) \ge C(1)$, implying $\beta(n-1) \ge C'\left(\frac{C(1)}{\alpha}\right)$. The proof is similar for an equilibrium to be strict. $\square$

Substituting the functional form $C(y) = \frac{1}{2}y^2$ into the condition $\beta(n-1) > C'\left(\frac{C(1)}{\alpha}\right)$ yields $2\alpha\beta > \frac{1}{n-1}$ as specified in Proposition 1.

Now consider a strict equilibrium (x,g). Suppose $x_1 \le x_2 \le \cdots \le x_n$. *i*) Suppose that there are players i and j such that $x_i < x_j$ and $g_{ji} = 1$. Pick the largest j with $x_i < x_j$ and $g_{ji} = 1$ for some $i < j$. Then, by Lemma 1, all players ℓ with $x_\ell \le x_j$ also i. In particular, all ℓ with $x_\ell = x_j > x_i$ follow i. Hence, our choice of the largest such j implies that, for each $\ell > j$, we must have $x_\ell > x_j$, while $g_{\ell i} = 1$ for each $\ell \le j$ where $\ell \ne i$. Hence, $d_i^{in} = j-1$. On the other hand, since all $\ell > j$ have $x_\ell > x_j$, no player $\ell > j$ follows j. Thus, $d_j^{in} \le j-1$. Hence, $d_j^{in} \le d_i^{in}$. However, by Lemma 2, this implies $x_i \ge x_j$, a contradiction. Hence, if $x_i < x_j$, then $g_{ji} = 0$.

Next, suppose that there are players i and j with $x_i < x_j$ and $g_{ij} = 0$. Pick the smallest j with this property and $g_{ij} = 0$ for player $i < j$. Then by Lemma 1, all those $\ell \ge i$ do not follow j. Thus, j's indegree is at most $i-1$. On the other hand, given that j is smallest number with $g_{i'j} = 0$ for some $i' < j$ and that $i < j$, all players $\ell < i$ follow i. Thus, i's indegree is at least $i-1$. Hence, $d_j^{in} \le i-1 \le d_i^{in}$. But then, by Lemma 2, $x_j \le x_i$, which is a contradiction. Hence, if $x_i < x_j$, then $g_{ij} = 1$.

ii) Suppose that $x_i = x_j$ and $g_{ij} = 1$, and consider players $\ell \ge k \ge i$, with $x_\ell = x_k \ge x_j = x_i$. Let $D(\cdot)$ be the inverse function of $C'(\cdot)$. That $g_{ij} = 1$ implies $x_\ell = x_k \ge x_i > 0$. Then, by the first-order condition $\frac{\partial u_i(x_i, x_{-i}, g)}{\partial x_i} = 0$, we obtain $x_i = D(\beta d_i^{in}) - d_i^{out}$. Substituting this back into $u_i(x_i, x_{-i}, g)$, we obtain

$$u_i(x_i, x_{-i}, g) = \alpha \sum_j g_{ij} x_j + \beta d_i^{in} D(\beta d_i^{in}) - \beta d_i^{in} d_i^{out} - C(D(\beta d_i^{in})).$$

Hence, if i deviates to $g_{ij} = 0$, the utility change is $-\alpha x_j + \beta d_i^{in}$. Given $g_{ij} = 1$ in the strict equilibrium and $x_j = x_i$, we obtain $\alpha x_i = \alpha D(\beta d_i^{in}) - \alpha d_i^{out} > \beta d_i^{in}$, leading to $f(d_i^{in}) \equiv D(\beta d_i^{in}) - \frac{\beta}{\alpha} d_i^{in} > d_i^{out}$. Note that $f'(d) = \beta[D'(\beta d) - \frac{1}{\alpha}] \ge 0$ if and only if $D'(\beta d) \ge \frac{1}{\alpha}$. By $C'(D(\beta d)) = \beta d$, we obtain $D'(\beta d) = \frac{1}{C''(D(\beta d))}$. Hence, $D'(\beta d) \ge \frac{1}{\alpha}$ if and only if $C''(D(\beta d)) \le \alpha$. Thus, the assumption $C'' \le \alpha$ implies $f'(d) \ge 0 \forall d > 0$. By Lemmas 1 and 2, $x_\ell = x_k \ge x_i$ implies $d_\ell^{in} = d_k^{in} \ge d_i^{in}$ and $d_\ell^{out} = d_k^{out} \le d_i^{out}$. Hence, $f(d_k^{in}) \ge f(d_i^{in}) > d_i^{out} \ge d_k^{out}$. Therefore, $D(\beta d_k^{in}) - \frac{\beta}{\alpha} d_k^{in} > d_k^{out}$, implying $\alpha x_\ell > \beta d_k^{in}$. Thus, holding other links fixed, player k's utility from linking to ℓ minus the utility from deleting the link is strictly positive: $\alpha x_\ell - \beta d_k^{in} > 0$. Therefore, k must link to ℓ in the strict equilibrium, since otherwise k would have incentive to deviate. This completes the proof.

Proof of Proposition 2

Consider a strict equilibrium $s = (x,g)$, we will show that $\rho(s)$ is lower than the cutoff as is specified in the proposition. First, by Proposition 1, g is a nested upward-linking network with $\bar{t}$ tiers in g and n_t

players in each tier t of g such that all players in the same tier provide the same content provision level in a strict equilibrium. We let $x_{[t]}$ denote the content provision of tier t players. Second, we will apply the following lemma to show that $g_{ij} = 0$ if $x_j \leq x_i$.

Lemma 4. *If $\alpha \leq 1$, then $g_{ij} = 0$ for each i, j such that $x_i \geq x_j$.*

Proof of Lemma 4. By Proposition 1, if $x_i > x_j$, $g_{ij} = 0$ and $g_{ji} = 1$. Now we will prove by contradiction that $g_{ij} = 0$ if $x_i = x_j$. Suppose that in contrary, there exists a tier t with $n_t \geq 2$ and $g_{ij} = 1$ for any player i and j that belong to tier t. Let $\overline{n}_t = \sum_{t'=1}^{t} n_{t'}$. Then $x_{[t]} = \beta(\overline{n}_t - 1) - n + \overline{n}_{t-1} + 1$, and for some player i that belongs to tier t,

$$\begin{aligned}
u_i(x,g) &= \alpha \sum_{x_j \geq x_i} x_j + \beta(\overline{n}_t - 1)x_i - \frac{1}{2}(\beta\overline{n}_t - \beta)^2 \\
&= \alpha \sum_{x_j > x_i} x_j + \alpha(n_t - 1)x_{[t]} + \beta(\overline{n}_t - 1)x_{[t]} - \frac{1}{2}(\beta\overline{n}_t - \beta)^2 \\
&\leq \alpha \sum_{x_j > x_i} x_j + (n_t - 1)x_{[t]} + \beta(\overline{n}_t - 1)x_{[t]} - \frac{1}{2}(\beta\overline{n}_t - \beta)^2 \\
&= \alpha \sum_{x_j > x_i} x_j + (n_t - 1)[\beta(\overline{n}_t - 1) - n + \overline{n}_{t-1} + 1] + \beta(\overline{n}_t - 1)x_{[t]} \\
&< \alpha \sum_{x_j > x_i} x_j + (n_t - 1)\beta(\overline{n}_t - 1) + \beta(\overline{n}_t - 1)x_{[t]} - \frac{1}{2}(\beta\overline{n}_t - \beta)^2 \\
&= \alpha \sum_{x_j > x_i} x_j + \beta(\overline{n}_t - 1)(x_{[t]} + n_t - 1) - \frac{1}{2}(\beta\overline{n}_t - \beta)^2
\end{aligned}$$

which is her utility when she deviates to $x_i' = \beta(\overline{n}_t - 1) - n + \overline{n}_t$, and deleting her links to players in the same tier t, where the third inequality holds since $\alpha \leq 1$ and the fifth holds since $n - \overline{n}_{t-1} \geq n_t - 1 \geq 1$. That is, player i has a weak incentive to deviate to deleting links to players in the same tier and higher content provision, which is contradiction to (x,g) being a strict equilibrium. Thus, $g_{ij} = 0$ if $x_j = x_i$. $\square$

Thus $x_{[1]} = 0$, $x_{[2]} = \beta n_1 - (n - n_1 - n_2)$, and for a player i in tier 1, $d_i^{out} = n - n_1$. To support s as a strict equilibrium, player i in tier 1 has no incentive to deviate by deleting a link, which requires that

$$\alpha[\beta n_1 - (n - n_1 - n_2)] > \frac{(d_i^{out})^2 - (d_i^{out} - 1)^2}{2} = n - n_1 - \frac{1}{2}$$

Since $n - n_1 - n_2 \geq 0$, we have $\alpha\beta n_1 > n - n_1 - \frac{1}{2}$, that is, $(\alpha\beta + 1)n_1 > n - \frac{1}{2}$. That is $1 - \rho(s) = \frac{n_1}{n} > \frac{1}{\alpha\beta+1} - \frac{1}{2(\alpha\beta+1)n} > \frac{1}{\alpha\beta+1} - \frac{1}{(\alpha\beta+1)n}$. This completes the proof.

Proof of Proposition 3

For sufficiently large n, a nonempty network equilibrium exists. When a nonempty network equilibrium exists, it is obvious that an empty network equilibrium is not a payoff-dominant equilibrium. Thus, it suffices to consider nonempty network equilibria. Consider a strict equilibrium $s=(x,g)$ with $\bar{t}$ tiers and n_t players in each tier t of g. According to Proposition 1, all players in the same tier provide the same level of content in a strict equilibrium. We let $x_{[t]}$ denote the content provision of tier t players. It suffices to show that there exists n^* and ρ^* such that, for each $n>n^*$, if $x_{[1]}=0$ and $\frac{n_1}{n}>1-\rho^*$, then (x,g) is not a payoff-dominant equilibrium. Let $\rho^L=\frac{\alpha^2\beta^2+\alpha\beta-\beta}{\alpha^2\beta^2+\alpha^2\beta+2\alpha\beta+\alpha}$. We will show that there exists n^* such that, for any $n>n^*$, if $\frac{n_1}{n}>1-\rho^L$ and $n_1\geq 3$, then (x,g) is not a payoff-dominant equilibrium. Note that since $\alpha^2\beta+\alpha>1$,

$$1-\rho^L=\frac{\alpha^2\beta+\alpha\beta+\alpha+\beta}{\alpha^2\beta^2+\alpha^2\beta+2\alpha\beta+\alpha}<1.$$

By Lemma 4, $g_{ij}=1$ if and only if $x_j>x_i$. Then, $x_{[1]}=0$, and $x_{[2]}=\beta n_1-(n-n_1-n_2)$. Now, we construct a strategy profile $s'=(x',g')$ which will be shown to be a Pareto improvement and a strict equilibrium. First, pick a positive integer $k<n_1$ that satisfies:

$$\frac{(\alpha\beta+\alpha+1)n_1-(\alpha+1)n}{\alpha\beta+1}>k>\frac{\alpha}{\beta}(n-n_2)-(\alpha+\frac{\alpha}{\beta}-1)n_1.$$

Since $\frac{n_1}{n}>1-\rho^L$,

$$\frac{(\alpha\beta+\alpha+1)n_1-(\alpha+1)n}{\alpha\beta+1}>\frac{\alpha}{\beta}n-(\alpha+\frac{\alpha}{\beta}-1)n_1>\frac{\alpha}{\beta}(n-n_2)-(\alpha+\frac{\alpha}{\beta}-1)n_1.$$

Moreover, $\frac{(\alpha\beta+\alpha+1)n_1-(\alpha+1)n}{\alpha\beta+1}<\frac{(\alpha\beta+1)n_1}{\alpha\beta+1}<n_1$. And since $\frac{n_1}{n}>1-\rho^L$ and $\alpha\leq 1$, we obtain

$$\begin{aligned}\frac{(\alpha\beta+\alpha+1)n_1-(\alpha+1)n}{\alpha\beta+1}&=n\frac{(\alpha\beta+\alpha+1)\frac{n_1}{n}-(\alpha+1)}{\alpha\beta+1}\\&>n\frac{(\alpha\beta+\alpha+1)(1-\rho^L)-(\alpha+1)}{\alpha\beta+1}\\&=n\frac{\beta}{\alpha^2\beta^2+\alpha^2\beta+2\alpha\beta+\alpha}\\&>0.\end{aligned}$$

Thus, there exists n^* such that for each $n>n^*$, such integer $k\geq 1$ exists. Second, g' has $\bar{t}'=\bar{t}+1$ tiers, such that (1) for all players in each of the tiers $t\geq 2$ in g, they are placed in tier $t+1$ in g' and provide exactly the same content as before, i.e., $x'_{[t+1]}=x_{[t]}$ for each $t\geq 2$; (2) for players in tier 1 in g, they are divided into two subsets, among which k players are placed in tier 2 in g', while the remaining players stay in tier 1 and provide $x_i = 0$. Let $N_2' \subset N$ denote the set of those k players in tier 2 in g', and we let $x_i' = \beta(n_1 - k) - (n - n_1)$ for each $i \in N_2'$. Let $N_1' \subset N$ denote the set of those $n_1 - k$ players in tier 1 in g'. Finally, if $x_j > x_i$ then set $g_{ij}' = 1$; otherwise, set $g_{ij}' = 0$. There are two steps.

In Step 1, we show that (x', g') is an equilibrium. First, as is shown in Lemma 4, no player i has incentive to deviate by adding links to a player j with$x_i \geq x_j$. Second, for each player i that belongs to tier $t \geq 2$ in g, since (x, g) is a strict equilibrium, and for any j, $x_i' = x_i$, $g_{ij}' = g_{ij}$, $g_{ji}' = g_{ji}$, player i has no incentive to deviate. Third, we show that player $i \in N_2'$ has no incentive to deviate. Suppose that she deletes l outward links. Then her optimal content provision would be $x_i'' = \beta(n_1 - k) - (n - n_1 - l)$, and her payoff would be weakly lower than

$$\begin{aligned} u_i(x', g') - \alpha l x_{[2]} + \beta(n_1 - k)l &= u_i(x', g') - l\{\alpha[\beta n_1 - (n - n_1 - n_2)] - \beta(n_1 - k)\} \\ &< u_i(x', g'), \end{aligned}$$

where the last inequality holds since $k > \frac{\alpha}{\beta}(n - n_2) - (\alpha + \frac{\alpha}{\beta} - 1)n_1$. Forth, we show that player $i \in N_1'$ has no incentive to deviate. It suffices to show that she does not want to deviate by deleting her links to l players in N_2', that is,

$$\begin{aligned} & -\alpha l[\beta(n_1 - k) - (n - n_1)] - \tfrac{1}{2}(n - n_1 + k - l)^2 + \tfrac{1}{2}(n - n_1 + k)^2 < 0 \\ \Leftrightarrow \quad & \alpha[\beta(n_1 - k) - (n - n_1)] > n - n_1 + k - \tfrac{l}{2} \\ \Leftrightarrow \quad & (\alpha\beta + \alpha + 1)n_1 - (\alpha + 1)n + \tfrac{l}{2} > (\alpha\beta + 1)k \end{aligned}$$

where the last inequality holds since $\frac{(\alpha\beta+\alpha+1)n_1 - (\alpha+1)n}{\alpha\beta+1} > k$.

In Step 2, we show that $u_i(x', g') \geq u_i(x, g)$ for each $i \in N$, and that the inequality holds strictly for each $i \in N_1'$. First, for any player i in tier $t > 2$ in g', $u_i(x', g') = u_i(x, g)$. Second, for each $i \in N_2'$, $u_i(x, g)$ equals to her payoff in (x', g') with deviation to zero content provision, which is lower than $u_i(x', g')$ by the definition of equilibrium. Third, for each $i \in N_1'$, $u_i(x, g)$ equals her payoff in (x', g') when deviating to deleting her links to all players in N_2',which is strictly lower than $u_i(x', g')$ as is shown in Step 1.

To conclude, for each $n > n^*$ and payoff-dominant equilibrium (x, g), we have $\frac{n_1}{n} \leq 1 - \rho^L$, equivalent to $\rho(s) \geq \rho^L$.

Proof of Footnote 15

By Proposition 3, if $\alpha^2\beta + \alpha > 1$, and $\alpha \leq 1$, then there is $\rho^* \in (0, 1)$ such that, for sufficiently large n, we have $\rho(s) \geq \rho^*$ for each payoff dominant strict equilibrium $s = (x, g)$. Suppose that there are $\bar{t}$ tiers in g and n_t players in each tier t of g. According to Proposition 1, all players in the same tier provide the same level of content in a strict equilibrium. We let $x_{[t]}$ denote the content provision of tier t players. It suffices to show that $\min\{x_{[t]} | x_{[t]} > 0\} > \frac{\rho^*}{2\alpha}n$. By Lemma 4, $x_{[1]} = 0$. Then $\min\{x_{[t]} | x_{[t]} > 0\} = x_{[2]}$.

To support (x,g) as a strict equilibrium, some player i in tier 1 has no incentive to delete a link to a player in tier 2, which requires that

$$\begin{aligned} & u_i(x,g) > u_i(x,g) - \alpha x_{[2]} + \tfrac{1}{2}(n-n_1)^2 - \tfrac{1}{2}(n-n_1-1)^2 \\ \Leftrightarrow \quad & \alpha x_{[2]} > \tfrac{1}{2}(2n-2n_1-1) \\ \Leftrightarrow \quad & x_{[2]} > \tfrac{1}{2\alpha}(2n-2n_1-1) \end{aligned}$$

Then for $n \leq \frac{1}{\rho^*}$, since $n-n_1 \geq 1$, $\frac{x_{[2]}}{n} > \frac{1}{2\alpha n}(2n-2n_1-1) \geq \frac{1}{2\alpha n} \geq \frac{\rho^*}{2\alpha} > 0$. For $n > \frac{1}{\rho^*}$, since $n-n_1 \geq \rho^* n$, $\frac{x_{[2]}}{n} > \frac{1}{2\alpha}(2\rho^* - \frac{1}{n}) > \frac{1}{2\alpha}(2\rho^* - \rho^*) = \frac{\rho^*}{2\alpha} > 0$.

Proof of Proposition 4

If $\alpha < 1$, then in any strict equilibrium network with $\bar{t} \geq 2$, we have $\alpha x_{[t]} - \beta \overline{n}_{t-1} < 0$ for all $t \geq 2$, where $x_{[t]} = \beta \overline{n}_{t-1} - (n - \overline{n}_t)$ is the optimal content provision of each tier-t player. This implies that if there are at least two players in tier $t \geq 2$, then they have no incentive to link each other in equilibrium. Hence, the payoff of each player in tier $t \geq 2$ is

$$u_{[t]} = \alpha \left(\sum_{k=t+1}^{\bar{t}} n_k x_{[k]} \right) + \beta \overline{n}_{t-1} x_{[t]} - \frac{1}{2} \left[x_{[t]} + (n - \overline{n}_t) \right]^2.$$

We now consider the following two cases separately: (i) $t \geq 3$ and (ii) $t = 2$. In the first case, given $t \geq 3$, we have $t-1 \geq 2$, and thus $x_{[t-1]} = \beta \overline{n}_{t-2} - (n - \overline{n}_{t-1}) > 0$. Hence,

$$\begin{aligned} u_{[t]} - u_{[t-1]} &= -\alpha n_t x_{[t]} + \frac{\beta^2}{2} (\overline{n}_{t-1} + \overline{n}_{t-2}) n_{t-1} - \beta \left[n n_{t-1} - \overline{n}_{t-1} (n_t + n_{t-1}) \right] \\ &> -\beta \overline{n}_{t-1} n_t + \frac{\beta^2}{2} (\overline{n}_{t-1} + \overline{n}_{t-2}) n_{t-1} - \beta \left[n n_{t-1} - \overline{n}_{t-1} (n_t + n_{t-1}) \right] \\ &= \frac{\beta^2}{2} (\overline{n}_{t-1} + \overline{n}_{t-2}) n_{t-1} - \beta (n - \overline{n}_{t-1}) n_{t-1} \\ &> \frac{\beta^2}{2} (\overline{n}_{t-1} + \overline{n}_{t-2}) n_{t-1} - \beta^2 \overline{n}_{t-2} n_{t-1} \\ &= \frac{\beta^2}{2} (n_{t-1})^2 \\ &> 0. \end{aligned}$$

In the second case, $t=2$. Then, $t-1=1$, $x_{[1]}=0$, and $u_{[1]}=\alpha\left(\sum_{k=2}^{\bar{t}} n_k x_{[k]}\right)-\frac{1}{2}\left(n-n_1\right)^2$. Hence,

$$
\begin{aligned}
u_{[2]}-u_{[1]} &= -\alpha n_2 x_{[2]}+\frac{1}{2}\left(\beta n_1\right)^2-\beta n_1\left(n-n_1-n_2\right)+\frac{1}{2}\left(n-n_1\right)^2 \\
&> -\beta n_1 n_2+\frac{1}{2}\left(\beta n_1\right)^2-\beta n_1\left(n-n_1-n_2\right)+\frac{1}{2}\left(n-n_1\right)^2 \\
&= \frac{1}{2}\left(\beta n_1\right)^2-\beta n_1\left(n-n_1\right)+\frac{1}{2}\left(n-n_1\right)^2 \\
&= \frac{1}{2}\left[\left(\beta n_1\right)-\left(n-n_1\right)\right]^2 \\
&\geq 0.
\end{aligned}
$$

In either case, $u_{[t]}>u_{[t-1]}$, completing the proof.

Proof of Proposition 5

Suppose $\alpha \geq 1$. Then for strict equilibrium networks with $\bar{t} \geq 2$, there are four cases for adjacent tiers t and $t-1$ to consider, where $t \geq 2$: (i) $t \geq 3$ where players link each other within tier $t-1$, (ii) $t \geq 3$ where players do not link each other within tier $t-1$ but players link each other within tier t, (iii) $t=2$ where players link each other within tier 2, and (iv) $t \geq 2$ where $n_t \geq 2$ and players do not link each other within tier t.

First, if players link each other within tier $t \geq 2$, the payoff of each player is

$$
u_{[t]}=\alpha\left(\sum_{k=t}^{\bar{t}} n_k x_{[k]}-x_{[t]}\right)+\beta\left(\bar{n}_t-1\right) x_{[t]}-\frac{1}{2}\left[x_{[t]}+\left(n-\bar{n}_{t-1}-1\right)\right]^2,
$$

where $x_{[t]}=\beta\left(\bar{n}_t-1\right)-\left(n-\bar{n}_{t-1}-1\right)$ is the optimal content provision of each tier-t player. If players do not link each other within tier $t \geq 2$, the payoff of each player is

$$
u_{[t]}=\alpha\left(\sum_{k=t+1}^{\bar{t}} n_k x_{[k]}\right)+\beta \bar{n}_{t-1} x_{[t]}-\frac{1}{2}\left[x_{[t]}+\left(n-\bar{n}_t\right)\right]^2,
$$

where $x_{[t]}=\beta \bar{n}_{t-1}-\left(n-\bar{n}_t\right)$ is the optimal content provision of each tier-t player.

For case (i), we have $x_{[t-1]}=\beta\left(\bar{n}_{t-1}-1\right)-\left(n-\bar{n}_{t-2}-1\right)>0$. Then

$$
\begin{aligned}
u_{[t]}-u_{[t-1]}= & -\alpha\left[\left(n_{t-1}-1\right) x_{[t-1]}+x_{[t]}\right]+\frac{\beta^2}{2}\left(\bar{n}_t+\bar{n}_{t-1}-2\right) n_t \\
& -\beta\left[\left(n-\bar{n}_{t-1}-1\right) n_t-\left(\bar{n}_{t-1}-1\right) n_{t-1}\right],
\end{aligned}
$$

which is strictly decreasing in α. Since $u_{[t]}-u_{[t-1]} \rightarrow-\infty$ as $\alpha \rightarrow+\infty$ and $u_{[t]}-u_{[t-1]} \rightarrow+\infty$ as $\alpha \rightarrow-\infty$, there exists a real number $\hat{\alpha}\left(\beta, n, n_t, n_{t-1}, \bar{n}_{t-1}\right) \in \mathbb{R}$ such that $u_{[t]}-u_{[t-1]}=0$. Hence,

$u_{[t]} \leq u_{[t-1]}$ if $\alpha \geq \max\{1, \hat{\alpha}(\beta, n, n_t, n_{t-1}, \bar{n}_{t-1})\}$ and $u_{[t]} > u_{[t-1]}$ if $\hat{\alpha}(\beta, n, n_t, n_{t-1}, \bar{n}_{t-1}) > 1$ and $1 \leq \alpha < \hat{\alpha}(\beta, n, n_t, n_{t-1}, \bar{n}_{t-1})$.

For case (ii), we have $x_{[t]} = \beta(\bar{n}_t - 1) - (n - \bar{n}_{t-1} - 1) > 0$. Then

$$\begin{aligned} u_{[t]} - u_{[t-1]} = &-\alpha x_{[t]} + \frac{\beta^2}{2}(\bar{n}_t + \bar{n}_{t-2} - 1)(n_t + n_{t-1} - 1) \\ &+ \beta\left[(\bar{n}_t - 1) - (n - \bar{n}_{t-1})(n_t + n_{t-1} - 1)\right], \end{aligned}$$

which is strictly decreasing in α. Since $u_{[t]} - u_{[t-1]} \to -\infty$ as $\alpha \to +\infty$ and $u_{[t]} - u_{[t-1]} \to +\infty$ as $\alpha \to -\infty$, there exists a real number $\hat{\alpha}(\beta, n, n_t, n_{t-1}, \bar{n}_{t-1}) \in \mathbb{R}$ such that $u_{[t]} - u_{[t-1]} = 0$. Hence, $u_{[t]} \leq u_{[t-1]}$ if $\alpha \geq \max\{1, \hat{\alpha}(\beta, n, n_t, n_{t-1}, \bar{n}_{t-1})\}$ and $u_{[t]} > u_{[t-1]}$ if $\hat{\alpha}(\beta, n, n_t, n_{t-1}, \bar{n}_{t-1}) > 1$ and $1 \leq \alpha < \hat{\alpha}(\beta, n, n_t, n_{t-1}, \bar{n}_{t-1})$.

For case (iii), we have $x_{[2]} > 0$, $x_{[1]} = 0$, and $u_{[1]} = \alpha\left(\sum_{k=2}^{\bar{t}} n_k x_{[k]}\right) - \frac{1}{2}(n - n_1)^2$. Then

$$u_{[2]} - u_{[1]} = -\alpha x_{[2]} + \beta(\bar{n}_2 - 1) + \frac{1}{2}\left[\beta(\bar{n}_2 - 1) - (n - n_1)\right]^2,$$

which is strictly decreasing in α. Since $u_{[2]} - u_{[1]} \to -\infty$ as $\alpha \to +\infty$ and $u_{[2]} - u_{[1]} \to +\infty$ as $\alpha \to -\infty$, there exists a real number $\hat{\alpha}(\beta, n, n_1, n_2) \in \mathbb{R}$ such that $u_{[2]} - u_{[1]} = 0$. Hence, $u_{[2]} \leq u_{[1]}$ if $\alpha \geq \max\{1, \hat{\alpha}(\beta, n, n_1, n_2)\}$ and $u_{[2]} > u_{[1]}$ if $\hat{\alpha}(\beta, n, n_1, n_2) > 1$ and $1 \leq \alpha < \hat{\alpha}(\beta, n, n_1, n_2)$.

For case (iv), since $n_t \geq 2$ and players do not link each other within tier t, we have $\alpha x_{[t]} - \beta \bar{n}_{t-1} < 0$, where $x_{[t]} = \beta \bar{n}_{t-1} - (n - \bar{n}_t)$. Then following the proof of Proposition 4, $u_{[t]} > u_{[t-1]}$.

Proof of Example 1

Considering each player's incentives to participate and link/not link another player. Below are the equilibrium conditions for each case.

For Network (a), the equilibrium conditions are

$$\begin{cases} \alpha x - \frac{1}{2} > 0 \\ (1-\delta)\alpha x - \frac{1}{2} < 0 \end{cases} \Rightarrow \begin{cases} \alpha\beta > \frac{1}{6} \\ 0 < 1 - \delta < \frac{1}{6\alpha\beta} \end{cases}$$

where $x = 3\beta$ is the optimal provision of player 1.

For Network (b), the equilibrium conditions are

$$\begin{cases} \alpha x - \frac{1}{2} > 0 \\ (1-\delta)\alpha x - \frac{3}{2} < 0 \\ (1-\delta)\alpha x - 3\beta < 0 \end{cases} \Rightarrow \begin{cases} \alpha\beta > \frac{1}{6} \\ 0 < 1 - \delta < \frac{1}{2\alpha\beta} \\ 0 < 1 - \delta < \frac{1}{\alpha} \end{cases}$$

where $x = 3\beta$ is the optimal provision of each player 1 and player 8.

For Network (c), the equilibrium conditions are

$$\begin{cases} \alpha x_1 - \frac{3}{2} > 0 \\ (1-\delta)\,\alpha x_8 - \frac{3}{2} > 0 \\ (1-\delta)\,\alpha x_1 - \frac{3}{2} < 0 \\ (1-\delta)\,\alpha x_8 - 3\beta < 0 \end{cases} \Rightarrow \begin{cases} \alpha\beta > \frac{1}{2} \\ 1-\delta > \frac{1}{4\alpha\beta} \\ 0 < 1-\delta < \frac{1}{2\alpha\beta} \\ 0 < 1-\delta < \frac{1}{2\alpha} \end{cases}$$

where $x_1 = 3\beta$ and $x_8 = 6\beta$ are the optimal provisions of players 1 and 8, respectively.

Note that there are no equilibrium conditions for all the three networks because Network (a) and Network (c) cannot coexist. Network (a) and Network (b) are strict equilibrium networks when $\alpha\beta > \frac{1}{6}$ and $0 < 1-\delta < \min\{\frac{1}{6\alpha\beta}, \frac{1}{\alpha}\}$, and Network (c) a strict equilibrium network when $\alpha\beta > \frac{1}{2}$, $\beta > \frac{1}{2}$, and $\frac{1}{4\alpha\beta} < 1-\delta < \min\{\frac{1}{2\alpha\beta}, \frac{1}{2\alpha}\}$.

Proof of Proposition 6

We prove using a general cost function $C(\cdot)$, with $C(0) \geq 0$, $C'(\cdot) > 0$, $C''(\cdot) > 0$, and $\lim_{y\to+\infty} \frac{C'(y)}{y} = \gamma \in (0, +\infty)$, where $C'(\cdot)$ is invertible and the inverse function is denoted by $D(\cdot) \equiv C'^{-1}(\cdot)$. Suppose $\beta(n-1) > C'\left(\frac{C(1)}{(1-\delta)\alpha}\right)$. Consider the periphery-sponsored star where player 1 provides $x_1 = D(\beta(n-1))$ of good A and follows no one, and all of the remaining $n-1$ players provide zero content and follow no other players but player 1. Apparently, player 1 has no incentive to deviate. And a follower in N_B has no incentive to deviate if and only if $(1-\delta)\,\alpha x_1 - C(1) > 0$, i.e., $\beta(n-1) > C'\left(\frac{C(1)}{(1-\delta)\alpha}\right)$. This condition also implies that no follower in N_A would deviate either. Hence, this periphery-sponsored star constitutes a strict nested upward-linking equilibrium network without isolated agents. Finally, we can substitute the functional form $C(y) = \frac{1}{2}y^2$ into the inequality $\beta(n-1) > C'\left(\frac{C(1)}{(1-\delta)\alpha}\right)$ to verify that it gives the condition $2\alpha\beta(1-\delta) > \frac{1}{n-1}$ specified in the proposition.

Proof of Proposition 7

Claim 1. For each $x \in \mathbb{R}^n_+$ and $g^* \in G^*(x)$, there exists a transfer profile τ such that $g(\tau) = g^*$, and it is pairwise stable and a Nash equilibrium in the Stage 2 subgame.

Consider $g^* \in G^*(x)$ and that $x_1 \geq \cdots \geq x_n$. Let d_i^{out} be i's outdegree in g^*. Let $z_i = \max\{x_j | j \in N, j \neq i, g^*_{ij} = 0\}$. By construction, $g^*_{ij} = 1$ implies $\alpha x_j - \Delta C(x_i, d_i^{out}) \geq \eta$ and $g^*_{ij} = 0$ implies $\alpha x_j - \Delta C(x_i, d) < \eta$ for each $d > d_i^{out}$. Consider the following $\tau \in \mathbb{T}$: 1) if $g^*_{ij} = g^*_{ji} = 1$, then let $\tau_{ij} = \tau_{ji} = \eta$; 2) if $g^*_{ij} = 1$ and $g^*_{ji} = 0$, then let $\tau_{ij} = \min\{\alpha(x_j - z_i) + \eta, \alpha x_j - \Delta C(x_i, d_i^{out})\}$ and $\tau_{ji} = \eta - \tau_{ij}$; 3) if $g^*_{ij} = g^*_{ji} = 0$, then let $\tau_{ij} = \tau_{ji} = 0$. Then, $g(\tau) = g^*$.

Step 1: We show that τ is a Nash equilibrium in the Stage 2 subgame.

Suppose that i deviates to $\tau'_i \neq \tau_i$ and i's utility change is Δu_i. First, suppose $g(\tau'_i, \tau_{-i}) = g^*$. Then, for each j with $g^*_{ij} = 0$, we have $\tau'_{ij} < \tau_{ij}$ and a change to any $\tau'_{ij} < \varepsilon$ does not change i's utility. For each j with $g^*_{ij} = 1$, we have $\tau'_{ij} \geq \tau_{ij}$, since $\tau_{ij} = \eta - \tau_{ji}$ is the minimum transfer to maintain the link from i to j. If $\tau'_{ij} = \tau_{ij}$ for each j with $g^*_{ij} = 1$, then $u'_i = u_i$. If $\tau'_{ij} > \tau_{ij}$ for some j with $g^*_{ij} = 1$, then $\Delta u_i < 0$, because i pays more transfers. Hence, i has no strict incentive to deviate to any τ'_i with $g(\tau'_i, \tau_{-i}) = g^*$.

Next, suppose $g(\tau'_i, \tau_{-i}) = g' \neq g^*$, and i's outdegree in g' is $d_i^{out}(g')$. First, suppose $d_i^{out}(g') = d_i^{out}(g^*)$. Then, $\tau'_{ij} < \tau_{ij}$ for some j with $g^*_{ij} = 1$ so that $g'_{ij} = 0$, and $\tau'_{ik} \geq \eta > \tau_{ik}$ for the same number of players k with with $g^*_{ik} = 0$ to make $g'_{ik} = 1$. Thus, we can match those j to those k so that the change in utility, Δu_i, sums up the utility changes due to adding a link to a k with $g^*_{ik} = 0$ and simultaneously deleting a link to j with $g^*_{ij} = 1$, while maintaining the same outdegree. For each of such switch of links, the change in i's utility is (weakly) negative:

$$-\alpha x_j + \tau_{ij} + \alpha x_k - \tau'_{ik} \leq \tau_{ij} - [\alpha(x_j - x_k) + \eta] \leq \tau_{ij} - [\alpha(x_j - z_i) + \eta] \leq 0,$$

where the first inequality follows from $\tau'_{ik} \geq \eta$, the second inequality follows from $x_k \leq z_i$ for each k with $g^*_{ik} = 0$, and the last inequality is due to $\tau_{ij} \leq \alpha(x_j - z_i) + \eta$ by construction. Hence, i cannot improve by deviating to any τ'_i with $d_i^{out}(g') = d_i^{out}(g^*)$.

Second, suppose $d_i^{out}(g') > d_i^{out}(g^*)$. We break down the change from τ_i to τ'_i into two steps and show that i cannot improve in either step. In the first step, i increases the transfers to some k with $g^*_{ik} = 0$ to make $g'_{ik} = 1$ and decreases the transfers to the same number of players j with $g^*_{ij} = 1$ to make $g'_{ij} = 0$. This step only involves switch of links and i's outdegree is not changed. In the second step, i increases transfers to a further subset of players k with $g^*_{ik} = 0$ to make $g'_{ik} = 1$. This step only involves adding new links. We have shown that no τ'_i with $d_i^{out}(g') = d_i^{out}(g^*)$ can strictly improve i's utility. Hence, i cannot improve in the first step. Consider the second step. Consider k with $g^*_{ik} = 0$ and let d_i^k be i's outdegree right after adding the link to k. Since $d_i^k \geq d_i^{out}(g^*) + 1$, forming each new link to k results in a negative change in i's utility:

$$\alpha x_k - \Delta C(x_i, d_i^k) - \tau'_{ik} \leq \alpha x_k - \Delta C(x_i, d_i^{out}(g^*) + 1) - \eta < 0.$$

Hence, i cannot improve in the second step.

Third, consider $d_i^{out}(g') < d_i^{out}(g^*)$. Similarly to the last case, we consider the change from τ_i to τ'_i involving two steps. The first step only involves switch of links. We have shown that this step cannot improve i's utility. The second step only involves deleting links to some j with $g^*_{ij} = 1$. Let d_i^j denote i's outdegree right before deleting the link to such j. Then, given $d_i^j \leq d_i^{out}(g^*)$ and $\tau_{ij} \leq \alpha x_j - \Delta C(x_i, d_i^{out}(g^*))$, player i's utility change due to deleting the link to j is

$$-\alpha x_j + \Delta C(x_i, d_i^j) + \tau_{ij} \leq -\alpha x_j + \Delta C(x_i, d_i^{out}(g^*)) + \tau_{ij} \leq 0.$$

Hence, i cannot strictly improve in the second step.

Step 2: We show that τ is pairwise stable.

Suppose not, and there are pairwise deviations, (τ_i', τ_j'), by i and j such that $u_i(\tau_{-ij}, \tau_i', \tau_j', x) > u_i(\tau, x)$ and $u_j(\tau_{-ij}, \tau_i', \tau_j', x) \geq u_j(\tau, x)$. Let $\Delta u_i = u_i(\tau_{-ij}, \tau_i', \tau_j', x) - u_i(\tau, x)$ and $g' = g(\tau_{-ij}, \tau_i', \tau_j')$. If $\tau_{ij}' = \tau_{ij}$ or $\tau_{ji}' = \tau_{ji}$, then the impact of (τ_i', τ_j') on u_i is the same as player i making a unilateral deviation to τ_i', i.e., $u_i(\tau_{-ij}, \tau_i', \tau_j', x) = u_i(\tau_{-i}, \tau_i', x)$. We have shown that i cannot strictly improve by any unilateral deviation. In all of the following cases, suppose $\tau_{ij}' \neq \tau_{ij}$ and $\tau_{ji}' \neq \tau_{ji}$.

First, suppose $g_{ij}' = g_{ji}' = 0$. Then regardless of whether $g_{ij}^* + g_{ji}^* \geq 1$ or $g_{ij}^* + g_{ji}^* = 0$, the pairwise deviations result in the same Δu_i as that of i making a unilateral deviation to some τ_i'' with $\tau_{ik}'' = \tau_{ik}'$ for each $k \neq j$ and with a sufficiently small $\tau_{ij}'' < 0$. Hence, i cannot strictly improve: $\Delta u_i \leq 0$. Second, suppose $g_{ij}' + g_{ji}' \geq 1$ and $g_{ij}^* = g_{ji}^* = 1$. In the case of $\tau_{ij}' < \tau_{ij} = \eta$, i breaks the link to j. In the case of $\tau_{ij}' > \eta$, i increases the transfer to j. In either case, the same utility for i can be achieved by unilaterally deviating to τ_i'. Hence, $\Delta u_i \leq 0$. Third, suppose $g_{ij}' + g_{ji}' \geq 1$, $g_{ij}^* = 0$, and $g_{ji}^* = 1$. If $g_{ij}' = 0$, then $g_{ji}' = 1$, and i can strictly improve only if $\tau_{ij}' < \tau_{ij}$ and $\tau_{ji}' \geq \eta - \tau_{ij}' > \eta - \tau_{ij} = \tau_{ji}$. But then, j is strictly worse off. If $g_{ij}' = 1$, then $\tau_{ij}' \geq \eta$, and $u_i(\tau_{-ij}, \tau_i', \tau_j', x) = u_i(\tau_{-i}, \tau_i', x)$. Hence, $\Delta u_i \leq 0$.

Finally, suppose $g_{ij}' + g_{ji}' \geq 1$, $g_{ij}^* = 1$, and $g_{ji}^* = 0$. If $g_{ij}' = 1$, then i is strictly better off only if $\tau_{ij}' < \tau_{ij}$ and $\tau_{ji}' \geq \eta - \tau_{ij}' > \eta - \tau_{ij} = \tau_{ji}$. Then, j is strictly worse off. Now, suppose $g_{ij}' = 0$. Then $g_{ji}' = 1$ and $\tau_{ji}' \geq \eta$, i.e., j forms a new link to i in g'. But then, $u_j(\tau_{-ij}, \tau_i', \tau_j', x) = u_j(\tau_{-j}, \tau_j', x)$. Hence, $\Delta u_j \leq 0$, and we have $\Delta u_j = 0$ only if j simultaneously deletes a link to a player k with $g_{jk}^* = 1$, so that $d_j^{out}(g') = d_j^{out}(g^*)$ and

$$\begin{aligned}\Delta u_j &= -\alpha x_k + \tau_{jk} + \alpha x_i - \tau_{ji}' + \tau_{ji} \\ &\leq -\alpha(x_k - x_i) + \alpha(x_k - z_j) + \eta - \tau_{ji}' + \tau_{ji} \\ &\leq -\alpha(z_j - x_i) - (\tau_{ji}' - \eta) - (\tau_{ij} - \eta).\end{aligned}$$

Hence, $\Delta u_j = 0$ only if $\tau_{ji}' = \tau_{ij} = \eta$, given $\tau_{ij} \geq \eta$ by $g_{ij}^* = 1$. And given $g_{ji}' = 1$, we have $\tau_{ij}' + \tau_{ji}' \geq \eta$. Thus, $\tau_{ij}' \geq \eta - \tau_{ji}' = 0$. In the most favorable case to i, $\tau_{ij}' = 0$. But then, $u_i(\tau_{-ij}, \tau_i', \tau_j', x) = u_i(\tau_{-i}, \tau_i', x)$. Therefore, $\Delta u_i \leq 0$.

Claim 2. If $n \geq 2\eta/\alpha^2 + 1$ and $\alpha \leq 1$, then there exists $(x^*, \hat{\tau}^*) \subset S^*(\alpha)$ such that i) $g(\hat{\tau}^*(x))$ is a nested upward-linking network, ii) $x_i > x_j = 0$ for some i and j, and iii) as $n \to \infty$, the fraction of players with $x_i > 0$ approaches $1 - \sqrt{\frac{1}{1+\alpha^2}}$.

Most part of our arguments apply to a general $C(\cdot)$ that satisfies $C(0) \geq 0$, $C'(\cdot) > 0$, $C''(\cdot) > 0$. We will indicate explicitly when we apply the functional form $C(y) = \frac{1}{2}y^2$. Consider $N_1, N_2 \subset N$ such that $N_2 = \{1, 2, \ldots, I\}$ and $N_1 = \{I+1, \ldots, n\}$, with $1 \leq I \leq n-1$. The exact number of I is determined in Step 1 below. Let $x^* \in \mathbb{R}_+^n$ be such that $x_i^* = 0 \forall i \in N_1$ and $x_i^* = D(\alpha(n-I)) \forall i \in N_2$. Let $\hat{\tau} = (\hat{\tau}_1, \ldots, \hat{\tau}_n)$ denote the transfer strategy profile considered in the proof of Claim 1, so that $g(\hat{\tau}(x)) \in G^*(x)$ for each $x \in \mathbb{R}_+^n$. Consider $\hat{\tau}^*$ such that for each $x \in \mathbb{R}_+^n$ and $i, j \in N$, $i \neq j$: if $x_j > D(\alpha(n-I))$, $g_{ij}(\hat{\tau}(x)) = 1$

and $g_{ji}(\hat{\tau}(x)) = 0$, then $\hat{\tau}^*_{ij}(x) = \min\{\hat{\tau}_{ij}(x), D(\alpha(n-I)) - \Delta C(0,I)\}$ and $\hat{\tau}^*_{ji}(x) = \eta - \hat{\tau}^*_{ij}(x)$; otherwise, $\hat{\tau}^*_{ij}(x) = \hat{\tau}_{ij}(x)$. Then, $g(\hat{\tau}^*(x)) = g(\hat{\tau}(x)) \in G^*(x) \forall \in \mathbb{R}^n_+$. We show that, with a suitably selected I, $(x^*, \hat{\tau}^*) \subset S^*(\alpha)$. Applying the same arguments we use to prove Claim 1, one can verify that, for each $x \in \mathbb{R}^n_+$, $\hat{\tau}^*(x) \in \mathbb{T}$ is pairwise stable and a Nash equilibrium given x. In the following, we show that given x^*_{-i}, x^*_i is a best-response in Stage 1 for each $i \in N$, given the anticipation that all players make transfers according to $\hat{\tau}^*(x)$ in Stage 2 for each $x \in \mathbb{R}^n_+$.

For each $i \in N$, given x^*_{-i}, let $u_i(x_i, d^{out}_i, d^{in}_i; g(\hat{\tau}(x_i, x^*_{-i})))$ denote i's utility given her provision x_i, her outdegree d^{out}_i, and indegree d^{in}_i in network $g(\hat{\tau}(x_i, x^*_{-i}))$. When the network $g(\hat{\tau}(x_i, x^*_{-i}))$ referred to is clear, we drop the network argument and write $u_i(x_i, d^{out}_i, d^{in}_i)$.

Step 1: Determine I^*.

Let $g^* = g(\hat{\tau}(x^*))$. Consider an $i \in N_2$. Let $m_1 = \min\{m \geq 4 | \alpha D(\alpha(m-I)) - \Delta C(0,I) \geq \eta\}$. Then, $\alpha D(\alpha(n-I)) - \Delta C(0,I) \geq \eta$ for each $n \geq m_1$. Assume $n \geq m_1$. Then $g^*_{ji} = 1$ for each $j \in N_1$ and $i \in N_2$. Moreover, given $C(T) = \frac{1}{2}T^2$ and $\alpha < 1$, we obtain for each $d \in \{1, \ldots, I-1\}$,

$$\begin{aligned}\alpha D(\alpha(n-I)) - \Delta C(D(\alpha(n-I)), d) &= \alpha^2(n-I) - \frac{1}{2}\left[2\alpha(n-I) + 2d - 1\right] \\ &\leq \alpha(\alpha-1)(n-I) - \frac{1}{2} \\ &< \eta.\end{aligned}$$

Hence, $g^*_{ij} = 0$ for each $i, j \in N_2$. Therefore, in g^*, player i follows no player and is followed by all players in N_1. Hence, i's utility given x^* and transfers $\hat{\tau}(x^*)$ is

$$u_i(D(\alpha(n-I)), 0, n-I) = -C(D(\alpha(n-I))) + (n-I)[\alpha D(\alpha(n-I)) - \Delta C(0,I) - \eta].$$

Suppose that i deviates to $x_i = 0$. Then, i loses all followers from N_1, and follows all other players in N_2, because $\alpha D(\alpha(n-I)) - \Delta C(0, I-1) > \alpha D(\alpha(n-I)) - \Delta C(0,I) \geq \eta$. Thus, i's utility from the deviation is

$$\begin{aligned}u_i(0, I-1, 0) &= (I-1)\alpha D(\alpha(n-I)) - C(I-1) - (I-1)[\alpha D(\alpha(n-I)) - \Delta C(0, I-1)] \\ &= \sum_{d=1}^{I-1}\Big[\Delta C(0, I-1) - \Delta C(0,d)\Big].\end{aligned}$$

Let I^* be the largest integer I such that $u_i(D(\alpha(n-I)), 0, n-I) \geq u_i(0, I-1, 0)$. Note that $u_i(0, I-1, 0)$ is increasing in I, and given $n \geq 3$, $u_i(D(\alpha(n-I)), 0, n-I) = 0 < u_i(0, I-1, 0)$ at $I = n$. If $n \geq m_2 = {}^{2\eta}/_{\alpha^2} + 1$, then $u_i(D(\alpha(n-I)), 0, n-I) \geq u_i(0, I-1, 0) = 0$ at $I = 1$. From $u_i(D(\alpha(n-I)), 0, n-I) \geq u_i(0, I-1, 0) \geq 0$, we obtain $\alpha D(\alpha(n-I)) - \Delta C(0,I) \geq \eta$. Thus, $m_2 \geq m_1$. Therefore, $n \geq m_2$ is sufficient to guarantee that a unique I^* exists, $1 \leq I^* \leq n-1$, and $u_i(D(\alpha(n-I)), 0, n-I) \geq 0$. Now, let $I = I^*$. Then by construction, no $i \in N_2$ has incentive to deviate to $x_i = 0$. Given $C(T) = \frac{1}{2}T^2$, we have

$$\lim_{n\to\infty}\frac{I^*}{n} = 1-\sqrt{\frac{1}{1+\alpha^2}}.$$

Step 2: Given $I = I^*$, no $i \in N_2$ has incentive to deviate from $x_i^* = D(\alpha(n-I))$.

Consider a deviation in a neighborhood of $D(\alpha(n-I))$. Let x^L be i's minimum provision for a player $j \in N_1$ to follow i, which satisfies $\alpha x^L - \Delta C(x^L, I) = \eta$ and $0 < x^L < D(\alpha(n-I))$. Then, for each $x_i \in [x^L, D(\alpha(n-I))]$, player i's utility from the deviation is

$$u_i(x_i, 0, n-I) = -C(x_i) + (n-I)[\alpha x_i - \Delta C(0,I) - \eta],$$

which is maximized at $x_i = D(\alpha(n-I))$. If i deviates to $x_i > D(\alpha(n-I))$, then by construction $\hat{\tau}^*_{ij}(x_i, x^*_{-i}) = D(\alpha(n-I)) - \Delta C(0,I)$ and $\hat{\tau}^*_{ji}(x_i, x^*_{-i}) = \eta - \hat{\tau}^*_{ij}(x_i, x^*_{-i})$, resulting in $u_i(x_i, 0, n-I) < u_i(D(\alpha(n-I)), 0, n-I)$. Therefore, i has no incentive to deviate to any $x_i \geq x^L$ and $x_i \neq D(\alpha(n-I))$.

Consider a deviation $x_i < x^L$, leading to network $g' = g(\hat{\tau}(x_i, x^*_{-i}))$. Then in g', player i loses all followers from N_1 and follows $d_i^{out}(x_i)$ other players in N_2, where $d_i^{out}(x_i) = \max\{d | \alpha D(\alpha(n-I)) - \Delta C(x_i, d) \geq \eta\}$. Define $y^\dagger > 0$ by $\alpha D(\alpha(n-I)) - \Delta C(y^\dagger, I-1) = \eta$. Then, for each $x_i > y^\dagger$, we have $d_i^{out}(x_i) \leq I-1$, and $\hat{\tau}_{ij}(x_i, x^*_{-i}) = \eta$ for each $j \in N_2$ with $g_{ij}(\hat{\tau}(x_i, x^*_{-i})) = 1$. If $I^* = 1$, then $d_i^{out}(x_i) = 0$ for each $x_i < x^L$ and $u_i(x_i, d_i^{out}(x_i), 0) = 0$. If $I^* \geq 2$, then i's utility from deviating to $x_i \in (y^\dagger, x^L)$ is

$$\begin{aligned} u_i(x_i, d_i^{out}(x_i), 0) &= d_i^{out}(x_i)\alpha D(\alpha(n-I)) - C(x_i + d_i^{out}(x_i)) - d_i^{out}(x_i)\eta \\ &= \sum_{d=1}^{d_i^{out}(x_i)} \Big[\alpha D(\alpha(n-I)) - \Delta C(x_i, d) - \eta\Big] - C(x_i), \end{aligned}$$

which is decreasing in x_i, because $d_i^{out}(x_i)$ is decreasing in x_i and $\alpha D(\alpha(n-I)) - \Delta C(x_i, d) \geq \eta$ for each $d \in \{1, \ldots, d_i^{out}(x_i)\}$. Thus, $u_i(y^\dagger, I-2, 0) \geq u_i(x_i, d_i^{out}(x_i), 0)$ for each x_i with $y^\dagger < x_i < x^L$.

Next, consider $x_i \leq y^\dagger$. We have $d_i^{out}(x_i) = I-1$ and $\hat{\tau}_{ij}(x_i, x^*_{-i}) = \alpha D(\alpha(n-I)) - \Delta C(x_i, I-1)$ for each $j \in N_2, j \neq i$. Thus, the utility for i when $x_i \leq y^\dagger$ is

$$\begin{aligned} u_i(x_i, I-1, 0) &= (I-1)\alpha D(\alpha(n-I)) - C(x_i + I - 1) - (I-1)[\alpha D(\alpha(n-I)) - \Delta C(x_i, I-1)] \\ &= \sum_{d=1}^{I-1} \Big[\Delta C(x_i, I-1) - \Delta C(x_i, d)\Big] - C(x_i) \\ &= \sum_{d=1}^{I-1} \Big[(x_i + I - \frac{3}{2}) - (x_i + d - \frac{1}{2})\Big] - C(x_i) \\ &= \sum_{d=1}^{I-1} (I-d-1) - C(x_i), \end{aligned}$$

where the third equality uses $C(T) = \frac{1}{2}T^2$. Hence, $u_i(0, I-1, 0) > u_i(x_i, I-1, 0)$ for each $0 < x_i \leq y^\dagger$. In particular, $u_i(0, I-1, 0) > u_i(y^\dagger, I-1, 0)$. And since $u_i(y^\dagger, I-1, 0) > u_i(y^\dagger, I-2, 0)$, we obtain

$u_i(0,I-1,0) > u_i(y^\dagger,I-2,0)$. By our choice of $I = I^*$, we have $u_i(D(\alpha(n-I)),0,n-I) \geq u_i(0,I-1,0)$. Thus, we establish that $i \in N_2$ has no incentive to deviate to any $x_i < D(\alpha(n-I))$.

Step 3: Given $I = I^*$, no $i \in N_1$ has incentive to deviate from $x_i = 0$.

First, by our choice of $I = I^*$, $u_i(D(\alpha(n-I-1)),0,n-I-1) < u_i(0,I,0)$ for each $i \in N_1$. Hence, no $i \in N_1$ has incentive to deviate to $x_i = D(\alpha(n-I-1))$. Next, observe that our arguments in Step 2 to show that for each $i \in N_2$, 1) $u_i(0,I-1,0) > u_i(x_i,I-1,0)$ for each $x_i \in (0,y^\dagger]$ and 2) $u_i(0,I-1,0) > u_i(y^\dagger,I-2,0) \geq u_i(x_i,d_i^{out}(x_i),0)$ for each $x_i \in (y^\dagger,x^L)$ apply to any $1 \leq I \leq n-1$. Hence, replacing $I-1$ with I, and $I-2$ with $I-1$, in the utility functions in the last sentence, we obtain that for each $i \in N_1$, 1) $u_i(0,I,0) > u_i(x_i,I,0)$ for each $x_i \in (0,y^\dagger]$ and 2) $u_i(0,I,0) > u_i(y^\dagger,I-1,0) \geq u_i(x_i,d_i^{out}(x_i'),0)$ for each $x_i \in (y^\dagger,x^L)$, given that $y^\dagger$ is defined by $\alpha D(\alpha(n-I)) - \Delta C(y^\dagger,I) = \eta$ and x^L is defined by $\alpha x^L - \Delta C(x^L,I-1) = \eta$ for each $i \in N_1$. Then, $i \in N_1$ has no incentive to deviate to any $x_i \in (0,x^L)$. And when $x_i \geq x^L$, the best choice for i is $x_i = D(\alpha(n-I-1))$, which has been shown to be worse than $x_i = 0$. Hence, each $i \in N_1$ has no incentive to deviate from $x_i = 0$.

References

Baetz, O. (2015). Social activity and network formation. *Theoretical Economics*, 10(2), 315–340. Number: 2.

Bakshy, E., Hofman, J. M., Mason, W. A., & Watts, D. J. (2011). Everyone's an influencer: quantifying influence on twitter. In *Proceedings of the fourth ACM international conference on Web search and data mining*, WSDM '11 (pp. 65–74). New York, NY, USA: Association for Computing Machinery.

Bala, V. & Goyal, S. (2000). A noncooperative model of network formation. *Econometrica*, 68(5), 1181–1229. Number: 5 Publisher: John Wiley & Sons, Ltd (10.1111).

Ballester, C., Calvo-Armengol, A., Zenou, Y., Calvó-Armengol, A., & Zenou, Y. (2006). Who's who in the network. wanted: the key player. *Econometrica*, 74(5), 1403–1417. Number: 5 Publisher: Wiley Online Library ISBN: 1570-6435.

Becker, J., Brackbill, D., & Centola, D. (2017). Network dynamics of social influence in the wisdom of crowds. *Proceedings of the National Academy of Sciences*, 114(26), E5070–E5076. Publisher: Proceedings of the National Academy of Sciences.

Belhaj, M., Bervoets, S., & Deroïan, F. (2016). Efficient networks in games with local complementarities. *Theoretical Economics*, 11(1), 357–380. Number: 1 Publisher: Wiley Online Library.

Bloch, F. & Jackson, M. O. (2007). The formation of networks with transfers among players. *Journal of Economic Theory*, 133(1), 83–110. Number: 1 Publisher: Academic Press.

Bojkov, N. (2023). How much money do instagram influencers make in 2023? Section: Blog.

Brown, A. & Freeman, A. (2022). The highest-paid YouTube stars: MrBeast, jake paul and markiplier score massive paydays. *Forbes*. Section: Social Media.

Burgess, J. & Green, J. (2018). *YouTube: online video and participatory culture*. John Wiley & Sons.

Bärtl, M. (2018). YouTube channels, uploads and views: a statistical analysis of the past 10 years. *Convergence*, 24(1), 16–32. Number: 1 Publisher: SAGE Publications Ltd.

Cabrales, A., Calvó-Armengol, A., & Zenou, Y. (2011). Social interactions and spillovers. *Games and Economic Behavior*, 72(2), 339–360. Number: 2 Publisher: Elsevier ISBN: 15424766.

Calvó-Armengol, A., Patacchini, E., & Zenou, Y. (2009). Peer effects and social networks in education. *Review of Economic Studies*, 76(4), 1239–1267. Number: 4 ISBN: 1467-937X.

Cha, M., Haddadi, H., Benevenuto, F., & Gummadi, K. (2010). Measuring user influence in twitter: the million follower fallacy. *Proceedings of the International AAAI Conference on Web and Social Media*, 4(1), 10–17. Number: 1.

Chen, Y.-J., Zenou, Y., & Zhou, J. (2018). Multiple activities in networks. *American Economic Journal: Microeconomics*, 10(3), 34–85. Number: 3.

Cong, L. W. & Li, S. (2023). A model of influencer economy. Issue: 31243.

Cui, Z. & Weidenholzer, S. (2021). Lock-in through passive connections. *Journal of Economic Theory*, 192, 105187. Publisher: Elsevier.

Cunningham, S., Craig, D., & Silver, J. (2016). YouTube, multichannel networks and the accelerated evolution of the new screen ecology. *Convergence*, 22(4), 376–391. Number: 4 Publisher: SAGE Publications Ltd.

Ding, S. (2022). Link investment substitutability: a factor influencing network formation. *Games and Economic Behavior*, 136, 340–359.

Duffy, B. E. (2020). Social media influencers. In *The International Encyclopedia of Gender, Media, and Communication* (pp. 1–4). John Wiley & Sons, Ltd. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/9781119429128.iegmc219.

Funk, M. (2020). How many YouTube channels are there?

Galeotti, A. & Goyal, S. (2010). The law of the few. *American Economic Review*, 100(4), 1468–1492. Number: 4 Publisher: American Economic Association ISBN: 1573-4471.

Golub, B. & Jackson, M. O. (2010a). Naïve learning in social networks and the wisdom of crowds. *American Economic Journal: Microeconomics*, 2(1), 112–149. Number: 1 ISBN: 0521698227.

Golub, B. & Jackson, M. O. (2010b). Using selection bias to explain the observed structure of internet diffusions. *Proceedings of the National Academy of Sciences*, 107(24), 10833–10836. Number: 24 Publisher: National Academy of Sciences ISBN: 1091-6490 (Electronic)\r0027-8424 (Linking).

Goyal, S. & Vega-Redondo, F. (2005). Network formation and social coordination. *Games and Economic Behavior*, 50(4), 178–207. Number: 4 Publisher: Elsevier.

Herskovic, B. & Ramos, J. (2020). Acquiring information through peers. *American Economic Review*, 110(7), 2128–2152. Number: 7.

Hiller, T. (2017). Peer effects in endogenous networks. *Games and Economic Behavior*, 105, 349–367.

Iyer, G. & Katona, Z. (2016). Competing for attention in social communication markets. *Management Science*, 62(8), 2304–2320. Number: 8 Publisher: INFORMS.

Jackson, M. O., Rogers, B., & Zenou, Y. (2016). The economic consequences of social network structure. *Social Science Electronic Publishing, Inc.*, (February), 1–55. Number: February.

Jackson, M. O. & Watts, A. (2002). On the formation of interaction networks in social coordination games. *Games and Economic Behavior*, 41(2), 265–291. Number: 2 Publisher: Academic Press.

Kandori, M., Mailath, G. J., & Rob, R. (1993). Learning, mutation, and long run equilibria in games. *Econometrica*, 61(1), 29–56. Number: 1 ISBN: 00129682.

Kim, S., Han, J., Yoo, S., & Gerla, M. (2017). How are social influencers connected in instagram? In G. L. Ciampaglia, A. Mashhadi, & T. Yasseri (Eds.), *Social Informatics*, Lecture Notes in Computer Science (pp. 257–264). Cham: Springer International Publishing.

Kinateder, M. & Merlino, L. P. (2017). Public goods in endogenous networks. *American Economic Journal: Microeconomics*, 9(3), 187–212. Number: 3.

Kor, R. & Zhou, J. (2023). Multi-activity influence and intervention. *Games and Economic Behavior*, 137, 91–115.

König, M. D., Tessone, C. J., & Zenou, Y. (2014). Nestedness in networks: a theoretical model and some applications. *Theoretical Economics*, 9(3), 695–752. Number: 3 Publisher: Wiley Online Library.

Lampel, J. & Bhalla, A. (2007). The role of status seeking in online communities: giving the gift of experience. *Journal of Computer-Mediated Communication*, 12(2), 434–455. Number: 2.

Li, X. (2023). Designing weighted and directed networks under complementarities. *Games and Economic Behavior*, 140, 556–574.

Marwick, A. E. (2015). Instafame: luxury selfies in the attention economy. *Public Culture*, 27(1 (75)), 137–160. Number: 1 (75).

Mohan, N. (2022). A look at 2022: community, collaboration, and commerce.

Muller, E. & Peres, R. (2019). The effect of social networks structure on innovation performance: A review and directions for research. *International Journal of Research in Marketing*, 36(1), 3–19.

Pagan, N., Mei, W., Li, C., & Dörfler, F. (2021). A meritocratic network formation model for the rise of social media influencers. *Nature Communications*, 12(1), 6865. Number: 1 Publisher: Nature Publishing Group.

Sadler, E. & Golub, B. (2022). Games on endogenous networks. Issue: arXiv:2102.01587 arXiv:2102.01587 [cs, econ].

Scorus, V. (2021). China tech express: bilibili releases 2020 influencer report - nativex.

Statista (2023). Biggest social media platforms 2023.

Staudigl, M. & Weidenholzer, S. (2014). Constrained interactions and social coordination. *Journal of Economic Theory*, 152, 41–63.

Sun, Y., Zhao, W., & Zhou, J. (2023). Structural interventions in networks. *International Economic Review*, 64(4), 1533–1563.

Toubia, O. & Stephen, A. T. (2013). Intrinsic vs. Image-related utility in social media: why do people contribute content to twitter? *Marketing Science*, 32(3), 368–392. Number: 3 Publisher: INFORMS.

Wu, S., Hofman, J. M., Mason, W. A., & Watts, D. J. (2011). Who says what to whom on twitter. In *Proceedings of the 20th international conference on World wide web*, WWW '11 (pp. 705–714). New York, NY, USA: Association for Computing Machinery.

Yanagizawa-Drott, D. (2014). Propaganda and Conflict: Evidence from the Rwandan Genocide *. *The Quarterly Journal of Economics*, 129(4), 1947–1994.

Young, H. P. (1993). The evolution of conventions. *Econometrica*, 61(1), 57–84. Number: 1 ISBN: 00129682 Publisher: JSTOR.

Zhang, X. M. & Zhu, F. (2011). Group size and incentives to contribute: a natural experiment at chinese wikipedia. *American Economic Review*, 101(4), 1601–1615. Number: 4.

Zhou, J. & Chen, Y.-J. (2015). Key leaders in social networks. *Journal of Economic Theory*, 157, 212–235.

Online Appendix

This supplementary material generalizes Propositions 2 and 3 and the remark in Footnote 15 in the main text to cover the case of $\alpha > 1$. Proposition 8 also considers an arbitrary $\delta \in [0,1)$ in the model with horizontal differentiation presented in Section 4.1 in the main text.

Generalization of Proposition 2

The following proposition extends Proposition 2 to the model with horizontal differentiation presented in Section 4.1 for any $\delta \in [0,1)$. Since the special case of $\delta = 0$ reduces to the baseline model without heterogenous preferences over content categories, Proposition 2 can also be proved as a corollary of the following proposition.

Proposition 8. *Consider the model with possible heterogenous preferences over content categories in Section 4.1.*

1. *If $\alpha > 1$ and $\alpha\beta - \alpha - \beta > 0$, then for each $\delta \in [0,1)$, there exists a strict equilibrium s with $\rho(s) = 1$ for each sufficiently large n.*
2. *If $\alpha > 1$, $\alpha\beta - \alpha - \beta \leq 0$, and $\delta = 0$, then $\rho(s) < \max\{\frac{\alpha\beta}{\alpha\beta+1} + \frac{1}{2(\alpha\beta+1)n}, \frac{\alpha\beta-\beta}{\alpha} + \frac{\alpha+\beta-\alpha\beta}{\alpha n}\}$.*
3. *If $\alpha \leq 1$, then for each $\delta \in [0,1)$, we have $\rho(s) < \frac{\alpha\beta}{\alpha\beta+1} + \frac{1}{(\alpha\beta+1)n}$ for each strict equilibrium s.*

The proof of Proposition 8 has three parts.

In **Part 1**, we prove that if $\alpha > 1$ and $\alpha\beta - \alpha - \beta > 0$, then there exists a strict equilibrium s such that $\rho(s) = 1$. Note that $\alpha\beta - \alpha - \beta > 0$ implies $\beta > 1$. There are three cases.

In Case 1, $(1-\delta)\alpha\beta - (1-\delta)\alpha - \beta > 0$. It suffices to show that the following strategy can be proved as a strict equilibrium: each player i provides $x_i = (\beta - 1)(n-1)$, and $g_{ij} = 1$ for any i, j. Player i has no incentive to deviate by deleting any l links and producing some x_i' since her payoff will be weakly lower than

$$\begin{aligned}
&\alpha x_i(\tfrac{n}{2} - 1) + (1-\delta)\alpha x_i(\tfrac{n}{2} - l) + \beta(n-1)x_i' - \tfrac{1}{2}(x_i' + n - 1 - l)^2 \\
\leq\ &\alpha x_i(\tfrac{n}{2} - 1) + (1-\delta)\alpha x_i(\tfrac{n}{2} - l) + \beta(n-1)(x_i + l) - \tfrac{1}{2}[\beta(n-1)]^2 \\
=\ &u_i(x,g) - [(1-\delta)\alpha x_i - \beta(n-1)]l \\
<\ &u_i(x,g).
\end{aligned}$$

In Case 2, $(1-\delta)\alpha\beta - (1-\delta)\alpha - \beta = 0$. It suffices to show that the following strategy can be proved as a strict equilibrium for sufficiently large n: for some integer $\frac{(\alpha\beta-\alpha-\beta)(n-2)}{2(\alpha\beta+\alpha-\beta)} > m >$

$\frac{(\alpha\beta-\alpha-2\beta)(n-2)-2\beta}{2(\alpha\beta+\alpha)}$,

$$x_i = \begin{cases} x_{[1]} = \beta(\frac{n}{2}-m-1)-(\frac{n}{2}+m-1) & \text{if } i \in \{1,...,\frac{n}{2}-m\} \cup \{\frac{n}{2}+1,...,n-m\} \\ x_{[2]} = \beta(n-1)-(2m-1) & \text{if } i \in \{\frac{n}{2}-m+1,...,\frac{n}{2}\} \cup \{n-m+1,...,n\} \end{cases},$$

$$\omega_i = \begin{cases} A & \text{if } i \in \{1,...,\frac{n}{2}\} \\ B & \text{if } i \in \{\frac{n}{2}+1,...,n\} \end{cases},$$

$$g_{ij} = \begin{cases} 1 & \text{if } x_j = x_{[2]} \\ 1 & \text{if } x_i = x_{[1]},\ \omega_j = \omega_i \\ 0 & \text{otherwise} \end{cases}.$$

Such $m \geq 1$ exists for sufficiently large n. First, similarly to Case 1, player i with $x_i = x_{[2]}$ has no incentive to deviate by deleting any l links and producing some x_i'. Second, player i with $x_i = x_{[2]}$ has no incentive to deviate by adding any l links and producing some x_i' since her payoff will be weakly lower than

$$\begin{aligned} & \alpha x_{[2]}(m-1)+(1-\delta)\alpha x_{[2]}m+\alpha x_{[1]}l+\beta(n-1)x_i'-\tfrac{1}{2}(x_i'+2m-1+l)^2 \\ \leq\ & \alpha x_{[2]}(m-1)+(1-\delta)\alpha x_{[2]}m+\alpha x_{[1]}l+\beta(n-1)(x_i-l)-\tfrac{1}{2}[\beta(n-1)]^2 \\ =\ & u_i(x,g)+[\alpha x_{[1]}-\beta(n-1)]l \\ <\ & u_i(x,g) \end{aligned}$$

where the last inequality holds since $m > \frac{(\alpha\beta-\alpha-2\beta)(n-2)-2\beta}{2(\alpha\beta+\alpha)}$. Similarly, since $(1-\delta)\alpha\beta-(1-\delta)\alpha-\beta = 0$, player i such that $x_i = x_{[1]}$ has no incentive to deviate by adding any l links and producing some x_i'. Third, player i with $x_i = x_{[1]}$ has no incentive to deviate by deleting any l links to some j with $x_j = x_{[1]}, \omega_j = \omega_i$ and producing some x_i' since her payoff will be weakly lower than

$$\begin{aligned} & \alpha x_{[2]}m+(1-\delta)\alpha x_{[2]}m+\alpha x_{[1]}(\tfrac{n}{2}-m-1-l)+\beta(\tfrac{n}{2}-m-1)x_i'-\tfrac{1}{2}(x_i'+\tfrac{n}{2}+m-1)^2 \\ \leq\ & \alpha x_{[2]}m+(1-\delta)\alpha x_{[2]}m+\alpha x_{[1]}(\tfrac{n}{2}-m-1-l)+\beta(\tfrac{n}{2}-m-1)(x_i+l)-\tfrac{1}{2}[\beta(\tfrac{n}{2}-m-1)]^2 \\ =\ & u_i(x,g)-[\alpha x_{[1]}-\beta(\tfrac{n}{2}-m-1)]l \\ <\ & u_i(x,g) \end{aligned}$$

where the last inequality holds since $m < \frac{(\alpha\beta-\alpha-\beta)(n-2)}{2(\alpha\beta+\alpha-\beta)}$. Forth, similarly to Case 1, player i with $x_i = x_{[1]}$ has no incentive to deviate by deleting any l links to some j with $x_j = x_{[2]}, \omega_j \neq \omega_i$.

In Case 3, $(1-\delta)\alpha\beta-(1-\delta)\alpha-\beta < 0$. It suffices to show that the following strategy can be proved as a strict equilibrium: for each player $i \in N_\omega$with $\omega = A, B$, $x_i = (\beta-1)(\frac{n}{2}-1)$,

$$g_{ij} = \begin{cases} 1 & \text{if } j \in N_\omega \\ 0 & \text{otherwise} \end{cases}.$$

Similarly to Case 1, player i has no incentive to deviate by deleting any l links and producing x_i'. Moreover, player i has no incentive to deviate by adding any l links and producing x_i' since her payoff will be

$$\begin{aligned} &\alpha x_i(\tfrac{n}{2}-1) + (1-\delta)\alpha x_i l + \beta(\tfrac{n}{2}-1)x_i' - \tfrac{1}{2}(x_i' + \tfrac{n}{2} - 1 + l)^2 \\ \le\ & \alpha x_i(\tfrac{n}{2}-1) + (1-\delta)\alpha x_i l + \beta(\tfrac{n}{2}-1)(x_i - l) - \tfrac{1}{2}[\beta(\tfrac{n}{2}-1)]^2 \\ =\ & u_i(x,g) + [(1-\delta)\alpha x_i - \beta(\tfrac{n}{2}-1)]l \\ <\ & u_i(x,g). \end{aligned}$$

In **Part 2**, we prove that if $\alpha > 1$, $\alpha\beta - \alpha - \beta \le 0$, and $\delta = 0$, then in a strict equilibrium s, $\rho(s)$ is lower than the cutoff as is specified in the proposition. By Proposition 1, g is a nested upward-linking network with $\bar{t}$ tiers and n_t players in each tier t. Since all players in the same tier provide the same level of content in a strict equilibrium, we let $x_{[t]}$ denote the content provision of tier t players. Moreover, $g_{ij} = 0$ if $x_j < x_i$. Similarly to Case 1 in Part 1, we can prove by contradiction that $x_{[1]} = 0$ by showing that otherwise each player in tier 1 has weak incentive to deviate by deleting a link. There are two cases. In Case 1, for each i and j in tier 2, $g_{ij} = 0$. Then $x_{[2]} = \beta n_1 - (n - n_1 - n_2)$. To support it as a strict equilibrium, each player i in tier 1 has no incentive to deviate by deleting a link, which requires that

$$\alpha x_{[2]} = \alpha[\beta n_1 - (n - n_1 - n_2)] > \tfrac{1}{2}\left[(n-n_1)^2 - (n-n_1-1)^2\right] = n - n_1 - \tfrac{1}{2}.$$

Since $n - n_1 - n_2 \ge 0$, we have $\alpha\beta n_1 > n - n_1 - \frac{1}{2}$. Thus, $1 - \rho(s) = \frac{n_1}{n} > \frac{1}{\alpha\beta+1} - \frac{1}{2(\alpha\beta+1)n}$. In Case 2, for each i and j in tier 2, $g_{ij} = 1$. Then $x_{[2]} = \beta(n_1 + n_2 - 1) - (n - n_1 - 1)$. To support it as a strict equilibrium, each player i in tier 2 has no incentive to deviate by deleting a link and deviating to $x_i' = x_{[2]} + 1$, which requires that

$$\alpha x_{[2]} - \beta(n_1 + n_2 - 1) = \alpha[\beta(n_1 + n_2 - 1) - (n - n_1 - 1)] - \beta(n_1 + n_2 - 1) > 0.$$

Thus $\beta(\alpha - 1)(n_1 + n_2 - 1) - \alpha(n - n_1 - 1) > 0$. Since $n_1 + n_2 \le n$, we have $\beta(\alpha-1)(n-1) - \alpha(n - n_1 - 1) > 0$. Thus, $1 - \rho(s) > \frac{\alpha+\beta-\alpha\beta}{\alpha} - \frac{\alpha+\beta-\alpha\beta}{\alpha n}$.

In **Part 3**, we prove that if $\alpha \le 1$, then in a strict equilibrium s, $\rho(s)$ is lower than the cutoff as is specified in the proposition. Similarly to Lemma 4, since $\alpha \le 1$, we have $g_{ij} = 0$ if $x_j \le x_i$. For $\omega = A$ or B, respectively, define a network g^ω which consists of players who provide good ω or does not provide but belongs to N_ω such that (1) we order g^ω according to content provision level; (2) it has $\bar{t}^\omega$ tiers and n_t^ω players in tier t. We let $x_{[t]}^\omega$ denote the content provision of tier t players in the network g^ω and $\bar{n}^\omega = \sum_{t=1}^{\bar{t}^\omega} n_t^\omega$. In a strict equilibrium, for i, j in the same tier t of the same network g^ω,$d_i^{out} = d_j^{out}$ and we denote it as $d_{[t]}^\omega$. By definition, each player i belongs to one tier of either g^Aor g^B. There are three cases.

In Case 1,for each ω and each player i in tier 1 of network g^ω, $g_{ij} = 0$ if j is in tier t of $g^{\omega'}$ where $t = \min\{\widetilde{t}|x_{\widetilde{t}}^{\omega'} > 0\}$. Then in each g_ω, $x_{[1]}^\omega$= 0, $x_{[2]}^\omega = \beta n_1^\omega - d_{[2]}^\omega$ and $g_{ij} = 1$ for each i in tier 1 of $g^\omega$and each j in tier $t \geq 2$ of g^ω. To support s as a strict equilibrium, player i in tier 1 of g^ωhas no incentive to deviate by deleting a link, which requires that

$$\alpha(\beta n_1^\omega - d_{[2]}^\omega) > \frac{1}{2}\left[(d_i^{out})^2 - (d_i^{out} - 1)^2\right] = d_i^{out} - \frac{1}{2} \geq \bar{n}^\omega - n_1^\omega - \frac{1}{2}$$

where the last inequality follows from $d_i^{out} \geq \bar{n}^\omega - n_1^\omega$. Since $d_{[2]}^\omega \geq 0$, we have $\alpha\beta n_1^\omega > \bar{n}^\omega - n_1^\omega - \frac{1}{2}$, that is, $(\alpha\beta + 1)n_1^\omega > \bar{n}^\omega - \frac{1}{2}$. Then $(\alpha\beta + 1)\sum_{\omega\in\{A,B\}} n_1^\omega > \sum_{\omega\in\{A,B\}} \bar{n}^\omega - 1 = n - 1$, that is $1 - \rho(s) = \frac{\sum_{\omega\in\{A,B\}} n_1^\omega}{n} > \frac{1}{\alpha\beta+1} - \frac{1}{(\alpha\beta+1)n}$.

In Case 2, there exists a unique ω such that a player i in tier 1 of network g^ω, $g_{ij} = 1$ if j is in tier t of $g^{\omega'}$ where $t = \min\{\widetilde{t}|x_{\widetilde{t}}^{\omega'} > 0\}$. Without loss of generality, let $\omega = A$. Thus, in g_A, $x_{[1]}^A$= 0, $x_{[2]}^A = \beta n_1^A - d_{[2]}^A$ and $g_{ij} = 1$ for each i in tier 1 of $g^A$and each j with $x_j > 0$. There are two subcases. In Subcase 1, $x_{[1]}^B > 0$. To support s as a strict equilibrium, player i in tier 1 of g^Ahas no incentive to deviate to deleting a link, which requires that

$$\alpha(\beta n_1^A - d_{[2]}^A) > \frac{1}{2}\left[(d_i^{out})^2 - (d_i^{out} - 1)^2\right] = \frac{1}{2}\left[(n - n_1^A)^2 - (n - n_1^A - 1)^2\right] = n - n_1^A - \frac{1}{2}.$$

Since $d_{[2]}^A \geq 0$, we have $\alpha\beta n_1^A > n - n_1^A - \frac{1}{2}$, that is, $(\alpha\beta + 1)n_1^A > n - \frac{1}{2}$. Then $1 - \rho(s) = \frac{n_1^A}{n} > \frac{1}{\alpha\beta+1} - \frac{1}{2(\alpha\beta+1)n}$. In Subcase 2, $x_{[1]}^B = 0$. To support s as a strict equilibrium, player i in tier 1 of g^B has no incentive to deviate to deleting a link to a player in tier 2 of g^B, which requires that

$$\begin{aligned}(1-\delta)\alpha[\beta(n_1^A + n_1^B) - d_{[2]}^B] &> \frac{1}{2}\left[(d_i^{out})^2 - (d_i^{out} - 1)^2\right] \\ &= \frac{1}{2}\left[(n - n_1^A - n_1^B)^2 - (n - n_1^A - n_1^B - 1)^2\right] \\ &= n - n_1^A - n_1^B - \frac{1}{2}.\end{aligned}$$

Since δ≥0 , $d_{[2]}^B \geq 0$, similarly to Subcase 1, we have $1 - \rho(s) = \frac{n_1^A + n_1^B}{n} > \frac{1}{\alpha\beta+1} - \frac{1}{2(\alpha\beta+1)n}$.

In Case 3, for each g^ω, there exists a player i in tier 1 of network g^ω, $g_{ij} = 1$ if j is in tier t of $g^{\omega'}$ where $t = \min\{\widetilde{t}|x_{\widetilde{t}}^{\omega'} > 0\}$. First, since $g_{ij} = 0$ if $x_j \leq x_i$, there exists ω such that $x_{[1]}^\omega$= 0. Without loss of generality, let $x_{[1]}^A$= 0. There are two subcases. In Subcase 1, $x_{[1]}^B$> 0. Then $g_{ij} = 1$ for each i in tier 1 of $g^A$and each j with $x_j > 0$. Then similarly to Subcase 1 in Case 2, $1 - \rho(s) > \frac{1}{\alpha\beta+1} - \frac{1}{2(\alpha\beta+1)n}$. In Subcase 2, $x_{[1]}^B$= 0. Since $g_{ij} = 0$ if $x_j \leq x_i$, there exists ω such that $g_{ij} = 0$ for each i in tier 2 of $g^\omega$and each j in tier 2 of $g^{\omega'}$. Without loss of generality, let g^A be such network. Then similarly to Subcase 2 in Case 2, $1-\rho(s) > \frac{1}{\alpha\beta+1} - \frac{1}{2(\alpha\beta+1)n}$.

Generalization of Proposition 3

The following proposition generalizes Proposition 3 to the case in which $\alpha > 1$. Then Proposition 3 is a special case of this proposition with $\alpha \leq 1$.

Proposition 9. *Suppose that $\alpha^2\beta+\alpha > 1$. Then there exists some sufficiently large n^* such that, for each $n \geq n^*$, we have*

$$\rho(s) \geq \begin{cases} \frac{\alpha^2\beta^2+\alpha\beta-\beta}{\alpha^2\beta^2+\alpha^2\beta+2\alpha\beta+\alpha} & \text{if } \alpha \leq 1 \\ \frac{\alpha\beta-\beta}{2(\alpha\beta+\alpha-\beta)} & \text{if } \alpha > 1 \text{ and } \alpha\beta-\alpha-\beta \leq 0 \\ \frac{\alpha\beta-\alpha-\beta}{2(\alpha\beta-\beta)} & \text{if } \alpha > 1 \text{ and } \alpha\beta-\alpha-\beta > 0 \end{cases}$$

for each payoff dominant strict equilibrium $s=(x,g)$.

Proof of Proposition 9: As in the proof for Proposition 3, it suffices to consider nonempty network equilibria. Consider a strict equilibrium $s=(x,g)$ with $\bar{t}$ tiers and n_t players in each tier t of g. According to Proposition 1, all players in the same tier provide the same level of content in a strict equilibrium. We let $x_{[t]}$ denote the content provision of tier t players. It suffices to show that there exists n^* and p^* such that, for each $n > n^*$, if $x_{[1]} = 0$ and $\frac{n_1}{n} > 1-p^*$, then (x,g) is not a payoff-dominant equilibrium. The case when $\alpha \leq 1$ is proved in Proposition 3. Then it remains to consider the case in which $\alpha > 1$. If $n_2 > 1$ and $g_{ij} = 0$ for any i, j that belong to tier 2, then $x_{[2]} = \beta n_1 - (n-n_1-n_2)$. To support it as a strict equilibrium, each player i in tier 2 has no incentive to deviate by adding a link to some j in the same tier and $x'_i = \beta n_1 - (n-n_1-n_2+1)$, that is,

$$\begin{aligned} & u_i(x',g') + \alpha x_{[2]} - \beta n_1 < u_i(x',g') \\ \Leftrightarrow\quad & \alpha x_{[2]} < \beta n_1 \\ \Rightarrow\quad & n_1 < \frac{\alpha n}{\alpha\beta+\alpha-\beta} \\ \Leftrightarrow\quad & \rho(s) > \frac{\alpha\beta-\beta}{\alpha\beta+\alpha-\beta} > \begin{cases} \frac{\alpha\beta-\beta}{2(\alpha\beta+\alpha-\beta)} & \text{if } \alpha\beta-\alpha-\beta \leq 0 \\ \frac{\alpha\beta-\alpha-\beta}{2(\alpha\beta-\beta)} & \text{if } \alpha\beta-\alpha-\beta > 0 \end{cases}. \end{aligned}$$

If $n_1 \leq n_2+1$, then

$$\frac{n_1}{n} \leq \frac{1}{2} + \frac{1}{2n} \Rightarrow \rho(s) \geq \frac{1}{2} - \frac{1}{2n}.$$

Since $\frac{\alpha\beta-\beta}{2(\alpha\beta+\alpha-\beta)} < \frac{1}{2}$ and $\frac{\alpha\beta-\alpha-\beta}{2(\alpha\beta-\beta)} < \frac{1}{2}$, there exists some m_1^* such that for $n > m_1^*$,

$$\frac{1}{2} - \frac{1}{2n} > \begin{cases} \frac{\alpha\beta-\beta}{2(\alpha\beta+\alpha-\beta)} & \text{if } \alpha\beta-\alpha-\beta \leq 0 \\ \frac{\alpha\beta-\alpha-\beta}{2(\alpha\beta-\beta)} & \text{if } \alpha\beta-\alpha-\beta > 0 \end{cases}.$$

Thus it remains to consider $n_1 > n_2 + 1$ with $n_2 = 1$ or $g_{ij} = 1$ for any i, j in tier 2 in g. There are three cases.

In **Case 1**, $\beta n_2 - \alpha - \frac{1}{2} > 0$ and $\alpha\beta - \alpha - \beta \leq 0$. We will show that there exists m_2^* such that, for any $n > m_2^*$, if $\frac{n_1}{n} > \frac{\alpha}{\alpha\beta+\alpha-\beta} + \frac{2\alpha\beta-2\alpha-2\beta+1}{2(\alpha\beta+\alpha-\beta)n}$, then (x, g) is not a payoff dominant equilibrium. Note that, if $2\alpha\beta - 2\alpha - 2\beta + 1 \leq 0$, then $\frac{\alpha}{\alpha\beta+\alpha-\beta} + \frac{2\alpha\beta-2\alpha-2\beta+1}{2(\alpha\beta+\alpha-\beta)n} \leq \frac{\alpha}{\alpha\beta+\alpha-\beta} < 1$ and $\frac{\alpha}{\alpha\beta+\alpha-\beta} + \frac{2\alpha\beta-2\alpha-2\beta+1}{2(\alpha\beta+\alpha-\beta)n} > \frac{2\alpha\beta-2\beta+1}{2(\alpha\beta+\alpha-\beta)} > 0$; if $2\alpha\beta - 2\alpha - 2\beta + 1 > 0$, then

$$0 < \frac{\alpha}{\alpha\beta+\alpha-\beta} + \frac{2\alpha\beta-2\alpha-2\beta+1}{2(\alpha\beta+\alpha-\beta)n} \leq \frac{2\alpha\beta-2\beta+1}{2(\alpha\beta+\alpha-\beta)} < 1.$$

Pick a positive integer k that satisfies

$$\frac{(\alpha\beta-\beta+\alpha)n_1-\alpha n+\alpha+\beta-\alpha\beta-\beta n_2}{\alpha} < k < \frac{2(\alpha\beta-\beta+\alpha)n_1-2\alpha n+2\alpha+2\beta-2\alpha\beta-1}{2\alpha}.$$

Note that since $\beta n_2 - \alpha - \frac{1}{2} > 0$,

$$\frac{(\alpha\beta-\beta+\alpha)n_1-\alpha n+\alpha+\beta-\alpha\beta-\beta n_2}{\alpha} < \frac{2(\alpha\beta-\beta+\alpha)n_1-2\alpha n+2\alpha+2\beta-2\alpha\beta-1}{2\alpha} - 1.$$

Moreover, since $\alpha\beta - \alpha - \beta \leq 0$ and $\alpha > 1$,

$$\begin{aligned} \frac{2(\alpha\beta-\beta+\alpha)n_1-2\alpha n+2\alpha+2\beta-2\alpha\beta-1}{2\alpha} &\leq \frac{4\alpha n_1-2\alpha n+2\alpha+2\beta-2\alpha\beta-1}{2\alpha} \\ &\leq \frac{2\alpha n_1-2\alpha+2\alpha+2\beta-2\alpha\beta-1}{2\alpha} \\ &= n_1 + \frac{2\beta-2\alpha\beta-1}{2\alpha} \\ &< n_1 \end{aligned}$$

where the second inequality holds since $n_1 \leq n-1$. And since $\frac{n_1}{n} > \frac{\alpha}{\alpha\beta+\alpha-\beta} + \frac{2\alpha\beta-2\alpha-2\beta+1}{2(\alpha\beta+\alpha-\beta)n}$, there exists some m_2^* such that for any $n > m_2^*$,

$$\begin{aligned} 2(\alpha\beta-\beta+\alpha)n_1-2\alpha n+2\alpha+2\beta-2\alpha\beta-1 &= 2n(\alpha\beta-\beta+\alpha)\left(\frac{n_1}{n} - \frac{\alpha}{\alpha\beta+\alpha-\beta} - \frac{2\alpha\beta-2\alpha-2\beta+1}{2(\alpha\beta+\alpha-\beta)n}\right) \\ &> 1 \end{aligned}$$

and therefore such integer $k \geq 1$ exists. Similarly to the proof for Proposition 3, we can construct strategy profile $s' = (x', g')$ which will be shown to be a Pareto improvement and an equilibrium. First, let g' has $\bar{t}' = \bar{t} + 1$ tiers, such that (1) for all players in each of the tiers $t \geq 2$ in g, they are placed in tier $t + 1$ in g' and provide exactly the same content as before, i.e., $x'_{[t+1]} = x_{[t]}$ for each $t \geq 2$; (2) for players in tier 1 in g, they are divided into two subsets, among which k players are placed in tier 2 in g', while the remaining players stay in tier 1 and provide $x_i = 0$. Let $N'_2 \subset N$ denote the set of those k players in tier 2 in g', and we let $x'_i = \beta(n_1 - 1) - (n - n_1 + k - 1)$ for each $i \in N'_2$. Let $N'_1 \subset N$ denote the set of those $n_1 - k$ players in tier 1 in g'. Second, set $g'_{ij} = 1$ if $i \in N'_1 \cup N'_2$ and $j \in N'_2$; otherwise, set $g'_{ij} = g_{ij}$. There are two steps.

In Step 1, we show that (x', g') is an equilibrium. First, for each player i that belongs to tier $t \geq 3$ in g', since (x, g) is a strict equilibrium, and $x'_i = x_i, g'_{ij} = g_{ij}$ and $g'_{ji} = g_{ji}$ for any j, it suffices to show that player does not have incentive to deviate to adding l links to players in N'_2 and $x'_i = \beta(n_1 + n_2 - 1) - (n - n_1 - 1 + l)$. Since $k > \frac{(\alpha\beta+\alpha-\beta)n_1-\alpha n+\alpha+\beta-\alpha\beta-\beta n_2}{\alpha}$, $\alpha x'_{[2]} < \beta(n_1 + n_2 - 1)$. Thus, player i has no incentive to deviate. Second, for each $i \in N'_2$, suppose that she deletes l outward links. Then her optimal content provision would be $x''_i = \beta(n_1 - 1) - (n - n_1 + k - 1 - l)$, and her payoff would be weakly lower than

$$\begin{aligned} u_i(x', g') - \alpha l x'_{[2]} + \beta(n_1 - 1)l &= u_i(x', g') - l\{\alpha[\beta(n_1 - 1) - (n - n_1 + k - 1)] - \beta(n_1 - 1)\} \\ &< u_i(x', g') \end{aligned}$$

where the last inequality holds since $k < \frac{2(\alpha\beta+\alpha-\beta)n_1-2\alpha n+2\alpha+2\beta-2\alpha\beta-1}{2\alpha} < \frac{(\alpha\beta+\alpha-\beta)n_1-\alpha n+\alpha+\beta-\alpha\beta}{\alpha}$. Third, each player i inN'_1 has no incentive to deviate by deleting l links to players in N'_2, that is,

$$\begin{aligned} \alpha[\beta(n_1 - 1) - (n - n_1 + k - 1)] &> \beta(n_1 - 1) + \tfrac{1}{2} \\ &> n - n_1 + k - 1 + \tfrac{1}{2} \\ &= n - n_1 + k - \tfrac{1}{2} \\ &> n - n_1 + k - \tfrac{l}{2} \end{aligned}$$

where the first inequality holds since $k < \frac{2(\alpha\beta+\alpha-\beta)n_1-2\alpha n+2\alpha+2\beta-2\alpha\beta-1}{2\alpha}$ and the third holds since $\beta(n_1 - 1) > n - n_1 + k - 1$. Thus,

$$-\alpha l[\beta(n_1 - 1) - (n - n_1 + k - 1)] - \tfrac{1}{2}(n - n_1 + k - l)^2 + \tfrac{1}{2}(n - n_1 + k)^2 < 0.$$

In Step 2, we can show that $u_i(x', g') \geq u_i(x, g)$ for each $i \in N$, and that the inequality holds strictly for each $i \in N'_1 \cup N'_2$. First, for any player i in tier $t > 2$ in g', $u_i(x', g') = u_i(x, g)$. Second, for each $i \in N'_2$, $u_i(x, g)$ equals to her payoff in (x', g') with deviation to zero content provision and deleting all links to players in N'_2, which is lower than her payoff in (x', g') with deviation to $x''_i = \beta(n_1 - 1) - (n - n_1)$ and deleting all links to players in N'_2, that is strictly lower than her payoff in (x', g') as is shown in Step 1. Third, for each $i \in N'_1$, $u_i(x, g)$ equals her payoff in (x', g') when deviated to deleting her links to all players in N_2',which is strictly lower than $u_i(x',g')$ as is shown in Step 1.

To conclude, for each payoff-dominant equilibrium (x,g) and each $n>m_2^*$, we have

$$\frac{n_1}{n}\leq\frac{\alpha}{\alpha\beta+\alpha-\beta}+\frac{2\alpha\beta-2\alpha-2\beta+1}{2(\alpha\beta+\alpha-\beta)n}.$$

Thus,

$$\begin{aligned}\rho(s)&\geq\frac{\alpha\beta-\beta}{\alpha\beta+\alpha-\beta}-\frac{2\alpha\beta-2\alpha-2\beta+1}{2(\alpha\beta+\alpha-\beta)n}\\&\geq\frac{\alpha\beta-\beta}{\alpha\beta+\alpha-\beta}-\frac{\alpha\beta-\beta}{2(\alpha\beta+\alpha-\beta)n}\\&>\frac{\alpha\beta-\beta}{\alpha\beta+\alpha-\beta}-\frac{\alpha\beta-\beta}{2(\alpha\beta+\alpha-\beta)}=\frac{\alpha\beta-\beta}{2(\alpha\beta+\alpha-\beta)}.\end{aligned}$$

In **Case 2**, $\beta n_2-\alpha-\frac{1}{2}\leq 0$ and $\alpha\beta-\alpha-\beta\leq 0$. We will show that there exists m_3^* such that, for any $n>m_3^*$, if $\frac{n_1}{n}>\frac{\alpha}{\alpha\beta+\alpha-\beta}+\frac{(\alpha\beta-\alpha-\beta)(1-n_2)}{(\alpha\beta+\alpha-\beta)n}$, then $n_1\geq 3$ and (x,g) is not a payoff dominant equilibrium. Note that, since $\alpha\beta-\alpha-\beta\leq 0$ and $\beta n_2-\alpha-\frac{1}{2}\leq 0$,

$$\frac{\alpha}{\alpha\beta+\alpha-\beta}+\frac{(\alpha\beta-\alpha-\beta)(1-n_2)}{(\alpha\beta+\alpha-\beta)n}<\frac{\alpha}{\alpha\beta+\alpha-\beta}-\frac{(\alpha\beta-\alpha-\beta)(\frac{\alpha+\frac{1}{2}}{\beta}-1)}{(\alpha\beta+\alpha-\beta)n}<1$$

for some m_4^* and any $n>m_4^*$. Moreover, there exists m_5^* such that, for any $n>m_5^*$, $\frac{\alpha}{\alpha\beta+\alpha-\beta}+\frac{(\alpha\beta-\alpha-\beta)(1-n_2)}{(\alpha\beta+\alpha-\beta)n}>\frac{1}{1+\beta}+\frac{(2\alpha-\beta)n_2+\beta}{(1+\beta)n}$. Thus,

$$\begin{aligned}\frac{n_1}{n}>\frac{1}{1+\beta}+\frac{(2\alpha-\beta)n_2+\beta}{(1+\beta)n}&\Rightarrow\beta(n_1+n_2-1)-(n-n_1)>2\alpha n_2\\&\Rightarrow[\beta(n_1+n_2-1)-(n-n_1)]^2>(2\alpha n_2)^2\geq 4\alpha n_2.\end{aligned}$$

Pick a positive integer $k<n_1$ that satisfies

$$2\leq k<\frac{(\alpha\beta-\beta+\alpha)n_1-\alpha n-(\alpha+\beta-\alpha\beta)n_2+\alpha+\beta-\alpha\beta}{\alpha}.$$

Note that since $\frac{n_1}{n}>\frac{\alpha}{\alpha\beta+\alpha-\beta}+\frac{(\alpha\beta-\alpha-\beta)(1-n_2)}{(\alpha\beta+\alpha-\beta)n}$, we have

$$(\alpha\beta-\beta+\alpha)n_1-\alpha n-(\alpha+\beta-\alpha\beta)n_2+\alpha+\beta-\alpha\beta>0.$$

Similarly to that in Case 1, there exists m_6^* such that, for any $n>m_6^*$ such integer k exists. Let $m_3^*=\max\{m_4^*,m_5^*,m_6^*\}$. We can construct strategy profile $s'=(x',g')$ which will be shown to be a Pareto improvement and an equilibrium. First, let g' has $\bar{t}$ tiers, such that (1) for all players in each of the tiers $t\geq 3$ in g, they are placed in tier t in g' and provide exactly the same content as before, i.e., $x'_{[t]}=x_{[t]}$ for each $t\geq 3$; (2) all players in tier 2 in g stay in tier 2 in g' and provide $x'_i=\beta(n_1+n_2-1)-(n-n_1+k-1)$; (3) players in tier 1 in g are divided into two subsets, among which k players are placed in tier 2 in g' and provide $\beta(n_1+n_2-1)-(n-n_1+k-1)$;, while the remaining players stay in tier 1 and provide $x_i=0$. Let $N_2'\subset N$ denote the set of those k players in tier 2 in g' who were in tier 1 in g. Let $N_1'\subset N$ denote the set of those n_1-k players in tier 1 in g'. Second, set $g'_{ij}=1$ for i in tier $t\leq 2$ in g and $j\in N_2'$; otherwise, set $g'_{ij}=g_{ij}$. Similarly to Case 1, there are two steps.

In Step 1, we show that (x',g') is an equilibrium. First, for each player i that belongs to tier 2 in g', since $k<\frac{(\alpha\beta-\beta+\alpha)n_1-\alpha n-(\alpha+\beta-\alpha\beta)n_2+\alpha+\beta-\alpha\beta}{2\alpha}$,

$$\alpha x'_{[2]}=\alpha[\beta(n_1+n_2-1)-(n-n_1+k-1)]>\beta(n_1+n_2-1)+\alpha n_2>\beta(n_1+n_2-1).$$

Thus, player i has no incentive to deviate by deleting l links to players in tier 2 in g' and $x'_i=\beta(n_1+n_2-1)-(n-n_1+k-1-l)$. As a result, player i has no incentive to deviate by deleting any links. Second, since

$$\begin{aligned}\alpha x'_{[2]}&=\alpha[\beta(n_1+n_2-1)-(n-n_1+k-1)]\\&>\beta(n_1+n_2-1)+\alpha n_2\\&\geq\beta(n_1+n_2-1)+1\\&>n-n_1+k-\tfrac{1}{2},\end{aligned}$$

each $i\in N_1'$ has no incentive to deviate by deleting her links. Third, for each player i that belongs to tier $t>2$ in g, since (x,g) is a strict equilibrium, $x'_i=x_i, x'_{[2]}<x_{[2]}, g'_{ij}=g_{ij}$ and $g'_{ji}=g_{ji}$ for each j, player i has no incentive to deviate by adding some links to player j in tier 2 in g' or any other possible deviations.

In Step 2, similarly to Step 2 in Case 1, we can show that $u_i(x',g')\geq u_i(x,g)$ for each $i\in N$, and that the inequality holds strictly for each i in tier 2 in g. First, for any player i in tier $t>2$ in g', $u_i(x',g')=u_i(x,g)$. Second, for each i in tier 2 in g, $u_i(x,g)$ equals to her payoff in (x',g') with deviation to deleting her links to players in N_2' and x_i, which is strictly higher than $u_i(x',g')$ as is shown in Step 1. Third, for each $i\in N_2'$,

$$\begin{aligned}&u_i(x',g')-u_i(x,g)\\&=\alpha n_2(x'_{[2]}-x_{[2]})+\alpha(k-1)x'_{[2]}+\beta(n_1+n_2-1)x'_{[2]}-\tfrac{1}{2}[\beta(n_1+n_2-1)]^2+\tfrac{1}{2}(n-n_1)^2\\&=-\alpha n_2k+\alpha(k-1)x'_{[2]}+\tfrac{1}{2}[\beta(n_1+n_2-1)-(n-n_1)]^2-(k-1)\beta(n_1+n_2-1)\\&>-\alpha n_2k+\alpha(k-1)x'_{[2]}+\alpha n_2-(k-1)\beta(n_1+n_2-1)\\&=(k-1)[\alpha x'_{[2]}-\alpha n_2-\beta(n_1+n_2-1)]\\&>0\end{aligned}$$

where the third inequality holds since $[\beta(n_1+n_2-1)-(n-n_1)]^2 > 4\alpha n_2$ and the last inequality holds since

$$\alpha x'_{[2]} = \alpha[\beta(n_1+n_2-1)-(n-n_1+k-1)] > \beta(n_1+n_2-1)+\alpha n_2.$$

Fourth, for each $i \in N'_1$,

$$\begin{aligned}
&u_i(x',g')-u_i(x,g)\\
&=\alpha n_2(x'_{[2]}-x_{[2]})+\alpha k x'_{[2]}-\tfrac{1}{2}(n-n_1+k)^2+\tfrac{1}{2}(n-n_1)^2\\
&=-\alpha n_2 k+\alpha k x'_{[2]}-k(n-n_1+\tfrac{k}{2})\\
&=k[\alpha x'_{[2]}-\alpha n_2-(n-n_1+\tfrac{k}{2})]\\
&>0
\end{aligned}$$

where the last inequality holds since

$$\begin{aligned}
\alpha x'_{[2]} &= \alpha[\beta(n_1+n_2-1)-(n-n_1+k-1)]\\
&> \beta(n_1+n_2-1)+\alpha n_2\\
&> n-n_1+k-1+\alpha n_2\\
&\geq n-n_1+\tfrac{k}{2}+\alpha n_2.
\end{aligned}$$

To conclude, for each payoff-dominant equilibrium (x,g) and each $n > m_3^*$, we have

$$\begin{aligned}
\tfrac{n_1}{n} &\leq \tfrac{\alpha}{\alpha\beta+\alpha-\beta}+\tfrac{(\alpha\beta-\alpha-\beta)(1-n_2)}{(\alpha\beta+\alpha-\beta)n}\\
&\leq \tfrac{\alpha}{\alpha\beta+\alpha-\beta}+\tfrac{(\alpha\beta-\alpha-\beta)(1-n+n_1)}{(\alpha\beta+\alpha-\beta)n}\\
&= \tfrac{-\alpha\beta+2\alpha+\beta}{\alpha\beta+\alpha-\beta}+\tfrac{\alpha\beta-\alpha-\beta}{(\alpha\beta+\alpha-\beta)n}+\tfrac{(\alpha\beta-\alpha-\beta)n_1}{(\alpha\beta+\alpha-\beta)n}.
\end{aligned}$$

Thus,

$$\tfrac{n_1}{n} \leq \tfrac{-\alpha\beta+2\alpha+\beta}{2\alpha}+\tfrac{\alpha\beta-\alpha-\beta}{2\alpha n} \Leftrightarrow \rho(s) \geq \tfrac{\alpha\beta-\beta}{2\alpha}+\tfrac{-\alpha\beta+\alpha+\beta}{2\alpha n} > \tfrac{\alpha\beta-\beta}{2\alpha} > \tfrac{\alpha\beta-\beta}{2(\alpha\beta+\alpha-\beta)}.$$

In **Case 3**, $\alpha\beta-\alpha-\beta > 0$. Thus $\alpha\beta > \alpha+\beta > \alpha$. We will show that there exists m_7^* such that, for any $n > m_7^*$, if $\frac{n_1}{n} > \frac{\alpha}{\alpha\beta-\beta}+\frac{2\alpha\beta-2\alpha-2\beta+1}{2(\alpha\beta-\beta)n}$, then $n_1 \geq 3$ and (x,g) is not a payoff dominant equilibrium. Note that for some m_8^* and each $n > m_8^*$,

$$\tfrac{2\alpha\beta-2\alpha-2\beta+1}{2(\alpha\beta-\beta)n} < \tfrac{\alpha\beta-\alpha-\beta}{2(\alpha\beta-\beta)}$$

and therefore

$$\tfrac{\alpha}{\alpha\beta-\beta}+\tfrac{2\alpha\beta-2\alpha-2\beta+1}{2(\alpha\beta-\beta)n} < \tfrac{\alpha}{\alpha\beta-\beta}+\tfrac{\alpha\beta-\alpha-\beta}{2(\alpha\beta-\beta)} < \tfrac{\alpha}{\alpha\beta-\beta}+\tfrac{\alpha\beta-\alpha-\beta}{\alpha\beta-\beta} = 1.$$

Let $m_7^* = \max\{m_2^*, m_3^*, m_8^*\}$. If $\frac{(\alpha\beta-\beta+\alpha)n_1-\beta n_2-\alpha n+\alpha+\beta-\alpha\beta}{\alpha} < n_1$, then we can construct strategy profile $s' = (x', g')$ as in Case 1 with $k = n_1$ such that (1) g' has $\bar{t}$ tiers, and all players in each of the tiers $t \geq 2$ in g are placed in tier t in g' and provide exactly the same content as before, i.e., $x'_{[t]} = x_{[t]}$ for each $t \geq 2$; (2) each player i in tier 1 in g stays in tier 1 and provides $x'_i = \beta(n_1 - 1) - (n-1)$; (3) $g'_{ij} = 1$ for i and j in tier 1 in g', and otherwise, $g'_{ij} = g_{ij}$. Since $\frac{n_1}{n} > \frac{\alpha}{\alpha\beta-\beta} + \frac{2\alpha\beta-2\alpha-2\beta+1}{2(\alpha\beta-\beta)n}$, then $k = n_1 < \frac{2(\alpha\beta-\beta+\alpha)n_1-2\alpha n+2\alpha+2\beta-2\alpha\beta-1}{2\alpha}$. Moreover, $\frac{(\alpha\beta-\beta+\alpha)n_1-\beta n_2-\alpha n+\alpha+\beta-\alpha\beta}{\alpha} < n_1 = k$. Thus, we can show (x', g') is a Pareto improvement and an equilibrium similarly to Case 1. If $\frac{(\alpha\beta-\beta+\alpha)n_1-\beta n_2-\alpha n+\alpha+\beta-\alpha\beta}{\alpha} \geq n_1$, then we can construct strategy profile $s' = (x', g')$ as in Case 2 with $2 \leq k < \frac{(\alpha\beta-\beta+\alpha)n_1-\alpha n-(\alpha+\beta-\alpha\beta)n_2+\alpha+\beta-\alpha\beta}{\alpha}$ and $k \leq n_1 - 1$ such that (1) g' has $\bar{t}$ tiers, and all players in each of the tiers $t \geq 3$ in g are placed in the same tier t in g' and provide exactly the same content as before, i.e., $x'_{[t]} = x_{[t]}$ for each $t \geq 3$; (2) each player i in tier 2 in g stays in tier 2 and provides $x'_i = \beta(n_1 + n_2 - 1) - (n - n_1 + k - 1)$; (3) players in tier 1 in g are divided into two subsets, among which k players are placed in tier 2 in g' and provide $\beta(n_1 + n_2 - 1) - (n - n_1 + k - 1)$, while the remaining players stay in tier 1 and provide $x_i = 0$; (4) $g'_{ij} = 1$ for i in tier $t \leq 2$ in g' and j in tier 2 in g', and otherwise, $g'_{ij} = g_{ij}$. Note that since $\frac{(\alpha\beta-\beta+\alpha)n_1-\beta n_2-\alpha n+\alpha+\beta-\alpha\beta}{\alpha} \geq n_1$,

$$
\begin{aligned}
(\alpha\beta - \beta + \alpha)n_1 - \alpha n - (\alpha + \beta - \alpha\beta)n_2 + \alpha + \beta - \alpha\beta &\geq \alpha n_1 - (\alpha - \alpha\beta)n_2 \\
&\geq \alpha n_2 + \alpha - (\alpha - \alpha\beta)n_2 \\
&= \alpha\beta n_2 + \alpha \\
&> 2\alpha
\end{aligned}
$$

where the second inequality holds since $n_1 \geq n_2+1$ and the last inequality holds since $\alpha\beta > \alpha$ and $n_2 \geq 1$. Thus, $\frac{(\alpha\beta-\beta+\alpha)n_1-\alpha n-(\alpha+\beta-\alpha\beta)n_2+\alpha+\beta-\alpha\beta}{\alpha} > 2$ and k exists. Then we can show (x', g') is a Pareto improvement and an equilibrium similarly to Case 2.

To conclude, for each payoff-dominant equilibrium (x, g) and each $n > m_7^*$, we have $\frac{n_1}{n} \leq \frac{\alpha}{\alpha\beta-\beta} + \frac{2\alpha\beta-2\alpha-2\beta+1}{2(\alpha\beta-\beta)n}$. Thus,

$$
\rho(s) \geq \frac{\alpha\beta-\alpha-\beta}{\alpha\beta-\beta} - \frac{2\alpha\beta-2\alpha-2\beta+1}{2(\alpha\beta-\beta)n} \geq \frac{\alpha\beta-\alpha-\beta}{\alpha\beta-\beta} - \frac{\alpha\beta-\alpha-\beta}{2(\alpha\beta-\beta)} = \frac{\alpha\beta-\alpha-\beta}{2(\alpha\beta-\beta)}.
$$

Combining the above three cases. Let $n^* = \max\{m_1^*, m_2^*, m_3^*, m_7^*\}$. Then for each payoff-dominant equilibrium (x, g) and each $n > n^*$, we have

$$
\rho(s) \geq \begin{cases} \frac{\alpha\beta-\beta}{2(\alpha\beta+\alpha-\beta)} & \text{if } \alpha\beta - \alpha - \beta \leq 0 \\ \frac{\alpha\beta-\alpha-\beta}{2(\alpha\beta-\beta)} & \text{if } \alpha\beta - \alpha - \beta > 0 \end{cases}.
$$

Generalization of Footnote 15

Here we will show that Proposition 9 holds if we alternatively we define influencers as those with content provision levels higher than a given threshold which can be even increasing in n.

Proof: By Proposition 9, if $\alpha^2\beta+\alpha>1$, then there exists $\rho^*\in(0,1)$ such that, for sufficiently large n, we have $\rho(s)\geq\rho^*$ for each payoff dominant strict equilibrium $s=(x,g)$. Suppose that there are $\bar{t}$ tiers in g and n_t players in each tier t of g. According to Proposition 1, all players in the same tier provide the same level of content in a strict equilibrium. We let $x_{[t]}$ denote the content provision of tier t players. It suffices to show that $\min\{x_{[t]}|x_{[t]}>0\}>bn$ for some $b>0$. There are two cases.

In Case 1, $x_{[1]}=0$. Then $\min\{x_{[t]}|x_{[t]}>0\}=x_{[2]}$. To support (x,g) as a strict equilibrium, some player i in tier 1 has no incentive to delete a link to a player in tier 2, which requires that

$$\begin{aligned}
&u_i(x,g)>u_i(x,g)-\alpha x_{[2]}+\tfrac{1}{2}(n-n_1)^2-\tfrac{1}{2}(n-n_1-1)^2\\
\Leftrightarrow\ &\alpha x_{[2]}>\tfrac{1}{2}(2n-2n_1-1)\\
\Leftrightarrow\ &x_{[2]}>\tfrac{1}{2\alpha}(2n-2n_1-1).
\end{aligned}$$

Then for $n\leq\frac{1}{\rho^*}$, since $n-n_1\geq1$, $\frac{x_{[2]}}{n}>\frac{1}{2\alpha n}(2n-2n_1-1)\geq\frac{1}{2\alpha n}\geq\frac{\rho^*}{2\alpha}>0$. For $n>\frac{1}{\rho^*}$, since $n-n_1\geq\rho^*n$, $\frac{x_{[2]}}{n}>\frac{1}{2\alpha}(2\rho^*-\frac{1}{n})>\frac{1}{2\alpha}(2\rho^*-\rho^*)=\frac{\rho^*}{2\alpha}>0$.

In Case 2, $x_{[1]}>0$. Then by Lemma 4, $\alpha\geq1$. Moreover, $\min\{x_{[t]}|x_{[t]}>0\}=x_{[1]}=\beta(n_1-1)-(n-1)$, and $n_1\geq2$. To support (x,g) as a strict equilibrium, some player i in tier 1 has no incentive to delete a link to a player in tier 1, which requires that

$$\begin{aligned}
&u_i(x,g)>u_i(x,g)-\alpha x_{[1]}+\beta(n_1-1)\\
\Leftrightarrow\ &\alpha x_{[1]}>\beta(n_1-1)\\
\Leftrightarrow\ &n_1-1>\tfrac{\alpha(n-1)}{(\alpha-1)\beta}\\
\Leftrightarrow\ &x_{[1]}>\tfrac{\alpha(n-1)}{\alpha-1}-(n-1)=\tfrac{n-1}{\alpha-1}\\
\Leftrightarrow\ &\frac{x_{[1]}}{n}>\tfrac{1}{\alpha-1}-\tfrac{1}{n}>0
\end{aligned}$$

for sufficiently large n. To conclude, there exists some $b>0$ such that the claim holds.