Communication networks, externalities and the price of information

Arnold Polanski*

Abstract

Information goods (or information for short) play an essential role in modern economies. We consider a setup where information has some idiosyncratic value for each consumer, exerts externalities and can be freely replicated and transmitted in a communication network. Prices paid for information are determined via the (asymmetric) Nash Bargaining Solution with endogenous disagreement points. This decentralized approach leads to unique prices and payoffs in any exogenous network. We use these payoffs to find connection structures that emerge under different externality regimes in pre-trade network formation stage. An application to citation graphs results in eigenvector-like measures of intellectual influence.

*University of East Anglia; A.Polanski@uea.ac.uk

1 Introduction

Information plays an ever more important role in modern economies. The growing information industry (or sector) comprises not only companies that produce information goods (e.g., media products, software) and services (e.g., consulting, education) but also companies that process (e.g., banking, insurance) and disseminate (e.g., telephone, internet) information. Nowadays, information created in this sector is traded predominantly in electronic form and appears in various manifestations, e.g., as music, e-books, patents or (fake) news. Following Shapiro and Varian (1999), we use the term information good (IG or information for short) very broadly. Essentially, anything that can be digitized - encoded as a stream of bits - is information. Muto (1986) identified the following distinctive properties of IGs: free replication, indivisibility, irreversibility of trade and negative external effects.[1] In this work, we assume the first three properties and generalize the last one to external effects of any sign (with no externalities as a special case). We posit further that individual consumption values of the IG and its externalities are known to all players before they acquire information. This premise of complete information extends to all other aspects of the model. Finally, we assume that information diffuses sufficiently fast - potentially, at the speed of light - to all prospective consumers. Then, we can neglect its depreciation and the discounting of (dis)utilities resulting from its consumption. This is a reasonable approximation for, e.g., automated transmission of digitized contents.

Information propagates through transmission channels that form a communication network, e.g., a distribution network for IGs, data transmission infrastructure or a virtual network implied by copyright regulations. Social and business contacts also serve as an ideal vehicle for information exchange. The importance of social networks for information diffusion is exemplified by the huge success of online networking communities such as *Facebook* and *Twitter*. Generally, a link in a communication network represents a channel through which a holder of an IG can transfer a copy of it to a connected agent.

In this article, we analyze the impact of communication networks, externalities and

[1] *Free replication*: Each trader can create identical replicas of the IG at no cost. *Indivisibility*: A possessor benefits from exactly one unit of the IG. *Irreversibility:* buyers cannot return the IG or cancel the trade. *Negative external effects*: Each possessor of the IG is negatively affected if others acquire this good.

valuations on the price of IGs that display the aforementioned properties. Our analysis is based on a (non-strategic) model of bilateral trade in networks. Like similar models, we assume that a seller and a prospective buyer can trade if and only if they are connected. The price paid in a bilateral transaction is calculated via the (asymmetric) Nash Bargaining Solution (e.g., Binmore et al., 1986) with endogenous disagreement points. As natural disagreement values, players in each trading pair use their respective (expected) payoffs from a hypothetical perpetual disagreement. This setup leads to unique prices and payoffs in any exogenous network. We use these payoffs to analyze a network formation stage that precedes information diffusion.

Our analysis yields the following main insights. Firstly, information diffuses to all players who can be reached along a (directed) path in the underlying network from the initial set of sellers. The order in which trades occur and information is transferred has, however, no impact on payoffs and prices of information. Secondly, we devise a recursive algorithm to calculate the unique prices and payoffs for any given network, externalities and initial set of sellers. We characterize the connectivity of nodes that obtain information for free and provide an explicit formula for the payoff to a single seller of information. This formula elucidates the role of externalities exerted along critical paths[2] from this seller to prospective buyers. Thirdly, we use the unique payoffs in fixed networks to find connection structures that emerge under different externality regimes in a pre-trade network formation stage. Finally, in an application to citation networks, we derive eigenvector-like measures (Bonacich and Lloyd, 2001) of intellectual influence.

In order to illustrate the broad spectrum of applications of the model, we consider the following stylized examples (see Figure 1 for their graphical representation). We consider generalizations of these examples in Section 5 in the context of network formation.

a *Positive externalities, tree network* (Figure 1a): A firm can use a medium (television, print, internet, etc.) to advertise its product in order to attract prospective customers. We model this situation as a (directed and rooted) tree with the root (advertiser) that is connected to an internal vertex (e.g., a TV station) that in turn

[2] We define critical paths in Section 3.

is connected to a set of leaf nodes (viewers, prospective customers). Whenever a prospective customer watches an ad, the probability that she will buy the product increases, which we interpret as a positive externality on the advertising firm. Interestingly, the ad itself has no (or has, perhaps, even a negative) intrinsic value for all agents.

b *Negative externalities, star network* (Figure 1b): In a bleak future scenario, a biotech company creates a deadly virus (and the antidote) and then offers its know-how to rival countries. Obviously, such a biological weapon in the arsenal of a country amounts to a threat (and a heavy cost) for its adversaries. A less bellicose example is motivated by the growing importance of markets for information and data brokers. Data brokers (or information re-sellers) collect, process and package data that they then sell to other firms. Accurate information about the business environment and market conditions can be hugely beneficial to a firm giving it an advantage over uninformed competitors. In a simplified form, we model this situation as a star network, where the center (data broker) is connected to a set of spokes (competing firms) and each spoke is harmed by information acquired by another spoke.

c *No externalities, complete network* (Figure 1c): Copyright regulations shape a virtual connection structure by defining property rights for IGs. Assume, for example, that an IG with negligible externalities is sold under the General Public Licence.[3] This licence allows each buyer to freely copy, distribute and modify the IG, provided that all copies and further developments are subject to the same licence. We interpret this "copyleft" covenant as a complete undirected network that connects all current and prospective possessors of the IG.

Figure 1.

The reminder of the article is structured as follows. Section 2 revises the related literature. Section 3 introduces the model, whereas Section 4 presents our results for exogenous communication networks. Based on these results, Section 5 studies network

[3] The terms of the GNU GPL are available at: http://www.gnu.org/copyleft/gpl.html

formation. Section 6 applies the model to citation graphs and Section 7 concludes. Proofs of the main results are relegated to the Appendix.

2 Related Literature

Information networks were first analyzed by graph theorists in the context of gossip and broadcast problems. In a gossip network every individual possess a unique piece of gossip which needs to be communicated to the others. In the broadcast version, one person wishes to communicate information to all others in the network. A survey of the seminal literature on gossip and broadcast networks can be found in Hedetniemi et al. (1988). Recent contributions consider spreading rumors - unverified or unconfirmed statements circulating in a community - as a strategic game (e.g., Bloch et al. 2018; Redlicki, 2015). These contributions belong to the large body of work on social learning. A classic setup in this literature assumes that some agents have private (incomplete) information about the state of the world which influences all players' utilities. Before engaging in a payoff relevant interaction, players decide strategically either to whom or/and how much of their private information to reveal (e.g., Gal-Or, 1985; Koessler and Renault, 2012), or they invest in pairwise communication channels (e.g., Calvó-Armengol et al., 2015) or simultaneously send cheap talk messages to each other (e.g., Hagenbach and Koessler, 2010; Galeotti et al., 2013). Alternatively, some authors consider non-strategic information transmission in the context of naive (e.g., Golub and Jackson, 2010) or Bayesian (e.g., Mueller-Frank, 2013) learning. In any case, the focus of these models is on the ability and efficiency of a population to aggregate information in a fixed or in an endogenously formed communication network. There is also a substantial work on the role of local strategic interactions in the diffusion of conventions, standards and inventions in socioeconomic networks (e.g., Chwe, 2000; Ellison, 1993; Bhaskar and Vega-Redondo, 2004).

Our work relates to the large literature on bilateral trade in networks (see Manea, 2016, for a recent survey). Two agents can trade with each other in these networks if and only if they are connected. In our framework, each holder of an IG can sell it to all adjacent buyers. However, unlike most models in this literature, a transaction does not entail the departure of the trading pair from the network and the deletion of all adjacent links. Instead, it ushers a new configuration with more sellers and fewer buyers. The present model complements the relevant literature by analyzing how the special properties of information goods (in particular, their externalities) interplay with the connection structure.

Most of the aforementioned work has focused on understanding the impact of an exogenously given (bipartite) network on the outcome of trade. Some authors study endogenous network formation. For example, Wang and Watts (2006) examine the formation of trade networks in quality-differentiated product markets, whereas Elliott (2015) proposes a two-stage game: First, players simultaneously decide their linking (relationship-specific) investments and, then, bargain and trade over the created network. Galeotti and Goyal (2010) develop a model where players decide simultaneously to acquire information and form connections with others to access their information. In this work, we will allow agents (social planner) to form (design) communication structures in a network formation stage that precedes information diffusion. We focus then on pairwise stable, optimal and (socially) efficient structures that emerge under different externality regimes.

Technological advances and the explosion of e-commerce have inspired a rich theoretical and empirical work on pricing policies for information goods. Varian (2000) is a general introduction to information goods and their pricing, whereas Linde (2009) is an example of new forms of price discrimination made possible by special economic features of information. Increasingly, firms turn to social networks to diffuse their products relying on word-of-mouth communication for advertising and exploiting consumption externalities among consumers. The advent of the internet has vastly increased the ease and scope of viral marketing. In this context, the question "who influences whom" is of fundamental interest. The development of new methods to identify influential and susceptible consumers from large data sets is an active research area in the intersection between business/marketing and information systems. Probst et al. (2013) is a survey of this literature, whereas Bloch (2016) revises recent theoretical work on targeting and pricing in social networks. Our work differs significantly from the main strand of this literature by explicitly assuming that a buyer of information will resell it to all other connected buyers.

Closest in spirit to this article are the works by Galbreth et al. (2012), Muto (1986), Ali et al. (2016), Polanski (2007) and Manea (2017). Galbreth et al. (2012) examine the effect of social sharing on the price of information goods under different network structures. Muto (1986) considers two types of markets for information goods: Markets with free resales and markets where resales are prohibited (i.e., complete networks and stars, respectively). He models the trading process as a multilateral bargaining in which each possessor of information offers simultaneously a price to every demander who can accept or refuse the trade. Ali et al. (2016) study a game with discounting, where sellers and prospective buyers bargain and trade information bilaterally, and the buyers may resell it. For the network structures they consider (complete networks and stars) their results coincide qualitatively with those obtained in the present work.

In Polanski (2007) and in Manea (2017), prices of a homogenous IG are calculated in a sequence of bilateral meetings of agents connected in the underlying network. Both authors use the (asymmetric) Nash Bargaining Solution to determine the terms of trade for the matched buyer-seller pairs in each configuration (state). Disagreement points for these pairs are set to their respective (expected) payoffs in the same configuration as the latter persists if they fail to agree. However, both authors impose a different rule when trade is possible in one link only. In this case, the disagreement points are set to zero. The present framework enriches both models by allowing network formation and idiosyncratic externalities of information. Unlike Polanski (2007) and Manea (2017), we assume immediate agreement in all matched links and, more importantly, we prescribe a uniform rule for all disagreement points: They are uniquely defined by the respective payoffs that the matched pair would obtain in case of (hypothetical) permanent disagreement in the current configuration. For fixed networks and negligible externalities, all these models produce identical results when evaluated with the same values for sellers' bargaining power and agents' valuations.

3 The model

■ **Defintions and notation.** We consider a graph (network) $\mathcal{G} \equiv \{\mathcal{N}, \mathcal{L}\}$ with the finite set of nodes (vertices) $\mathcal{N} \equiv \{1, ..., n\}$ and the set of directed edges $\mathcal{L} \subseteq \{vw : v, w \in \mathcal{N}\}$.

If $vw \in \mathcal{L}$ and $wv \in \mathcal{L}$ there exists an undirected edge $\overline{vw}$ between v and w. A (directed) *path* from a vertex v_1 to a vertex v_k in $\mathcal{G}$ is a sequence of nodes $(v_1, v_2, ..., v_k)$, $v_i \in \mathcal{N}$, such that successive vertices v_i and v_{i+1} are endpoints of the intermediate edge $v_i v_{i+1} \in \mathcal{L}$. In a *connected component* a path exists between any two nodes. A *cycle* is a path $(v_i, ..., v_k, ..., v_i)$ in which all but the first and the last node are pairwise different. An *acyclic graph* has no cycles. In the *complete graph*, the shortest path from vertex v to another node w is (v, w). A vertex $v \in \mathcal{N}$ can be reached (is *accessible)* in $\mathcal{G}$ from a subset $\mathcal{V} \subseteq \mathcal{N}$ if a path from a node $k \in \mathcal{V}$ to v exists. Only in this case, information can flow from $\mathcal{V}$ to v. Furthermore, we say that a node $v \in \mathcal{N}$ is *connected* (*adjacent*) to another vertex $w \in \mathcal{N}$ if $vw \in \mathcal{L}$. The set of all vertices adjacent to the node v will be denoted $N_v(\mathcal{G}) \equiv \{w \in \mathcal{N} : vw \in \mathcal{L}\}$ and occasionally referred to as v's neighbors.

A crucial role in our analysis is played by critical paths. Critical path $\delta^{v \to w} \equiv \delta^{v \to w}(\mathcal{G})$ in the network $\mathcal{G}$ is the longest path (i.e., a sequence of nodes) that starts at v and is shared by all paths from v to w. For example, in undirected trees $\delta^{v \to w} = (v, ..., w)$ is the unique path from v *to* w, whereas in the complete network $\delta^{v \to w} = (v)$. When no path from v to w exists or $w = v$, we set $\delta^{v \to w} \equiv (v)$. Figure 2 and Figure 3 illustrate critical paths in directed and undirected networks, respectively.

Figure 2.

We denote by $\mathcal{S} \subseteq \mathcal{N}$ the subset of holders (possessors, sellers) of a perfect copy of an underlying IG. When there is no risk of confusion, we will use the shorthand $\mathcal{C}$ for $(\mathcal{G}, \mathcal{S})$ and refer to $\mathcal{C}$ as *configuration.* We define further $\mathcal{B} \equiv \mathcal{B}(\mathcal{C}) \subseteq \mathcal{N}$ as the subset of prospective buyers of the IG that can be reached in $\mathcal{G}$ from the set $\mathcal{S}$. Hence, $\mathcal{B} \cap \mathcal{S} = \emptyset$, $\mathcal{B} \cup \mathcal{S} \subseteq \mathcal{N}$ and for each $b \in \mathcal{B}$ there is some $s \in \mathcal{S}$ such that a path from s to b exists in $\mathcal{G}$. Note that $\mathcal{B}(\mathcal{C}) = \emptyset$ when $\mathcal{S} = \emptyset$ or $\mathcal{S} = \mathcal{N}$. If a subset $\mathcal{X} \subseteq \mathcal{B}$ of buyers acquires information in configuration $\mathcal{C}$, the ensuing configuration will be denoted by $\mathcal{C} \oplus \mathcal{X} \equiv (\mathcal{G}, \mathcal{S} \cup \mathcal{X})$. When $\mathcal{X} = \{x\}$, we simply write $\mathcal{C} \oplus x$, $\mathcal{S} \cup x$, etc. Ordered seller-buyer pairs $sb \in (\mathcal{S} \times \mathcal{B}) \cap \mathcal{L}$ that are connected in $\mathcal{L}$ form the set $L(\mathcal{C})$ of *active pairs (links).* Each active pair can trade information. Note that $L(\mathcal{C}) = \emptyset$ if and only if $\mathcal{B}(\mathcal{C}) = \emptyset$. An *active player* is a node covered by at least one active link. Moreover, we write $\#\mathcal{X}$ for the number of elements in the set or sequence $\mathcal{X}$ and define the indicator function $\mathbf{I}_c \equiv 1$ (0) when the condition c is true (false).

Information acquired by a player has a (possibly negative) consumption value for her but it also may impose externalities on other players. We collect the valuations and externalities in the matrix $W = (\varpi_{bk})_{b,k\in\mathcal{N}}$. Each $\varpi_{bk} \in R$ (with the special case of b's intrinsic value or valuation ϖ_{bb}) stands for the one-off (dis)utility that player $k \in \mathcal{N}$ realizes when player $b \in \mathcal{B}$ consumes the IG (learns the relevant piece of information). As we neglect depreciation and discounting, this utility will not depend on the date of information acquisition by agent b. In addition to the examples $a-b$ in the Introduction, this type of externalities result, e.g., from revelation of confidential data on platforms such as *Wikileaks* or in telecommunication networks, when both the caller and the receiver benefit from a call (Jeon et al., 2004).

It is important to stress that not all forms of externalities are captured by the approach embodied in the matrix W. For example, information may exhibit network (external) effects (e.g. Shapiro and Varian, 1999; Crémer, 2000). In this case, the value of IGs (or other goods) is affected by the number of agents possessing the same good. Typical example are the adoption of new information technologies, standards or social behavior. Generally, externalities could depend on the set of players that possess the relevant piece of information.[4] In our framework, the intrinsic value of an IG is independent of any other holders of this good. However, when each possessor exerts externalities, their increasing number will have a cumulative impact on agents' total payoffs.

■ **Pairwise matching and bilateral trading.** In a configuration $\mathcal{C}$ with $L(\mathcal{C}) \neq \emptyset$ at most one active pair meets at each (discrete) date $t = 1, 2, ...$ In the matched pair $sb \in L(\mathcal{C})$, the seller $s \in \mathcal{S}$ transfers the IG to the buyer $b \in \mathcal{B}$, who pays the price[5] $p_{sb}(\mathcal{C})$, each player $k \in \mathcal{N}$ obtains the (dis)utility ϖ_{bk} and the configuration $\mathcal{C} \oplus b$ with the set of sellers $\mathcal{S} \cup b$ (and the set of buyers $\mathcal{B} \backslash b$) ensues in $t+1$. Crucially, we allow each information holder to produce an arbitrary number of perfect copies of it at zero cost that can be transferred to any neighbour. Information diffusion ends when a configuration $\mathcal{C}$ with $L(\mathcal{C}) = \mathcal{B}(\mathcal{C}) = \emptyset$

[4] Fainmesser and Galeotti (2015) distinguish between global network effects (when an individual is affected by the consumption of the entire population) and local network effects (when an individual is affected by the consumption of a subset of the population).

[5] In principle, this price could depend on past events and the date t. As this is never the case in our model, we simply write $p_{sb}(\mathcal{C})$.

is reached.

We need not make any assumptions on how a particular pair is selected other than supposing that every link $sb \in L(\mathcal{C})$ is chosen with some probability that is bounded away from zero.[6] This requirement implies that information diffuses to all accessible buyers with probability one and ensures that the connection structure is preserved during the diffusion process. If an active link never traded the resulting outcome would reflect the connection pattern of a reduced graph.

The price $p_{sb}(\mathcal{C})$ paid in the active link sb for the IG is determined via the asymmetric Nash Bargaining Solution (NBS) with exogenous bargaining powers and endogenous disagreement points.[7] We collect the bargaining powers in the matrix $T = (\theta_{ik})_{i,k \in \mathcal{N}}$, where $\theta_{ik} \in (0, 1)$ is the share of the net surplus, created when i sells the IG to k, that accrues to seller i. The remainder $1 - \theta_{ik}$ of this share goes to the buyer k. Note that this specification allows for role-dependent bargaining powers. For example, player's i bargaining position vis-à-vis player k may become stronger when she sells information ($\theta_{ik} > \theta_{ki}$). We say that i and k have symmetric bargaining positions when $\theta_{ik} = \theta_{ki} = 1/2$. For a pair $sb \in \mathcal{S} \times \mathcal{B}$ and $k \in sb$, we define further $\theta_k^{sb} \equiv \theta_{sb}$ when $k = s$ and $\theta_k^{sb} \equiv 1 - \theta_{sb}$ when $k = b$.

Regarding disagreement points, we make the following assumptions. If sb is the only active pair in the configuration $\mathcal{C}$, i.e., $L(\mathcal{C}) = \{sb\}$, their hypothetical perpetual disagreement would lead to zero payoffs for all players as information would not spread to any prospective buyer. Hence, we set disagreement values to zero for both s and b. On the other hand, if $sb \in L(\mathcal{C})$ decided to perpetually disagree in configuration $\mathcal{C}$, when $L(\mathcal{C})$ contains more than one active pair, then some other link would eventually meet and trade (remember that each active pair trades with positive probability). Thus, we assume that the disagreement values for s and b are the respective payoffs resulting from a trade elsewhere. As we show below, the latter payoffs do not depend on the trading link. Importantly, disagreement considered by an active link is hypothetical and treated as such by all players given the stipulation that information flows through all active links with positive probability.

[6] In order to minimize inessential notation, we do not formally define a matching protocol.

[7] Asymmetric NBS (e.g., Binmore et al., 1986) for the bilateral bargaining problem with the disagreement points d_1 and d_2, bargaining powers $\theta_1 \in [0, 1]$ and $\theta_2 = 1 - \theta_1$ and the surplus S is the solution $x_1^* = d_1 + \theta_1(S - d_1 - d_2)$ and $x_2^* = d_2 + \theta_2(S - d_1 - d_2)$ to $\max_{x_1, x_2}(x_1 - d_1)^{\theta_1}(x_2 - d_2)^{\theta_2}$, *s.t.* $x_1 + x_2 = S$.

This protocol ignores the possibility of strategic no-trade matchings. This simplification allows for sharp predictions and straightforward applications of our trading model. Although innocuous in many situations, one can easily construct examples (see Section 5) where a matched pair may prefer to disagree on the terms of trade, i.e., to dissolve the match without the transfer of information. A rational (dis)agreement, however, will often depend on (dis)agreement decisions in other links, which leads to a complex strategic interaction. In Section 5, we study a network formation stage that precedes information trading.

■ **Agreement payoff and transferable surplus.** For a fixed configuration $\mathcal{C}$, we will denote by $x_k^{sb}(\mathcal{C})$ the payoff to player $k \in \mathcal{N}$, when trade takes place in the link $sb \in L(\mathcal{C})$. Similarly, $x_k(\mathcal{C})$ is the expected or ex-ante payoff in configuration $\mathcal{C}$. As we show below, $x_k^{sb}(\mathcal{C})$ and $x_k(\mathcal{C})$ are unique for each $\mathcal{C}$. In particular, $x_k(\mathcal{C}) = 0$ when $\mathcal{C}$ does not allow for information diffusion, i.e., when $L(\mathcal{C}) = \mathcal{B}(\mathcal{C}) = \emptyset$. When $\mathcal{C}$ admits trading in the link $sb \in L(\mathcal{C})$, we compute the payoff $x_k^{sb}(\mathcal{C})$ as,

$$x_k^{sb}(\mathcal{C}) = \varpi_{bk} + x_k(\mathcal{C} \oplus b) + (\mathbf{I}_{k=s} - \mathbf{I}_{k=b})p_{sb}(\mathcal{C}), \quad \forall k \in \mathcal{N}, \tag{1}$$

which simply states that each player k obtains the (dis)utility ϖ_{bk} and expects the continuation payoff $x_k(\mathcal{C} \oplus b)$ in configuration $\mathcal{C} \oplus b$ that ensues after player b has acquired information (note that the continuation payoffs are not discounted). Moreover, if $k = s$ $(k = b)$, then k also receives (pays) the price $p_{sb}(\mathcal{C})$ for the IG. This price is determined by (1) and the NBS sharing rule,

$$x_k^{sb}(\mathcal{C}) = d_k^{sb}(\mathcal{C}) + \theta_k^{sb}(S_{sb}(\mathcal{C}) - d_s^{sb}(\mathcal{C}) - d_b^{sb}(\mathcal{C})), \tag{2}$$

where $d_k^{sb}(\mathcal{C})$ is player's $k \in sb$ (endogenous) disagreement point in the link sb and the transferable surplus $S_{sb}(\mathcal{C})$ is the sum of payoffs (1) for s and b following their agreement in $\mathcal{C}$,

$$S_{sb}(\mathcal{C}) \equiv x_s^{sb}(\mathcal{C}) + x_b^{sb}(\mathcal{C}) = \varpi_{bs} + x_s(\mathcal{C} \oplus b) + \varpi_{bb} + x_b(C \oplus b). \tag{3}$$

Hence, the surplus $S_{sb}(\mathcal{C})$ consists of b's valuation ϖ_{bb}, this player's externality ϖ_{bs} exerted on s and the continuation payoffs for both players in the ensuing configuration $\mathcal{C} \oplus b$.

As advanced earlier, for the disagreement value of player $k \in sb \in L(\mathcal{C})$, we consider two cases: $d_k^{sb} = 0$ when $\{sb\} = L(\mathcal{C})$ and $d_k^{sb} = x_k^{s'b'}(\mathcal{C})$ for some active link $s'b' \neq sb$ when $\{sb, s'b'\} \subseteq L(\mathcal{C})$. The former case reflects the fact that no further (dis)utility would be created if s and b perpetually disagreed in a configuration with the single active link sb. The latter case prescribes for the link sb the disagreement payoffs that result from a trade in another link $s'b' \neq sb$. As we show in the next section, the payoff (2) does not depend on the choice of $s'b'$.

4 Exogenous networks

■ **Order independence and recursive payoff computation.** Our liberal assumptions on the matching procedure may suggest a multiplicity of (expected) payoffs. Fortunately, our first result dispels this possibility. This result, which we refer to as *order independence,* shows that the payoff to a player does not depend on the link that agrees and trades. Hence, the paths of information diffusion are irrelevant for the payoffs.

Proposition 1. *(order independence) There are unique payoffs $\{x_k(\mathcal{C})\}_{k \in \mathcal{N}}$ for any configuration $\mathcal{C}$. These payoffs neither depend on the (non-vanishing) matching probabilities nor on the trading pair:*

$$x_k(\mathcal{C}) = x_k^{sb}(\mathcal{C}), \quad \forall k \in \mathcal{N}, \quad \forall sb \in L(\mathcal{C}) \neq \emptyset.$$

Order independence implies that none of the players loses valuable trade opportunities when transactions occur in non-adjacent links. In particular, buyers are not harmed by transactions elsewhere as the latter can only improve their bargaining position, whereas delays to trade are inconsequential for valuations and externalities. Then, each buyer b acquires the IG at a price that she would obtain after all trades not involving her had taken place. In the corresponding configuration $\mathcal{C}$, there is either only one active link $sb \in L(\mathcal{C})$ and the NBS (2) yields $x_k^{sb}(\mathcal{C}) = \theta_k^{sb} S_{sb}(\mathcal{C})$ for $k \in \{s, b\}$ or the competition among sellers drives the price to zero and (1) boils down to $x_k^{sb}(\mathcal{C}) = \varpi_{bk} + x_k(\mathcal{C} \oplus b)$ for any active link $sb \in L(\mathcal{C})$. Therefore, the ex-ante payoff to any player $k \in \mathcal{N}$ in configuration $\mathcal{C}$ is computed recursively by (1),

$$x_k(\mathcal{C}) = \varpi_{bk} + x_k(\mathcal{C} \oplus b) + (\mathbf{I}_{k=s} - \mathbf{I}_{k=b})p_{sb}(\mathcal{C}), \quad \forall k \in \mathcal{N}, \forall sb \in L(\mathcal{C}). \tag{4}$$

with the price $p_{sb}(\mathcal{C})$ determined in our next result.

Proposition 2. *In configuration $\mathcal{C}$ with at least one active buyer, the price paid in the link $sb \in L(\mathcal{C})$ for information verifies,*

$$p_{sb}(\mathcal{C}) = \begin{cases} \varpi_{bb} + x_b(\mathcal{C} \oplus b) - \varpi_{b'b} - x_b(\mathcal{C} \oplus b'), & \forall s'b' \in L(\mathcal{C}),\ s'b' \neq sb, \\ \theta_{sb} \cdot S_{sb}(\mathcal{C}) - \varpi_{bs} - x_s(\mathcal{C} \oplus b), & if\ L(\mathcal{C}) = \{sb\}, \end{cases} \tag{5}$$

where $S_{sb}(\mathcal{C})$ is defined in (3). Moreover, $p_{sb}(\mathcal{C}) = p_{sb}(\mathcal{C} \oplus b')$ for all $s'b' \in L(\mathcal{C})$ with $b \neq b'$.

We can apply (4)-(5) recursively to obtain unique payoffs to all players in configuration $\mathcal{C}$. Specifically, if $L(\mathcal{C}) \neq \emptyset$, we first find all active links. If there are more than one, we select any two of them, say sb and $s'b'$, and compute $p_{sb}(\mathcal{C})$ by the first line in (5). Otherwise, the price in the single active link is given by the second line in (5). The computed price is, then, substituted into (4). By order independence, the resulting payoffs will not depend on the selected links. This algorithm expresses each payoff in $\mathcal{C}$ as a linear combination of externalities and payoffs in configurations with one more seller. Note that the recursion is closed as $x_k(\mathcal{C}) = 0$ when $\mathcal{B}(\mathcal{C}) = \emptyset$, i.e., when no active buyers (and links) remain. Although easily implemented as a computer program, the formulae (4)-(5) offer no direct insights into the interplay of the network, valuations and externalities in the determination of information prices and payoffs. Our next results go some way in this direction. First, we need the following definition.

Definition 1. *Given a configuration $\mathcal{C} = (\mathcal{G}, \mathcal{S})$, we say that a buyer $b \in \mathcal{B}(\mathcal{C})$ satisfies the two paths property (with respect to $\mathcal{S}$) when it has at least two neighbors $v, w \in \mathcal{N}$, $b \in N_v(\mathcal{G}) \cap N_w(\mathcal{G})$, and each of them either belongs to the set $\mathcal{S}$ or it can be reached from $\mathcal{S}$ by a path that does not include b.*

The two paths property is illustrated in Figure 3 and its implication is presented in the next proposition.

Figure 3.

Proposition 3. *Each buyer that satisfies the two paths property obtains the IG for free.*

When buyer b has two seller neighbours in configuration $\mathcal{C}$, say s and s', and $sb, s'b \in L(\mathcal{C})$, this result follows directly from (5) in Proposition 2. Otherwise, the two paths property and order independence ensure that such a configuration can be reached from the initial set of sellers without affecting b's payoffs. Intuitively, the absence of opportunity costs (no discounting) gives a strong bargaining position to buyers satisfying the two paths property. These buyers can wait until at least two of their neighbors become sellers and compete with each other driving the price to zero. For example, in the network in Figure 3, player 2 satisfies the two paths property when node 3 is the single IG seller. If this player waits for the buyer 4 to acquire the IG from 3, then she will face two sellers of a perfect substitute. On the other hand, player 2 does not satisfy the two paths property when vertex 1 is the single seller. In this case, player 2 has no choice but to split up the surplus with this seller.

A particularly transparent situation arises when all buyers satisfy the two paths property as is the case of the complete network. Then, all payoffs are computed by summing up the columns of the matrix W.

Corollary 1. *If all buyers in the configuration $\mathcal{C}$ satisfy the two paths property, then,*

$$x_k(\mathcal{C}) = \textstyle\sum_{b\in\mathcal{B}(\mathcal{C})} \varpi_{bk}, \quad \forall k \in \mathcal{N}.$$

The proof of this corollary follows directly by the recursive expansion of (4) with all prices equal to zero. Interestingly, the bargaining powers are irrelevant for the payoffs in this case.[8]

The next, somehow polar, scenario contemplates arbitrary undirected networks and "local externalities". The latter refer to situations, where only consumer b and her linked

[8] As a consequence of Proposition 1, in configurations with at least two active links, each trading player obtains her disagreement value. In these configurations, the bargaining power is irrelevant for price and payoff computation.

acquaintances are affected by b's information consumption. Think, for instance, of posts on *Facebook* that intend to impress friends. Formally, W displays local externalities when,

$$\varpi_{bk} \neq 0 \Rightarrow \left\{ b = k \quad or \quad \overline{kb} \in \mathcal{L} \right\}. \tag{6}$$

The payoffs in this scenario highlight the role of bridges or cut-edges for information pricing. We say that a link $\overline{vw} \in \mathcal{L}$ is a bridge (cut-edge) in an undirected graph $\mathcal{G}$ when its deletion increases the number of connected components. The concept of critical path conveys this property succinctly: $\overline{vw} \in \mathcal{L}$ is a cut-edge if and only if $\delta^{v \to w} = (v, w)$ (or, equivalently, $\delta^{w \to v} = (w, v)$). Bridges are the only edges across which non-zero prices are paid for information.

Corollary 2. *For matrix W that satisfies (6) and undirected graph $\mathcal{G}$, the total payoff to the single seller k in configuration $\mathcal{C} = (\mathcal{G}, \{k\})$ is computed by the recursive formula,*

$$x_k(\mathcal{C}) = \sum_{\overline{kb} \in \mathcal{L} : \delta^{k \to b} = (k)} \varpi_{bk} + \sum_{\overline{kb} \in \mathcal{L} : \delta^{k \to b} = (k,b)} \theta_{kb} \left\{ \varpi_{bb} + \varpi_{bk} + x_b(\mathcal{C} \oplus b) \right\}, \quad \forall k \in \mathcal{N}. \tag{7}$$

To show (7), we note first that any neighbor b of k such that $\delta^{k \to b} = (k)$ exerts the externality ϖ_{bk} but obtains the IG for free as b satisfies the two paths property. For k's neighbor b such that $\delta^{k \to b} = (k, b)$, we invoke order independence and assume that all active links but $\overline{kb}$ have traded. Then, the surplus to divide between k and b is equal to ϖ_{bb}+ ϖ_{bk}+ $x_b(\mathcal{C} \oplus b)$ because after buying the IG, b will be the only seller in the connected component of $\mathcal{G}$ obtained by cutting the bridge $\overline{kb}$ (k will be inactive after this transaction). In this case, k receives the share θ_{kb} of this surplus by the NBS.

The last formula shows, in particular, that only players connected by cut-edges will be able to extract a positive surplus from their neighbors either as information creators or intermediaries. In the latter case, they will also have to pay a non-zero price for information that arrives through a cut-edge. In the next subsection, we generalize the case of a single seller to any externality structure and directed networks.

■ **The case of a single seller.** Probably the most interesting case for practical applications is the situation of a single creator of an IG, who wants to sell it to a network of prospective buyers. This case is of particular importance for the evaluation of incentives to create IGs. The next result shows that the profit (or loss) of a single seller of an IG is intimately related to externalities exerted along the critical paths from this seller to the accessible buyers.

Proposition 4. *The payoff to the single seller s in configuration $\mathcal{C} = (\mathcal{G}, \{s\})$ is given by,*

$$x_s(\mathcal{C}) = \sum_{b \in \mathcal{B}(\mathcal{C})} \left\{ \sum_{k \in \delta^{s \to b}} \varpi_{bk} \cdot \theta^{\min\{\#\delta^{s \to k}, \#\delta^{s \to b} - 1\}} \right\}, \tag{8}$$

when $\theta = \theta_{ik} \in (0,1)$ for all $i, k \in \mathcal{N}$ and $x_s(\mathcal{C}) = 0$ if $\mathcal{B}(\mathcal{C}) = \emptyset$.

The formula (8) can be interpreted as follows: For each accessible buyer $b \in \mathcal{B}(\mathcal{C})$, the single seller $s \in S$ extracts a share of the externalities that this buyer exerts on each node $k \in \mathcal{N}$ along the critical path $\delta^{s \to b}$ (including b's "externality" ϖ_{bb} on itself if $b \in \delta^{s \to b}$). Seller's share falls geometrically in the length of the critical path to k, as each ϖ_{bk} is weighted by (essentially) $\theta^{\#\delta^{s \to k}}$. For example, an unit increase in ϖ_{bk} changes seller's payoff only when k lies on the critical path from s to b. In this case, seller's payoff increases by $\theta^{\#\delta^{s \to k}}$ when $k \neq b$ and by $\theta^{\#\delta^{s \to k} - 1}$ when $k = b$.

We apply (8) to single-seller configurations $\mathcal{C} = (\mathcal{G}, \{s\})$, where $\mathcal{G}$ is the undirected network depicted in Figure 3. As the critical path from $s = 1$ to any other node b in $\mathcal{G}$ is $\delta^{s \to b} = (1, 2)$, the formula (8) boils down to,

$$x_1(\mathcal{C}) = \theta \textstyle\sum_{b=2}^{4} (\varpi_{b1} + \varpi_{b2}).$$

Similarly, for $s = 2$, the critical paths $\delta^{s \to 1} = (2, 1)$ and $\delta^{s \to b} = (2)$ for $b = 3, 4$ imply,

$$x_2(\mathcal{C}) = \theta(\varpi_{12} + \varpi_{11}) + \textstyle\sum_{b=3}^{4} \varpi_{b2}.$$

Finally, for $s \in \{3, 4\}$, the critical path to any other vertex $b \neq s$ is $\delta^{s \to b} = (s)$ and,

$$x_s(\mathcal{C}) = \textstyle\sum_{b \neq s} \varpi_{bs}.$$

5 Endogenous networks

Our crucial assumption (see Section 3) that each active link trades information when matched may lead to irrational transactions. In the simplest network with a single seller connected to a single buyer, a trade is not compatible with rational behavior if the negative externality, imposed on the seller by the informed buyer, exceeds the utility of the latter player from information acquisition. In this case, NBS implies negative payoffs for both players. Obviously, these players could be better off by never trading but our model prescribes a transaction for each matched pair. Then, the only way to resolve this conflict is to permanently delete the common link.

In this section, we will allow agents (social planner) to form (design) communication structures in a network formation stage that precedes information diffusion. Specifically, we shall focus on pairwise stable, optimal and efficient networks under different externality regimes.

■ **Externality regimes.** First, we generalize our motivating examples in the Introduction to three different externality scenarios. In each scenario, only player 1 will initially possess the relevant IG.

A) *Positive externalities*: Building on the advertisement example *a*) in Section 1, we assume that each player (viewer) $i \in \mathcal{N}\backslash\{1\}$ experiences the disutility ϖ_{ii} from acquiring information (watching the ad) and exerts the positive externality ϖ_{i1} on the IG seller (advertising firm). The positive externality of the ad outweighs the disutility from being exposed to it. Formally,

$$\varpi_{ii} \leq 0, \quad \varpi_{i1} > 0, \quad \varpi_{ik} = 0, \quad \varpi_{i1} + \varpi_{ii} > 0, \quad \forall i, k \in \mathcal{N}\backslash\{1\}, i \neq k. \tag{9}$$

B) *Negative externalities*: Generalizing the example *b*) of market for information in Section 1, we assume that each firm $i \in \mathcal{N}\backslash\{1\}$ earns the profit ϖ_{ii} when acquiring information from the data broker (player 1) and exerts the negative externality ϖ_{ik} on any other firm $k \neq i$. We assume further that externalities are sufficiently strong, i.e. they exceed any intrinsic value,

$$\varpi_{ii} > 0, \quad \varpi_{i1} = 0, \quad \varpi_{ik} < 0, \quad |\varpi_{ik}| > \max_{v} \varpi_{vv}, \quad \forall i, k \in \mathcal{N} \backslash \{1\}, i \neq k. \qquad (10)$$

C) *No externalities*: In the n-player version of the example c), each player $i \in \mathcal{N}$ has a positive valuation ϖ_{ii} for the IG and there are no externalities,

$$\varpi_{ii} > 0, \quad \varpi_{ik} = 0, \quad \forall i, k \in \mathcal{N} \backslash \{1\}, i \neq k. \qquad (11)$$

Our aim is to find (pairwise) stable, optimal and efficient connection structures in each of these scenarios. First, we need to specify what benefits and costs agents anticipate from their alternative linking decisions. Benefits, on the one hand, are quite naturally identified with the payoffs that an agent foresees in the information diffusion stage in a formed network. Linking costs, on the other hand, are assumed to be positive but infinitesimally small: every agent bears a cost of each of her links but this cost is orders of magnitude lower than any payoff she receives in a formed network. This assumption on connection costs is intuitively appealing and helps eliminate "superfluous" links. We assume further undirected networks (i.e., two-way information flow), symmetric bargaining powers $\theta_{ik} = 1/2$ for all $i, k \in \mathcal{N} = \{1, ..., n\}$, and at least three players ($n \geq 3$).

■ **Pairwise stability.** There are many approaches to modeling decentralized network formation. An obvious one is simply to model it explicitly as a non-cooperative game. Alternatively, one may dispense with the specifics of a noncooperative game and define a notion of a stable network directly. Jackson et al. (2005) and Mauleon and Vannetelbosch (2016) are excellent surveys of both approaches. Here, we apply pairwise stability which is probably the most popular network stability concept. Intuitively, a network is pairwise stable if no player benefits from severing one of their links and no two players benefit from adding a link between them, with one benefiting strictly and the other at least weakly. Pairwise stability is simple and tractable but it is a weak concept that does not eliminate many implausible networks. Moreover, pairwise stable networks do not always exist.

The next proposition reports, for each scenario, pairwise stable networks and the corresponding payoff $x_1(\mathcal{C})$ to the initial seller. This payoff omits the (negligible) linking costs.

Proposition 5. *A pairwise stable network $\mathcal{G}$ in the configuration $\mathcal{C} = (\mathcal{G}, \{1\})$ is connected and in scenario*

A) $\mathcal{G}$ is the seller-centered star; $x_1(\mathcal{C}) = \sum_{b=2}^{n} \varpi_{b1}/2$.

B) $\mathcal{G}$ is a collection a cycles, any two of which sharing at most one node; $x_1(\mathcal{C}) = 0$.

C) $\mathcal{G}$ is a line where the seller has a single link; $x_1(\mathcal{C}) = \sum_{b=2}^{n} \varpi_{bb}/2^{b-1}$.

In scenario A, the advertising firm (center of the star) pays $\varpi_{i1}/2 > 0$ to each viewer (spoke) $i \in \mathcal{N}\backslash\{1\}$. We can think of this arrangement as direct marketing, where the advertiser communicates directly with potential customers (e.g., via text messages) offering them promotional codes. In scenario B, all pairwise stable structures, with the cycle $(1, 2, ..., n, 1)$ as the simplest example, are Eulerian graphs, i.e. have a cycle that goes through all edges exactly once. Then, each firm satisfies the two paths property and obtains the information for free. It follows that the seller does not earn any revenue, whereas all firms suffer the full extent of externalities. Finally, in scenario C, each (re)seller sells the IG to one buyer only. Hence, all (re)sellers achieve positive prices but the price for the initial seller is below the revenue earned in the star. For all scenarios, pairwise stable networks are connected, i.e. information diffuses to all nodes, and the payoff to the initial seller is well-defined, i.e. unique.

Interestingly, considering the examples in Section 1 as special cases of the three scenarios, we observe that none of the networks depicted in Figure 1 is pairwise stable. In the tree in example a), each viewer would delete her link to the TV station to save the linking cost without decreasing her payoff. In example b), any two spokes would form a link in order to get the IG for free, whereas in example c) the initial seller would delete one of her links to save the linking costs without reducing her (zero) profits.

■ **Optimality.** An important question for a single seller of an IG is which connection structure - representing, e.g., patent legislation - maximizes the profit from selling this IG. Formally, we say that the network $\mathcal{G}$ is optimal, i.e. profit maximizing for the single seller s, given the set of prospective buyers $\mathcal{N}\backslash\{s\}$, if $x_s((\mathcal{G}, \{s\})) \geq x_s((\mathcal{G}', \{s\}))$ for all networks

$\mathcal{G}'$ with the set of nodes $\mathcal{N}$. It is clear that optimal networks always exist given that the set of relevant networks is finite. The following proposition reports the optimal network(s) for each scenario and the corresponding payoff $x_1(\mathcal{C})$ to the initial seller. This payoff is net of the (negligible) linking costs.

Proposition 6. *An optimal network $\mathcal{G}$ in the configuration $\mathcal{C} = (\mathcal{G}, \{1\})$ is connected and in scenario*

A) the seller is covered by a cycle and has no other links; $x_1(\mathcal{C}) = \sum_{b=2}^{n} \varpi_{b1}$.

B) $\mathcal{G}$ is the seller-centered star; $x_1(\mathcal{C}) = \sum_{b=2}^{n} \varpi_{bb}/2$.

C) $\mathcal{G}$ is the seller-centered star; $x_1(\mathcal{C}) = \sum_{b=2}^{n} \varpi_{bb}/2$.

In scenarios B and C with positive consumption values but non-positive externalities, the seller-centered star is the unique optimal structure. In the context of property rights, we can think of star networks as a rule that declares illegal any resales of IGs. As a diametrically opposed provision, the General Public Licence can be interpreted as the complete (undirected) network that connects all current and prospective possessors of the IG. The single seller 1 prefers the GPL (complete network $\mathcal{G}_c$) to the traditional copyright (star network $\mathcal{G}_s$) when,

$$x_1((\mathcal{G}_c, \{1\})) = \sum_{b=2}^{n} \varpi_{b1} > x_1((\mathcal{G}_s, \{1\})) = \sum_{b=2}^{n} (\varpi_{b1} + \varpi_{bb})/2 \qquad (12)$$
$$\Leftrightarrow \sum_{b=2}^{n} \varpi_{b1} > \sum_{b=2}^{n} \varpi_{bb}.$$

The last inequality holds in Scenario A due to our assumptions in (9). This scenario illustrates that in information goods markets the strongest copyright protection is not necessarily the same as profit maximization (see, e.g., Shapiro and Varian, 1999).

■ **Efficiency.** Another important consideration related to network formation is efficiency. Following Jackson et al. (2005), we call a network efficient if it generates the largest value among all possible networks. When we ignore the linking costs, the value of a network (for a given set of sellers $\mathcal{S} \subseteq \mathcal{N}$) is easily computed in our context as the sum of all valuations and externalities,[9]

$$v(\mathcal{G},\mathcal{S}) \equiv \sum_{k\in\mathcal{N}} x_k((\mathcal{G},\mathcal{S})) = \sum_{b\in\mathcal{B}(\mathcal{G},\mathcal{S})} \sum_{k\in\mathcal{N}} \varpi_{bk}, \tag{13}$$

where the last equality follows by the iterative expansion of (4). A network $\mathcal{G}$ is, then, efficient relative to $v(\mathcal{G},\mathcal{S})$ if $v(\mathcal{G},\mathcal{S}) \geq v(\mathcal{G}',\mathcal{S})$ for all networks $\mathcal{G}'$ with the set of vertices $\mathcal{N}$. It is clear that there always exists at least one efficient network given that the set of relevant networks is finite. Alternatively, one can consider the standard notion of Pareto efficiency. Adapted to our context, network $\mathcal{G}$ Pareto dominates another network $\mathcal{G}'$ if

$$x_k((\mathcal{G},\mathcal{S})) \geq x_k((\mathcal{G}',\mathcal{S})), \quad \forall k \in \mathcal{N},$$

with at least one strict inequality.

The following proposition reports efficient connection structures and their values (net of the linking costs) for each scenario.

Proposition 7. *An efficient network $\mathcal{G}$ in the configuration $\mathcal{C} = (\mathcal{G}, \{1\})$ is in scenario*

A) a connected tree with $n-1$ links; $v(\mathcal{G},\{1\}) = \sum_{b=2}^{n}(\varpi_{b1} + \varpi_{bb})$.

B) the empty network; $v(\mathcal{G},\{1\}) = 0$.

C) a connected tree with $n-1$ links; $v(\mathcal{G},\{1\}) = \sum_{b=2}^{n} \varpi_{bb}$.

The results in this section illustrate a possible tension between profit maximization, pairwise stability and efficiency. For example, in scenario C (no externalities), the seller-centered star maximizes seller's revenue. Although this network is also efficient, it is not pairwise stable as only a line satisfies this property. On the other hand, in scenario B (negative externalities), the empty network is the unique efficient structure, whereas the seller-centered star maximizes seller's revenue. Neither network is pairwise stable as the latter criterion requires a connected network, where all buyers are covered by a cycle.

Our results offer some practical insights into copyright regulations. For example, the inequalities (12) show that distributing an IG under the GPL tends to be more beneficial

[9] In the terminology of Jackson et al. (2005), $x_k(\mathcal{G},\mathcal{S})$ is an *allocation rule* and $v(\mathcal{G},\mathcal{S}) \equiv \sum_{k\in\mathcal{N}} x_k(\mathcal{G},\mathcal{S})$ is a *value function.*

for its creator than the exclusive copyright when the sum of positive externalities exerted by the consumers on the creator outweighs the sum of consumers' valuations. This is likely the case for advertising but also for open source projects. In an open source project (e.g., Linux operating system), a lead-developer delivers an initial program to the community which is free to modify and to distribute it under the same licence terms. As a result, each contributor (including the lead-developer) acquires for free successive software releases. Similarly, GPL-like licences may be preferred by IG creators when the latter have direct access to only few prospective buyers, who in turn are connected to other buyers and so on. On the other hand, Scenario B in Proposition 7 suggests that some restrictions on information dissemination (e.g., censorship) may be necessary in order to achieve socially efficient outcomes in the presence of negative externalities.

6 An application to citation networks

A citation network is a directed graph in which each vertex represents a document and each directed edge maps a citation from one document to another.[10] Academic articles, court judgements, patents, web pages, etc. can be embedded in a citation graph. A typical application for citation graphs is the calculation of an impact measure of a document. An important impact metric is the citation count. For example, Trajtenberg (1990) shows that patent citations are indicative of the value of an innovation, whereas Hall et al. (2005) demonstrate that they significantly affect market value of the patent holder. However, citations exploit only a small portion of information contained in a citation graph. We apply our model to construct an impact index that takes advantage of the whole structure of such networks. Intuitively, our method captures the direct and (discounted) indirect impact of a vertex, i.e., it accounts for citations to this vertex, citations to the citations, etc. In the context of the information pricing model, we interpret the constructed index as the total price that an article generates in a citation graph.[11]

Specifically, we assume that each node $i \in \mathcal{N}$ in a directed acyclic graph (DAG)

[10] An illustration of a directed citation graph is provided in Figure 2.

[11] In a similar vein, Du et al. (2015) use equilibrium prices in a competitive economy to rank items in a (citation) network.

$\mathcal{G}$ ={$\mathcal{N}, \mathcal{L}$} creates its own IG (article). This article is then "sold" to all vertices that cite it, i.e. to all $b \in \mathcal{N}$ such that $ib \in \mathcal{L}$ ($b \in N_i(\mathcal{G})$). Each article has some value ϖ_{bb} for the buyer b that cites it and serves as an input, i.e., b can create its own article only after acquiring (reading) all articles that it cites. In other words, each buyer node resells the acquired articles after transforming (recombining) them into its own output. The matching and bargaining in $\mathcal{G}$ unfolds as in the original model. We can calculate, then, the impact index of an article i as the total price that the vertex i obtains from its sale, where we use i's bargaining power θ_{ib} vis-à-vis each buyer $b \in N_i(\mathcal{G})$ and ignore any prices that i has paid to the nodes that it cites.

Proposition 8. *For a directed acyclic graph $\mathcal{G}$, the total price that node i obtains from selling its article is computed by,*

$$p_i(\mathcal{G}) = \sum\nolimits_{b \in N_i(\mathcal{G})} \theta_{ib} \left\{ \varpi_{bb} + p_b(\mathcal{G}) \right\}, \tag{14}$$

Hence, the article i is rewarded not just for its direct but also for its indirect citations, where the weight of the latter decreases geometrically in their geodesic distance from the node i in graph $\mathcal{G}$. The price index (14) is reminiscent of the formula (7) and it subsumes a class of centrality measures for (directed and acyclic) graphs. For $\theta_{ik} = \theta$, it can be written in the vectorial form as,

$$\mathbf{p} = \theta G \left\{ \boldsymbol{\varpi} + \mathbf{p} \right\} = \theta G \boldsymbol{\varpi} + \theta G \mathbf{p},$$

where G is the adjacency matrix of the network $\mathcal{G}$, $\mathbf{p} = (p_i(\mathcal{G}))_{i \in \mathcal{N}}$ and $\boldsymbol{\varpi} = (\varpi_{bb})_{b \in \mathcal{N}}$. In this case, (14) belongs to the class of eigenvector-like (or generalized eigenvector) centralities defined in Bonacich and Lloyd (2001) by the system $\mathbf{c} = \mathbf{e} + \alpha G \mathbf{c}$ for an adjacency matrix G, scalar α and vector $\mathbf{e}$. In particular, (14) boils down to the Katz-Bonacich centrality ($\mathbf{e} = \mathbf{1}$) when,

$$\theta < 1/\lambda_{\max}(\mathcal{G}), \quad \sum\nolimits_{b \in N_i(\mathcal{G})} \varpi_{bb} = 1/\theta, \quad \forall i \in \mathcal{N},$$

and to the eigenvector centrality ($\mathbf{e} = \mathbf{0}$) when,

$$\theta = 1/\lambda_{\max}(\mathcal{G}), \quad \varpi_{ii} = 0, \quad \forall i \in \mathcal{N},$$

where $\lambda_{\max}(\mathcal{G})$ is the largest eigenvalue of the adjacency matrix of $\mathcal{G}$.

Centrality measures based on eigenvector methods can be used to rank vertices and can be also applied to weighted adjacency matrices such as citations across journals or hyperlinks between webpages. For example, Palacios-Huerta and Volij (2004) find a set of cardinal properties that axiomatize the invariant method for ranking journals, whereas Altman and Tennenholtz (2005) identify a set of ordinal properties that fully characterize the PageRank algorithm for ranking webpages. Both methods generalize the eigenvector centrality in a similar vein to (14). Although a rigorous discussion of network centralities and their application to measurement of intellectual influence is beyond the scope of this work, the interested reader is referred, e.g., to Jackson (2008, Chapter 2.2) and Bloch et al. (2017).

7 Concluding remarks

We consider an information trading framework, where information has some idiosyncratic value for its consumers, exerts externalities and is transmitted through links in a (directed) network. Bilateral trading leads in our model to unique information prices and players' payoffs in any fixed network. We use these payoffs to analyze a two-stage setting, where a communication network is formed before information diffuses.

This model has many prospective applications to, e.g., copyright and licensing regulations, internet based commerce (e-commerce), intelligence networks or data brokerage. For example, we show that an optimal copyright provision may depend on externalities exerted by buyers on the creator of an IG. In particular, the strongest protection does not necessarily maximize the total profit for the content creator. Furthermore, our scenario B suggests that in markets, where firms impose negative externalities on their competitors if they acquire information, these externalities are not internalized. If sufficiently strong, they can cannibalize any benefits from information acquisition. When we allow for decentralized link formation in this scenario, we observe that competing firms are able to obtain information for free but create inefficient networks. Finally, our scenario A suggests that direct marketing, where businesses communicate directly to customers (e.g., via text messages with promotional codes), is a more stable arrangement than mediated marketing, when the latter does not add value to advertisement. Referral bonus programs and advertising on social media are other promising areas where the present framework can be applied.

Appendix

Lemma 1. *Assume configuration $\mathcal{C}$ with $sb, s'b \in L(\mathcal{C})$ for $s \neq s'$ and fix the payoffs $x_k(\mathcal{C} \oplus b)$ for all $k \in \mathcal{N}$. Then, the NBS implies prices $p_{sb}(\mathcal{C}) = p_{s'b}(\mathcal{C}) = 0$ and payoffs $x_k(\mathcal{C}) = x_k^{sb}(\mathcal{C}) = x_k^{s'b}(\mathcal{C}) = \varpi_{bk} + x_k(\mathcal{C} \oplus b)$ for all $k \in \mathcal{N}$.*

Proof. *The NBS (2) applied to the link $sb \in L(\mathcal{C})$ with the payoffs from the trade in the link $s'b \in L(\mathcal{C})$ as disagreement values implies,*

$$x_s^{sb}(\mathcal{C}) - x_s^{s'b}(\mathcal{C}) = \theta_{sb}(S_{sb}(\mathcal{C}) - x_s^{s'b}(\mathcal{C}) - x_b^{s'b}(\mathcal{C}))$$
$$= \theta_{sb}(x_s(\mathcal{C}\oplus b) + \varpi_{bs} + x_b(\mathcal{C}\oplus b) + \varpi_{bb} - x_s^{s'b}(\mathcal{C}) - x_b^{s'b}(\mathcal{C})),$$

where we replaced $S_{sb}(\mathcal{C})$ from its definition (3). By substituting for $x_s^{sb}(\mathcal{C})$ and for $x_k^{s'b}(\mathcal{C})$, $k \in sb$, from (1), we can write the last equation as,

$$x_s^{sb}(\mathcal{C}) - x_s^{s'b}(\mathcal{C}) = x_s(\mathcal{C}\oplus b) + \varpi_{bs} + p_{sb}(\mathcal{C}) - (x_s(\mathcal{C}\oplus b) + \varpi_{bs}) =$$
$$\theta_{sb}(x_s(\mathcal{C}\oplus b) + \varpi_{bs} + x_b(\mathcal{C}\oplus b) + \varpi_{bb} - x_s(\mathcal{C}\oplus b) - \varpi_{bs} - x_b(\mathcal{C}\oplus b) - \varpi_{bb} + p_{s'b}(\mathcal{C}))$$
$$\Rightarrow p_{sb}(\mathcal{C}) = \theta_{sb} \cdot p_{s'b}(\mathcal{C}).$$

A symmetric condition can be derived for $p_{s'b}(\mathcal{C}) = \theta_{s'b} \cdot p_{sb}(\mathcal{C})$. Hence,

$$p_{sb}(\mathcal{C}) = \theta_{sb} \cdot p_{s'b}(\mathcal{C}) = \theta_{sb} \cdot \theta_{s'b} \cdot p_{sb}(\mathcal{C}) \Rightarrow p_{sb}(\mathcal{C}) = p_{s'b}(\mathcal{C}) = 0,$$

as $\theta_{sb}, \theta_{s'b} \in (0,1)$. The payoffs in $\mathcal{C}$ follow, then, from (1). ■

Proof. ***Proposition 1:*** *For a configuration $\mathcal{C}$ such that $\mathcal{B}(\mathcal{C}) = L(\mathcal{C}) = \emptyset$ (e.g., a configuration with no accessible buyers), $x_k(\mathcal{C}) = 0$ for all $k \in \mathcal{N}$ as no surplus is created by information trading. When $\mathcal{B}(\mathcal{C}) \neq \emptyset$, we assume as our inductive hypothesis that (4) specifies a unique, order-independent payoff $x_k(\mathcal{C}')$ to each player k in any configuration $\mathcal{C}'$ such that $\#\mathcal{B}(\mathcal{C}') < \#\mathcal{B}(\mathcal{C})$. In particular, in a configuration $\mathcal{C}'$ with $\mathcal{B}(\mathcal{C}') = \{b\}$ and $L(\mathcal{C}') = \{sb\}$, the NBS (2) with zero payoffs as disagreement points results in,*

$$x_k(\mathcal{C}') = x_k^{sb}(\mathcal{C}') = \mathbf{I}_{k \in sb} \cdot \theta_k^{sb}(\varpi_{bs} + \varpi_{bb}), \quad \forall k \in \mathcal{N}.$$

When at least two links, say sb and $s'b$, belong to $L(\mathcal{C}')$, Lemma 1 implies,

$$x_k(\mathcal{C}') = x_k^{sb}(\mathcal{C}') = x_k^{s'b}(\mathcal{C}') = \varpi_{bk}, \quad \forall k \in \mathcal{N}.$$

Importantly, all active links have been assumed to trade with probabilities bounded away from zero. We prove now a unique, order-independent payoff $x_k(\mathcal{C})$ for any configuration $\mathcal{C}$ and player $k \in \mathcal{N}$.

(1) Case $L(\mathcal{C}) = \{sb\}$: The claim follows directly from (1), the NBS (2) and our inductive hypothesis,

$$x_k(\mathcal{C}) = x_k^{sb}(\mathcal{C}) = \mathbf{I}_{k \in sb} \cdot \theta_k^{sb}(\varpi_{bs} + x_s(\mathcal{C} \oplus b) + \varpi_{bb} + x_b(\mathcal{C} \oplus b)).$$

(2) Case $\#L(\mathcal{C}) > 1$: We consider $sb, s'b' \in L(\mathcal{C})$ where $b \neq b'$ (the case $b = b'$ is covered in Lemma 1). When the link sb trades in $\mathcal{C}$, then $x_k^{sb}(\mathcal{C})$ is given by (1),

$$x_k^{sb}(\mathcal{C}) = \varpi_{bk} + (\mathbf{I}_{k=s} - \mathbf{I}_{k=b})p_{sb}(\mathcal{C}) + x_k(\mathcal{C} \oplus b) = \tag{A.1}$$
$$\varpi_{bk} + (\mathbf{I}_{k=s} - \mathbf{I}_{k=b})p_{sb}(\mathcal{C}) + \varpi_{b'k} + (\mathbf{I}_{k=s'} - \mathbf{I}_{k=b'})p_{s'b'}(\mathcal{C} \oplus b) + x_k(\mathcal{C} \oplus \{b, b'\}),$$

where in the second line in (A.1), we applied our inductive hypothesis to expand $x_k(\mathcal{C} \oplus b)$

according to (4) for the link $s'b' \in L(\mathcal{C} \oplus b)$*. By a similar argument,*

$$x_k^{s'b'}(\mathcal{C}) = \varpi_{b'k} + (\mathbf{I}_{k=s'} - \mathbf{I}_{k=b'})p_{s'b'}(\mathcal{C}) + x_k(\mathcal{C}\oplus b') = \tag{A.2}$$
$$\varpi_{b'k} + (\mathbf{I}_{k=s'} - \mathbf{I}_{k=b'})p_{s'b'}(\mathcal{C}) + \varpi_{bk} + (\mathbf{I}_{k=s} - \mathbf{I}_{k=b})p_{sb}(\mathcal{C}\oplus b') + x_k(\mathcal{C}\oplus\{b, b'\}).$$

From (A.1) and (A.2), we compute the difference,

$$x_k^{sb}(\mathcal{C}) - x_k^{s'b'}(\mathcal{C}) = (\mathbf{I}_{k=s} - \mathbf{I}_{k=b})D - (\mathbf{I}_{k=s'} - \mathbf{I}_{k=b'})D', \quad \textit{where}, \tag{A.3}$$
$$D \equiv p_{sb}(\mathcal{C}) - p_{sb}(\mathcal{C}\oplus b'), \quad D' \equiv p_{s'b'}(\mathcal{C}) - p_{s'b'}(\mathcal{C}\oplus b).$$

From (A.3), we obtain,

$$x_b^{sb}(\mathcal{C}) = x_b^{s'b'}(\mathcal{C}) - D, \quad x_s^{sb}(\mathcal{C}) = x_s^{s'b'}(\mathcal{C}) + D - \mathbf{I}_{s=s'}D', \tag{A.4}$$
$$x_{b'}^{s'b'}(\mathcal{C}) = x_{b'}^{sb}(\mathcal{C}) - D', \quad x_{s'}^{s'b'}(\mathcal{C}) = x_{s'}^{sb}(\mathcal{C}) + D' - \mathbf{I}_{s=s'}D.$$

We can then compute the net surpluses,

$$\begin{aligned} S_{sb} - d_s^{sb} - d_b^{sb} &= x_s^{sb}(\mathcal{C}) + x_b^{sb}(\mathcal{C}) - (x_s^{s'b'}(\mathcal{C}) + x_b^{s'b'}(\mathcal{C})) = -\mathbf{I}_{s=s'}D', \\ S_{s'b'} - d_{s'}^{s'b'} - d_{b'}^{s'b'} &= x_{s'}^{s'b'}(\mathcal{C}) + x_{b'}^{s'b'}(\mathcal{C}) - (x_{s'}^{sb}(\mathcal{C}) + x_{b'}^{sb}(\mathcal{C})) = -\mathbf{I}_{s=s'}D, \end{aligned} \tag{A.5}$$

and the NBS (2) payoffs,

$$\begin{aligned} x_b^{sb}(\mathcal{C}) &= x_b^{s'b'}(\mathcal{C}) - (1-\theta_{sb})\mathbf{I}_{s=s'}D', \quad x_s^{sb}(\mathcal{C}) = x_s^{s'b'}(\mathcal{C}) - \theta_{sb}\mathbf{I}_{s=s'}D', \\ x_{b'}^{s'b'}(\mathcal{C}) &= x_{b'}^{sb}(\mathcal{C}) - (1-\theta_{s'b'})\mathbf{I}_{s=s'}D, \quad x_{s'}^{sb}(\mathcal{C}) = x_{s'}^{sb}(\mathcal{C}) - \theta_{s'b'}\mathbf{I}_{s=s'}D. \end{aligned} \tag{A.6}$$

For the case $s \neq s'$*, the claim* $x_k^{sb}(\mathcal{C}) = x_k^{s'b'}(\mathcal{C})$ *for* $k \in \{s, s', b, b'\}$ *follows immediately from (A.6) as* $\mathbf{I}_{s=s'} = 0$*. When* $s = s'$*, we compare the differences* $x_b^{sb}(\mathcal{C}) - x_b^{s'b'}(\mathcal{C})$ *and* $x_{b'}^{s'b'}(\mathcal{C}) - x_{b'}^{sb}(\mathcal{C})$ *in (A.4) and in (A.6), which leads to the system of two equations in D* *and D' and its unique solution,*

$$\begin{cases} D = (1-\theta_{sb})D' \\ D' = (1-\theta_{s'b'})D \end{cases} \Rightarrow D' = D = 0,$$

because θ_{sb}, $\theta_{s'b'} \in (0,1)$. Hence, we have shown that $x_k^{sb}(\mathcal{C}) = x_k^{s'b'}(\mathcal{C})$ for all $k \in \{s,b,s',b'\}$ and any $sb, s'b' \in L(\mathcal{C})$ with $sb \neq s'b'$. For $k \notin \{s,b,s',b'\}$ order independence follows from our inductive hypothesis and (1),

$$x_k^{sb}(\mathcal{C}) = \varpi_{bk} + x_k(\mathcal{C}\oplus b) = \varpi_{bk} + \varpi_{b'k} + x_k(\mathcal{C}\oplus\{b,b'\}) = \varpi_{b'k} + x_k(\mathcal{C}\oplus\{b'\}) = x_k^{s'b'}(\mathcal{C}).$$

Hence, without the loss of generality, we can assume, that $sb \in L(\mathcal{C})$ trades in $\mathcal{C}$ and compute the payoff to player $k \in \mathcal{N}$ by (1),

$$x_k(\mathcal{C}) = x_k^{sb}(\mathcal{C}) = \varpi_{bk} + x_k(\mathcal{C}\oplus b) + (\mathbf{I}_{k=s} - \mathbf{I}_{k=b})p_{sb}(\mathcal{C}),$$

where $p_{sb}(\mathcal{C})$ is computed by (1) and order independence,

$$x_b^{sb}(\mathcal{C}) = x_b^{s'b'}(\mathcal{C}) \Rightarrow p_{sb}(\mathcal{C}) = \varpi_{bb} + x_b(\mathcal{C}\oplus b) - \varpi_{b'b} - x_b(\mathcal{C}\oplus b').$$

Hence, $x_k(\mathcal{C})$ is a linear combination of payoffs in configurations with one more buyer. These payoffs are unique by our inductive hypothesis. We note that none of the arguments in this proof depends on particular matching or agreement probabilities as long as these probabilities are bounded away from zero. ■

Proof. ***Proposition 2:***

Case 1: $sb, s'b' \in L(\mathcal{C})$ and $sb \neq s'b'$. If $b \neq b'$, then the claim follows from Proposition 1 and (1):

$$\begin{aligned} x_b^{sb}(\mathcal{C}) &= x_b^{sb'}(\mathcal{C}) \Rightarrow \varpi_{bb} + x_b(\mathcal{C}\oplus b) - p_{sb}(\mathcal{C}) = \varpi_{b'b} + x_b(\mathcal{C}\oplus b') \\ &\Rightarrow p_{sb}(\mathcal{C}) = \varpi_{bb} + x_b(\mathcal{C}\oplus b) - \varpi_{b'b} - x_b(\mathcal{C}\oplus b'). \end{aligned} \tag{A.7}$$

If $b = b'$ *(and, hence,* $s \neq s'$*) then* $p_{sb}(\mathcal{C}) = 0$ *by Lemma 1 and (5) yields correctly:*

$$p_{sb}(\mathcal{C}) = \varpi_{bb} + x_b(\mathcal{C}\oplus b) - \varpi_{bb} - x_b(\mathcal{C}\oplus b) = 0.$$

Case 2: $L(\mathcal{C}) = \{sb\}$*: First, we compute* $x_s(\mathcal{C})$ *by (2) with* $d_s^{sb}(\mathcal{C}) = d_b^{sb}(\mathcal{C}) = 0$,

$$x_s(\mathcal{C}) = \theta_{sb} S_{sb}(\mathcal{C}) = \theta_{sb}(\varpi_{bs} + x_s(\mathcal{C}\oplus b) + \varpi_{bb} + x_b(\mathcal{C}\oplus b)). \tag{A.8}$$

Then, the price $p_{sb}(\mathcal{C})$ *is readily computed from (4),*

$$p_{sb}(\mathcal{C}) = x_s(\mathcal{C}) - \varpi_{bs} - x_s(\mathcal{C}\oplus b) = \theta_{sb} S_{sb}(\mathcal{C}) - \varpi_{bs} - x_s(\mathcal{C}\oplus b). \tag{A.9}$$

Finally, in order to prove $p_{sb}(\mathcal{C}) = p_{sb}(\mathcal{C}\oplus b')$ *when* $b \neq b'$*, we use order independence,*

$$x_b^{sb}(\mathcal{C}) = x_b^{s'b'}(\mathcal{C}) \Rightarrow x_b(\mathcal{C}\oplus b) + \varpi_{bb} - p_{sb}(\mathcal{C}) = x_b(\mathcal{C}\oplus b') + \varpi_{b'b},$$

and expand $x_b(\mathcal{C}\oplus b)$ *and* $x_b(\mathcal{C}\oplus b')$ *by (4),*

$$x_b(\mathcal{C}\oplus\{b, b'\}) + \varpi_{b'b} + \varpi_{bb} - p_{sb}(\mathcal{C}) = x_b(\mathcal{C}\oplus\{b', b\}) + \varpi_{bb} - p_{sb}(\mathcal{C}\oplus b') + \varpi_{b'b}.$$

From the last equation, we obtain $p_{sb}(\mathcal{C}) = p_{sb}(\mathcal{C} \oplus b')$.

■

Proof. ***Proposition 3:*** *By order independence (Proposition 1), the price that* b *pays for information is independent of the order of trades. Hence, we compute this price in configuration* $\mathcal{C}'$ *where two neighbors of* b *but not* b *herself have acquired information. Such a configuration can be reached by information diffusion from the original configuration* $\mathcal{C}$ *due to* b*'s two paths property. By Lemma 1, the price that* b *pays in configuration* $\mathcal{C}'$ *is zero.*

■

Proof. ***Proposition 4****: For a configuration* $\mathcal{C} = (\mathcal{G}, \{s\})$ *with the single seller* s*, we define the configuration* $\mathcal{C}^b = (\mathcal{G}^b, \{s\})$*, where* $\mathcal{G}^b$ *is a (possibly disconnected) subnetwork of* $\mathcal{G}$ *containing the seller* s *and the largest set* $\mathcal{B}^b \subseteq \mathcal{B} \equiv \mathcal{B}(\mathcal{C})$ *of buyers that are accessible in*

$\mathcal{G}$ from $\{s\}$ via b only,

$$\mathcal{B}^b \equiv \{k \in \mathcal{B} : b \in \delta^{s \to k}(\mathcal{G})\}.$$

We define further the set of buyers $\mathcal{D}_i^d \subseteq \mathcal{B}(\mathcal{C})$ at the "δ-distance" $d \geq 1$ from $i \in \mathcal{N}$,

$$\mathcal{D}_i^d \equiv \{k \in \mathcal{B} : \#\delta^{i \to k} = d\}.$$

Moreover, to ease the notation, we define the set $\Psi \equiv N_s(\mathcal{G}) \cap \mathcal{D}_s^2$ of direct neighbors of s accessible from $\{s\}$ by one path only. If $\Psi = \emptyset$, then each neighbor $b \in N_s(\mathcal{G})$ of s satisfies the two paths property and gets the IG for free. Hence, the recursive expansion of (4) yields,

$$x_s(\mathcal{C}) = x_s(\mathcal{C} \oplus b) + \varpi_{bs} = ... = \sum_{b \in \mathcal{B}} \varpi_{bs},$$

which confirms (4) because $\delta^{s \to b} = (s)$ for all $b \in N_s(\mathcal{G})$ implies $\delta^{s \to k} = (s)$ for all $k \in \mathcal{B}$.

If $\Psi \neq \emptyset$, we expand (4) recursively for each $b \in \Psi$,

$$\begin{aligned} x_s(\mathcal{C}) &= x_s(\mathcal{C} \oplus b) + \varpi_{bs} + p_{sb}(\mathcal{C}) = ... \\ &= x_s(\mathcal{C} \oplus \Psi) + \sum_{b \in \Psi} \varpi_{bs} + \sum_{b \in \Psi} p_{sb}(\mathcal{C}). \end{aligned} \tag{A.10}$$

where we used the fact that $p_{sb}(\mathcal{C})$ does not depend on the configuration (by the iterative application of $p_{sb}(\mathcal{C}) = p_{sb}(\mathcal{C} \oplus b')$ for $b \neq b'$ proved in Proposition 2). Specifically, we compute $p_{sb}(\mathcal{C}) = p_{sb}(\mathcal{C}^b)$ in configuration $\mathcal{C}^b$ with the single active link sb from (5),

$$\begin{aligned} p_{sb}(\mathcal{C}) &= \theta \cdot S_{sb}(\mathcal{C}^b) - \varpi_{bs} - x_s(\mathcal{C}^b \oplus b) \\ &= \theta \left\{ x_b(\mathcal{C}^b \oplus b) + x_s(\mathcal{C}^b \oplus b) + \varpi_{bb} + \varpi_{bs} \right\} - \varpi_{bs} - x_s(\mathcal{C}^b \oplus b), \end{aligned}$$

and substitute the computed prices into (A.10),

$$x_s(\mathcal{C}) = x_s(\mathcal{C}\oplus\Psi) + \theta \sum_{b\in\Psi} S_{sb}(\mathcal{C}^b) - \sum_{b\in\Psi} x_s(\mathcal{C}^b\oplus b) \tag{A.11}$$
$$= \sum_{b\in\mathcal{B}\backslash\Psi} \varpi_{bs} + \theta \sum_{b\in\Psi} S_{sb}(\mathcal{C}^b) - \sum_{b\in\Psi} \sum_{v\in\mathcal{B}^b\backslash b} \varpi_{vs}$$
$$= \theta \sum_{b\in\Psi} \left\{ x_b(\mathcal{C}^b\oplus b) + x_s(\mathcal{C}^b\oplus b) + \varpi_{bb} + \varpi_{bs} \right\} + \sum_{b\in\mathcal{D}_s^1} \varpi_{bs}.$$

The second line in (A.11) follows by (4) because seller s gets the price of zero from buyers in $\mathcal{B}(\mathcal{C}\oplus\Psi) = \mathcal{B}\backslash\Psi$ and never trades with buyers in $\mathcal{B}(\mathcal{C}^b\oplus b) = \mathcal{B}^b\backslash b$ for each $b\in\Psi$. The third line follows because $\mathcal{B}\backslash\Psi = (\cup_{b\in\Psi}\mathcal{B}^b\backslash b)\cup\mathcal{D}_s^1$. In order to evaluate (A.11), we simplify notation by defining $\rho_k^{s\to b} \equiv \min\{\#\delta^{s\to k}, \#\delta^{s\to b} - 1)$ and calculate first,

$$\theta \sum_{b\in\Psi} x_b(\mathcal{C}^b\oplus b) = \sum_{b\in\Psi} \sum_{v\in\mathcal{B}^b\backslash b} \left\{ \sum_{k\in\delta^{b\to v}} \theta\cdot\theta^{\rho_k^{b\to v}}\cdot\varpi_{vk} \right\} =$$
$$\sum_{v\in\cup_{b\in\Psi}\mathcal{B}^b\backslash b} \left\{ \sum_{k\in\delta^{s\to v}} \theta^{\rho_k^{s\to v}}\cdot\varpi_{vk} - \theta\varpi_{vs} \right\}, \tag{A.12}$$

by applying (8) to each $x_b(\mathcal{C}^b\oplus b)$. Then, we evaluate the remaining term in (A.11),

$$\theta \sum_{b\in\Psi} \left\{ x_s(\mathcal{C}^b\oplus b) + \varpi_{bb} + \varpi_{bs} \right\} + \sum_{b\in\mathcal{D}_s^1} \varpi_{bs}$$
$$= \theta \sum_{b\in\Psi} \sum_{v\in\mathcal{B}^b\backslash b} \varpi_{vs} + \theta \sum_{b\in\Psi} \{\varpi_{bb} + \varpi_{bs}\} + \sum_{b\in\mathcal{D}_s^1} \varpi_{bs}$$
$$= \sum_{v\in\cup_{b\in\Psi}\mathcal{B}^b\backslash b} \theta\varpi_{vs} + \sum_{v\in\Psi\cup\mathcal{D}_s^1} \left\{ \sum_{k\in\delta^{s\to v}} \theta^{\rho_k^{s\to v}}\cdot\varpi_{vk} \right\}. \tag{A.13}$$

The second line in (A.13) follows by (4) because seller s never trades with buyers in $\mathcal{B}(\mathcal{C}^b\oplus b) = \mathcal{B}^b\backslash b$ for each $b\in\Psi$. The third line follows by the fact that $\delta^{s\to v} = (s,v)$ and $\rho_s^{s\to v} = \rho_v^{s\to v} = 1$ for each $v\in\Psi$ and $\delta^{s\to v} = (s)$ and $\rho_s^{s\to v} = 0$ for each $v\in\mathcal{D}_s^1$. Then, the sum of (A.12) and (A.13) yields the claim because $\mathcal{B} = (\cup_{b\in\Psi}\mathcal{B}^b\backslash b)\cup\ \Psi\cup\mathcal{D}_s^1$. ■

Proof. ***Proposition 5.*** *Pairwise stable (PS) structures in the relevant scenarios are shown below. The corresponding seller's payoffs follow then from (8).*

A) (positive externalities): First, we show that PS networks (PSN) are connected. For the sake of contradiction, suppose there is no path between the initial seller 1 and some *node* $i \in \mathcal{N}\backslash\{1\}$. *Then,* i *never obtains the IG and will not maintain any (costly) links. Hence,* i *is an isolated singleton. But this is incompatible with PS as* 1 *and* i *would benefit by creating the link* $\overline{1i}$ *with the value* $\varpi_{ii} + \varpi_{i1} > 0$ *to share among themselves.*

Secondly, we show that there are no links in a PSN between prospective buyers $v, w \in \mathcal{N}\backslash\{1\}$. *To see this, we use the order independence and consider the configuration* $\mathcal{C}$ *that arises after the initial seller (and only this player) has traded with all his linked neighbours. Then, the surplus for any link* $\overline{vw}$, $v, w \in \mathcal{N}\backslash\{1\}$, *when* v *acts as seller, verifies,*

$$S_{vw}(\mathcal{C}) = \varpi_{wv} + x_v(\mathcal{C} \oplus w) + \varpi_{ww} + x_w(C \oplus w) \leq 0,$$

as only non-positive values and externalities are created for v *and* w *in* C *or in any ensuing configuration due to the assumption in (9) that* $\varpi_{vw} \leq 0$ *for all* $v, w \in \mathcal{N}\backslash\{1\}$. *Hence,* v *and* w *would benefit from the deletion of* $\overline{vw}$.

B) (negative externalities): The proof of connectivity is the same as in scenario A.

Next, we show that trade across a bridge in this scenario leads to negative payoffs to the involved players. For a bridge $\overline{sb}$ *consider the configuration* C *where* $\overline{sb}$ *is the only active link. Then, the surplus,*

$$S_{sb}(\mathcal{C}) = \varpi_{bs} + x_s(\mathcal{C} \oplus b) + \varpi_{bb} + x_b(C \oplus b) = \varpi_{bs} + \varpi_{bb} + x_b(C \oplus b),$$

is negative by our assumption of sufficiently strong externalities, $|\varpi_{ik}| > \max_i \varpi_{ii}$ *for all* $i, k \in \mathcal{N}\backslash\{1\} : i \neq k$ *and by Proposition 4, which decomposes* $x_b(C \oplus b)$ *into a (weighted) sum of (negative) externalities. Given their disagreement points of zero, the NBS implies the share of* $S_{sb}(\mathcal{C})/2 \leq 0$ *to both* s *and* b.

We show now that a connected collection $\mathcal{G}$ *of cycles, any two of which sharing at most one node, is PS. As every buyer node in* $\mathcal{G}$ *is covered by a cycle, all of them satisfy the two paths property and get the IG for free. None of the players will then benefit from adding a link. On the other hand, if node* i *that is covered by the cycle* $(i, v, ..., w, i)$ *cut one of its links, say* $\overline{iw}$, *then* $\overline{iv}$ *would become a bridge. Otherwise, there would be a node* r *and a cycle* $(i, v, ..., r, i)$. *But then, two cycles in* $\mathcal{G}$, $(i, v, ..., w, i)$ *and* $(i, v, ..., r, i)$, *would share* *two nodes, i and v, which contradicts the definition of $\mathcal{G}$. Hence, we conclude that $\mathcal{G}$ is PS because adding or deleting links to $\mathcal{G}$ decreases the payoffs to the involved nodes. Any other structures cannot be PS as they either contain a bridge or allow for link deletion without destroying the two paths property of some node.*

C) (no externalities): We show first that the line $\mathcal{G}_l$ with the set of links $\mathcal{L}_l = \{i(i+1), i = 1, ..., n-1\}$ is pairwise stable. First, we note that a node that deletes one of its links in $\mathcal{G}_l$ either looses the access to the IG or to the (resale) market. Given symmetric bargaining powers and strictly positive valuations, link deletion results then in lower payoffs to the involved nodes. On the other hand, if a new link ik, $i < i+1 < k$, is created then i acts as the seller whenever information flows through this link. However, i will be forced to sell at the price of zero to k and to $i+1$ as these players satisfy the two paths property. Although $i > 1$ will pay now a reduced price to $i-1$ for the IG, this reduction will compensate only for the half of the loss of the resale value. Hence, the overall profit to i will be lower than in the line. This violates the condition on link addition in the definition of pairwise stability.

Secondly, we show that any PS network is a line, where node 1 has one link: For the single initial seller, node 1, we note that in a PSN it holds for $i = 1$:

1) For any link $\overline{ik}$, where $k \in \{i+1, ..., n\}$ is a buyer, the price paid by k to i for the IG is strictly positive as otherwise i would delete the link $\overline{ik}$.

2) Node i has only one link, say with the player $i+1$. If there existed two links, $\overline{ij}$ and $\overline{ik}$ for $j \neq k$, then by 1), j and k would pay strictly positive prices for the IG. This is, however, incompatible with PS as j and k could get the IG for free by creating the link $\overline{jk}$.

We can repeat the arguments 1) − 2) for each IG reseller $i \in \{2, ..., n-1\}$. ■

Proof. ***Proposition 6.*** *Optimal structures in the relevant scenarios are shown below. The corresponding seller's payoffs follow then from (8).*

A) (positive externalities): The maximum value that the seller can extract in this scenario is the sum $\varpi_{21} + ... + \varpi_{n1}$ of positive externalities. This is only possible if the network is connected (i.e. information diffuses to all players) and when there are no positive transfers from the seller to the buyers. The latter condition is satisfied when all buyers connected to the seller are covered by a cycle as in this case they acquire the IG for free. Moreover, *the seller can save linking costs by connecting to just two such buyers.*

B) (negative externalities): See scenario C and note that, by (8), the seller in the seller-centered star does not internalize any share of the negative externalities.

C) (no externalities): First, we note that (8) simplifies to the following expression when $\varpi_{ki} = 0$ *for all* $i \neq k$,

$$x_s(\mathcal{C}) = \sum_{b \in \mathcal{B}(\mathcal{C}): b \in \delta^{s \to b}} \varpi_{bb} \cdot \theta^{\#\delta^{s \to b} - 1}.$$

Hence, an optimal network must satisfy two properties: Each buyer b *must belong to the critical path* $\delta^{s \to b}$ *(in order to extract a share of* $\varpi_{bb} > 0$*) and this path must be as short as possible (in order to reduce the discounting by* $\theta^{\#\delta^{s \to b} - 1}$*). These properties imply a direct link from* s *to each buyer* b*. Moreover, no connections between buyers should exist. For if such a link existed, then the involved buyers would satisfy the two paths property and could not belong to a critical path starting at* s*. The only network with these characteristics is the seller-centered star .* ■

Proof. ***Proposition 7.*** *Efficient structures in the relevant scenarios are shown below. The corresponding network values follow then from (13).*

A) (positive externalities): As each buyer b *creates the value* $\varpi_{bb} + \varpi_{b1} > 0$ *by acquiring the IG, an efficient network must be connected (i.e. information diffuses to all players). This connectivity should be achieved with the smallest number of links in order to save linking costs.*

B) (negative externalities): As each buyer b *creates the value* $\varpi_{bb} + \sum_{k \neq b} \varpi_{bk} < 0$ *by acquiring the IG, an efficient network must be fully disconnected (i.e. no information diffusion).*

C) (no externalities): As each buyer b *creates the value* $\varpi_{bb} > 0$ *by acquiring the IG, an efficient network must be connected (i.e. information diffuses to all players). This connectivity should be achieved with the smallest number of links in order to save linking costs.* ■

Proof. ***Proposition 8****: First, we show that all nodes in* $\mathcal{G}$ *with some outgoing links sell their articles in the information trading stage. For each node* i*, we define the subgraph* $\mathcal{G}_i = \{\mathcal{N}_i, \mathcal{L}_i\}$ *that starts with* i *and includes all nodes accessible from it in* $\mathcal{G}$*. As* $\mathcal{G}_i$ *is* *also a DAG, it has a (not necessarily unique) topological ordering* $(v^1, ..., v^{\#\mathcal{N}_i})$ *of the set of its nodes, i.e., for each link* $v^s v^k \in \mathcal{L}_i$*, we have* $s < k$*. An inductive algorithm to find such an ordering works as follows: Start with the graph* $\mathcal{G}_i^1 = \{\mathcal{N}_i^1, \mathcal{L}_i^1\} = \mathcal{G}_i$ *and the empty ordering* O*. For each* $k = 1, 2..., \#\mathcal{N}_i$ *repeat the following steps: If* $\mathcal{N}_i^k \neq \emptyset$ *find the set* $\{s \in \mathcal{N}_i^k : \nexists vs \in \mathcal{L}_i^k\}$ *of nodes in* $\mathcal{G}_i^k$ *with no incoming links. As each* $\mathcal{G}_i^k$ *is a DAG, the set of such nodes is non-empty. Append these nodes (in any order) to* O *and remove them from* $\mathcal{G}_i^k$*. This leads to the subgraph* $\mathcal{G}_i^{k+1} = \{\mathcal{N}_i^{k+1}, \mathcal{L}_i^{k+1}\}$ *of* $\mathcal{G}_i^k$*. This procedure ends after a finite number of steps (when* $\mathcal{N}_i^k = \emptyset$*) with a topological ordering* O*.*

When only i *but none of its followers in* O *has created an article, the existence of a topological ordering implies that the trading process in* $\mathcal{G}_i$ *only stops after every other node* $k \in \mathcal{N}_i$ *has produced and sold its article to all neighbor nodes in* $N_k(\mathcal{G}_i)$*. To see this, we observe that each node* v^k *in the topological ordering* $O = (v^1, ..., v^{\#\mathcal{N}_i})$ *is able to produce and sell its article when all its predecessors* v^s*,* $s < k$*, in* $\{v^s : v^s v^k \in \mathcal{L}_i\}$ *have sold their articles to* v^k*. Hence, if* v^k *is unable to produce its output it is because* v^k *has not yet traded with all of its predecessors in* $\{v^s : v^s v^k \in \mathcal{L}_i\}$*. But then, either* k *can trade with the missing predecessors or at least one of them has not yet created its IG. We take first such predecessor in* O *and repeat the above argument. Eventually, we will reach the first node* $v^1 = i$*, which, by construction of* $\mathcal{G}_i$*, already possesses its article. Hence, at least one pair can trade in* $\mathcal{G}_i$ *whenever not all nodes have sold their articles to their neighbor successors.*

Secondly, we have to show that order independence holds in the present context and it leads to unique payoffs. The proof follows the steps of the proof of Proposition 1 but it requires additional notation to account for several IGs (articles) traded in $\mathcal{G}_i$*. For the sake of brevity, we omit here this mechanical exercise.*

Finally, the pricing formula (14) is a direct analogue of (7) in the current context of directed graphs and multiple IGs (articles) that do not exert externalities. As each node s *is the unique seller of her article, the relevant network for trading this article is the* s*-centered star and the critical path to any buyer* $b \in N_s(\mathcal{G})$ *in this star is* $\delta^{s \to b} = (s, b)$*. The formula (14) follows then from the arguments below the formula (7).* ■

References

Ali, N.S., Chen-Zion, A. and Lillethun, E. "Reselling Information." working paper, 2017.

Altman, A., Tennenholtz, M. "Ranking systems: The PageRank axioms." *Proceedings of the 6th ACM conference on Electronic commerce (EC-05).* 2005.

Bhaskar, V. and Vega-Redondo, F. "Migration and the evolution of conventions." *Journal of Economic Behavior & Organization*, Vol. 55 (2004), pp. 397-418.

Binmore, K., Rubinstein, A., Wolinsky, A. "The Nash Bargaining Solution in Economic Modelling." *RAND Journal of Economics*, 17 (1986), pp. 176-188.

Bloch, F., G. Demange and R. Kranton "Rumors and Social Networks." *International Economic Review*, Vol. 59 (2018), pp. 421-448.

Bloch, F., Jackson, M. and P. Tebaldi "Centrality measures in networks." working paper, 2017.

Bloch, F. "Targeting and pricing in social networks." In Bramoullé, Y., Galeotti, A. and Rogers, B., eds., *The Oxford Handbook of the Economics of Networks*, Oxford University Press, 2016.

Bonacich, P., Lloyd, P. "Eigenvector-like measures of centrality for asymmetric relations." *Social Networks*, Vol. 23 (2001), pp. 191-201.

Calvó-Armengol, A., de Martí, J. and Prat, A. "Communication and influence." *Theoretical Economics*, Vol. 10 (2015), pp. 649–690.

Chwe, K. "Communication and coordination in social networks." *Review of Economic Studies*, Vol. 67 (2000), pp. 1-16.

Crémer, J. "Network externalities and universal service obligation in the internet." *European Economic Review*, Vol. 44 (2000), pp. 1021-1031.

Du, Y., Lehrer, E. and Pauzner, A. "Competitive Economy as a Ranking Device Over Networks." *Games and Economic Behavior* Vol. 91 (2015), pp. 1-13.

Elliott, M. "Inefficiencies in Networked Markets." *American Economic Journal: Microeconomics*, Vol. 7 (2015), pp. 43-82.

Ellison, G. "Learning, local interaction and coordination." *Econometrica*, Vol. 61

(1993), pp. 1047-1071.

Fainmesser, I.P., Galeotti, A. "Pricing network effects." *Review of Economic Studies*, Vol. 83 (2015), pp. 165-198.

Galbreth, M.R., Ghosh, B. "Social Sharing of Information Goods: Implications for Pricing and Profits." *Marketing Science*, Vol. 31 (2012), pp. 603–620.

Galeotti, A., Ghiglinoz, C., Squintani, F. "Strategic Information Transmission in Networks." *Journal of Economic Theory*, Vol. 148 (2013), pp. 1751–1769.

Galeotti, A., Goyal, S. "The law of the few." *American Economic Review*, Vol. 100 (2010), pp. 1468-1492.

Gal-Or, E. "Information sharing in oligopoly." *Econometrica*, Vol. 53 (1985), pp. 329–343.

Hagenbach, J, and F. Koessler, "Strategic Communication Networks." *Review of Economic Studies*, Vol. 77 (2010), pp. 1072-1099.

Hedetniemi, S., Liestman, A. "A survey of gossiping and broadcasting in communication networks." *Networks*, Vol. 18 (1988), pp. 319-349.

Hall, B., Jaffe, A. and Trajtenberg, M. "Market value and patent citations." *RAND Journal of Economics*, Vol. 36 (2005), pp. 16-38.

Jackson M., Demange G., Wooders M. "A survey of network formation models: Stability and efficiency." In *Group Formation in Economics: Networks, Clubs, and Coalitions*, Cambridge University Press. 2005.

Jackson, M. *Social and economic networks*, Princeton: Princeton University Press. 2008.

Jeon, D., Laffont, J., Tirole, J. "On the receiver-pays principle." *RAND Journal of Economics*, Vol. 35 (2004), pp. 85-110.

Koessler, F., Renault, R. "When does a firm disclose product information?" *RAND Journal of Economics*, Vol 43 (2012), p. 630.

Manea, M. "Models of bilateral trade in networks." In Bramoullé, Y., Galeotti, A. and Rogers, B., eds., *The Oxford Handbook of the Economics of Networks*, Oxford University Press, 2016.

Manea, M. "Bottleneck Links, Essential Intermediaries, and Competing Paths of Diffusion in Networks." working paper, 2017.

Mauleon, A., Vannetelbosch, V. "Network formation games." In Bramoullé, Y., Galeotti, A. and Rogers, B., eds., *The Oxford Handbook of the Economics of Networks*, Oxford University Press, 2016.

Mueller-Frank, M. "A general framework for rational learning in social networks." *Theoretical Economics*, Vol. 8 (2013), pp. 1–40.

Muto, S. "An information good market with symmetric externalities." *Econometrica*, Vol. 54 (1986), pp. 295-312.

Palacios-Huerta, I. Volij, O. "The measurement of intellectual influence." *Econometrica*, Vol. 72 (2004), pp. 963-977.

Polanski, A. "A decentralized model of information pricing in networks." *Journal of Economic Theory*, Vol. 136 (2007), pp. 497–512.

Probst, F., Grosswiele, L. and Peger, R. "Who will lead and who will follow: Identifying Influential Users in Online Social Networks." *Business & Information Systems Engineering* Vol. 5 (2013), pp. 179-193.

Redlicki, B. "Rumours and Cheap Talk." working paper, 2015.

Shapiro, C. and Varian, H. *Information Rules: A Strategic Guide to the Network Economy.* Boston, Massachusettes: Harvard Business School Press, 1999.

Trajtenberg, M. "A Penny for Your Quotes: Patent Citations and the Value of Innovations." *RAND Journal of Economics*, Vol. 21 (1990), pp. 172–189.

Varian, H. "Buying, selling and renting information goods." *Journal of Industrial Economics*, Vol. XLVIII (2000), pp. 473-488.

Wang, P., Watts, A. "Formation of buyer-seller trade networks in a quality-differentiated product market." *Canadian Journal of Economics*, Vol. 39 (2006), pp. 971–1004.

Figure 1: Networks with a single information seller (dark node 1) and three prospective buyers (light nodes 2-4). Solid lines - information transmission links, dotted lines - positive externalities, dashed lines - negative externalities.

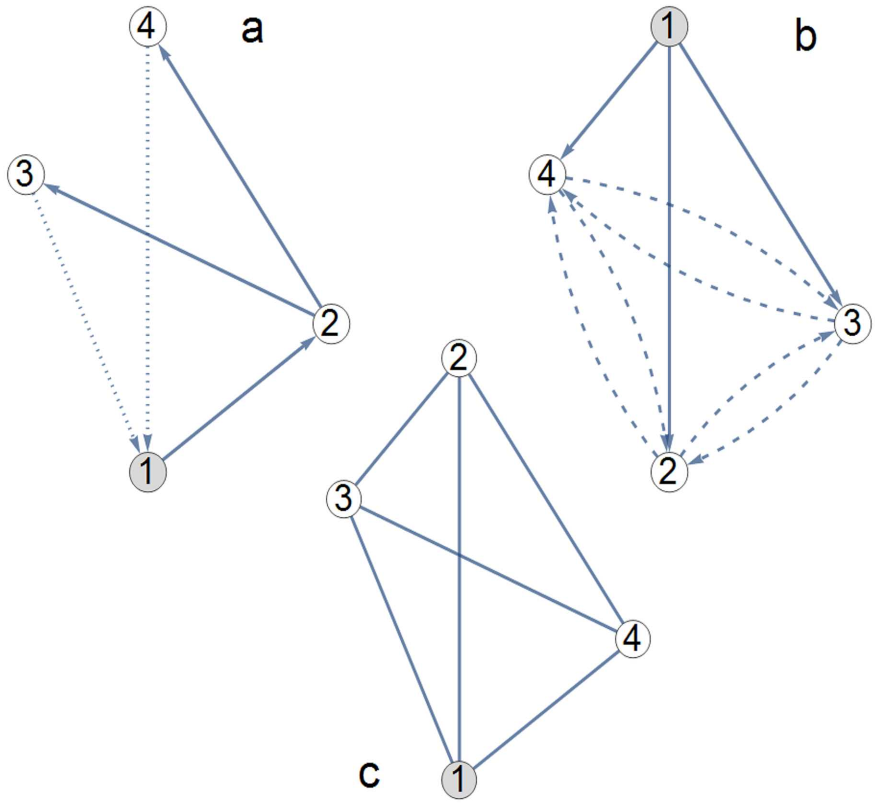

Figure 2: Critical paths in a directed (citation) graph: $\delta^{71\to 9232} = (71, 3412, 7149)$, $\delta^{71\to 7565} = (71)$.

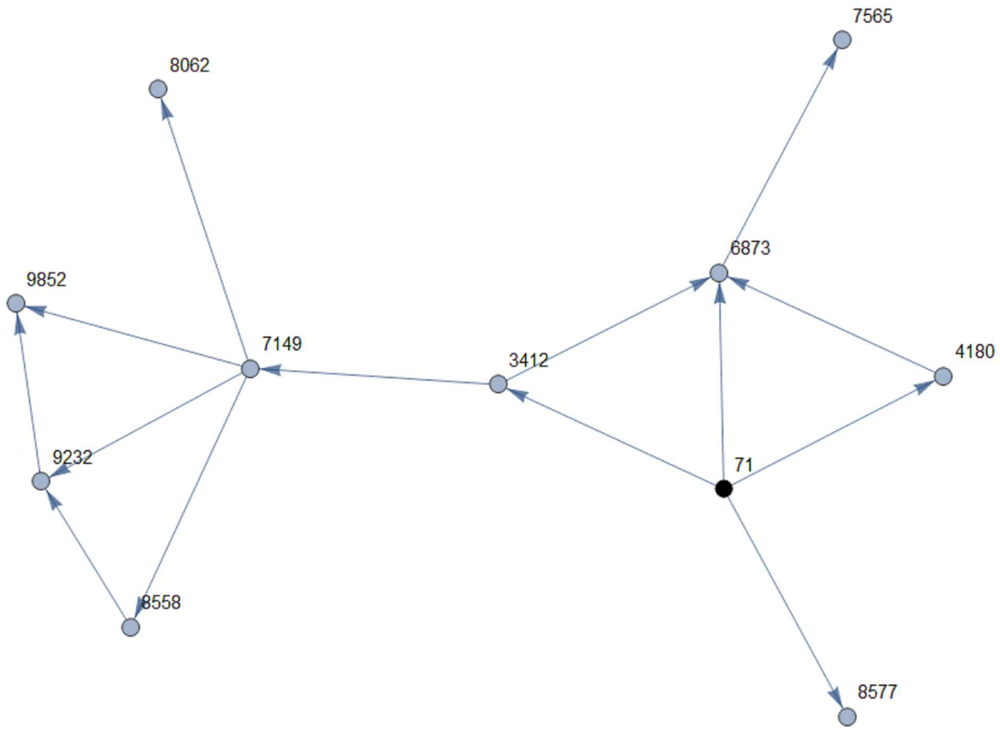

Figure 3: Two paths property (TPP): Buyer $k \in \{3, 4\}$ satisfies TPP with respect to any single seller $i \neq k$. Buyer 1 does not satisfy TPP with respect to any single seller $i \neq 1$. Buyer 2 satisfies TPP with respect to $k \in \{3, 4\}$ but not with respect to 1. Critical paths: $\delta^{1\to v} = (1, 2)$ for all $v \in \mathcal{N} \backslash \{1\}$ and $\delta^{v\to w} = (v)$ for $v \in \{3, 4\}$ and $w \in \mathcal{N}$.