Zhang, Yang and Chen, Ying-Ju (2020) *Optimal Nonlinear Pricing in Social Networks under Asymmetric Network Information.* Operations Research, 68 (3). pp. 818-833. ISSN 0030-364X.

Optimal nonlinear pricing in social networks under asymmetric network information

Yang Zhang[*] Ying-Ju Chen[†]

Forthcoming, ***Operations Research***

Abstract

We study the optimal nonlinear pricing of products and services in social networks, where customers are strategic and their consumption exhibits local externality. Customers know about their local network characteristics (which are positively affiliated across neighbors) but the selling firm only has knowledge of the global network. We develop a solution approach based on calculus of variations and positive neighbor affiliation to tackle this non-standard principal-agent problem faced by the firm. We show that the optimal pricing compromises the capitalization of the susceptibility to neighbor consumption with the motivation of one's own consumption, which gives rise to a menu of quantity premium or quantity discount. In the Erdös and Rényi graph (a special case of the social network model we use), we find that the pricing scheme does not screen network positions; consequently, the firm can offer a simple uniform price. We conduct robustness checks of our results with two way connections, where the firm-optimal consumption becomes linear in customer degree in the scale free network. Compared to linear pricing, we show that nonlinear pricing allows the firm to respond more effectively to the changes of network topology and economic factors.

Keywords: network formation, local network effects, game theory, information asymmetry, nonlinear pricing.

[*]Global Asia Institute, National University of Singapore, Singapore; email: yangzhang.buz@gmail.com.

[†]School of Business and Management & School of Engineering, The Hong Kong University of Science and Technology, Hong Kong; email: imchen@ust.hk.

1 Introduction

Social networks reflect the complications inherent to relationships among humans, and one's daily life decisions are frequently affected by the choices of her social contacts. Examples abound, ranging from video games, cell phones, movies, game consoles, to social clubs and online platform services. This state of affairs challenges a long line of research on the economics of networks, where individual activities (e.g. adoption, consumption) have global impacts on the activities of others (see Rohlfs 1974, Katz and Shapiro 1985, Farrell and Saloner 1986, and the references therein). Consider, for example, a person who engages in online games. His or her time spent on the game might be positively associated with the time that his or her friends spend on the same game, but at the same time unaffected by the play of the game from strangers. In other words, the externalities of network activities are localized within social neighborhoods.

Being aware of these localized externalites, firms shall devise their marketing, promotion, and pricing strategies accordingly. In 1990, for example, MCI Communications proposed its Friends and Family Plans, which charged customers different prices based on their number of friends. A more recent instance is that of Airbnb, which offers coupons for those who invite friends to join its platform. Facebook, Twitter, Klout, and Vocalpoint all have implemented sophisticated tools to facilitate such network-based pricing; see Bloch and Quérou (2013), Candogan et al. (2012), and Fainmesser and Galeotti (2016) for comprehensive overviews of these target marketing and pricing efforts.

Notwithstanding the advanced technologies now available, firms are still greatly challenged by network pricing, particularly because of *lack of precise network information*. On the one hand, social networks are so complicated that an individual customer usually has limited knowledge about the entire network structure (i.e. who is connected to whom). Hence, a customer may have to make her consumption decisions with only local information and some speculation about other customers' choices. This creates *information heterogeneity* among customers owing to their heterogeneous network positions. On the other hand, the firm is at an information disadvantage with respect to individual customers because it cannot directly observe their respective positions in the network. This leads to *information asymmetry* between the firm and customers. An additional obstacle of pricing arises from *social comparison*. Even for the firms facing no information asymmetry (such as Facebook or Twitter, who know well enough about the connectivity of their users), it might be still difficult to price directly based on the customer's network position, because the customer may easily compare the price she got with that offered to her friend, and rejects any price that she thinks is "unfair".

The presence of these two hurdles, information asymmetry and customers' counteractions, partially explains why in the era of modern Internet, we still see seemingly coarse pricing schemes for products whose local externalities are widely documented. In the aforementioned examples, movies are predominantly sold via fixed pricing or volume discounts; e.g., Goodrich quality theaters in the United States offer their movie tickets at a discount through VIP cards, but the benefit can be redeemed with a minimum order of 50 tickets.[1] For mobile carriers like China Mobile in Hong Kong, although there may be some sophisticated combinations between cellphone pricing and data usages, the major differentiator among these plans remains volume-based: e.g., 6GB vs 10GB/ month for data usage, 2,000 vs 3,700 minutes of local calls, and etc.[2] These practical plans strike us as firms should definitely be aware of how crucial the local network structure is to customers who purchase movie tickets or use cellphones, and yet it remains a nomenclature that they, unwillingly maybe, abandon more sophisticated offers.

This paper acknowledges the above issues and investigates the firm's optimal *nonlinear pricing* problem when customers are embedded in a network and possess private information about their network positions. The pricing scheme is nonlinear in that the average unit price for a customer is not constant, depending on her total consumption. Nonlinear pricing emerges as the solution because in the presence of information asymmetry, the firm can offer a *menu* that couples quantity with price, from which customers choose by themselves. Through self-selection, a customer's choice then reveals her network position. Nonlinear pricing also bypasses the issues of social comparison because the *discrimination is implicit*: the same quantity-price menu is offered to all customers. At the same time, as the means to implement second-degree price discrimination, nonlinear pricing is superior to a fixed price offer because it solicits customers' private information more effectively.

We consider a model of social network that evolves via the framework put forth by Jackson and Rogers (2007). The stochastic network formation process describes how each new node identifies extant nodes at random and further searches the neighborhoods of these nodes, and admits a tractable mean field approximation of the degree distribution. This framework highlights the *correlation across neighbor degrees*, and entails several notable special cases, including the graphs studied by Callaway et al. (2001) and Erdös and Rényi (1960), a variant of the preferential attachment model of Barabási and Albert (1999), and the model studied by Price (1976). We build up our model upon the network topology generated by Jackson and Rogers (2007) for its generality.[3]

[1] See the details via https://www.goodrichqualitytheaters.com/vip-cards.

[2] See details via
http://www.hk.chinamobile.com/en/corporate_information/Service_Plan s/Supreme_Series/SupremeGlobal_ServicePlan.html and http://www.hk.chinamobile.com/en/corporate_information/Service_Plans/4.5G_Service_Plan/4Glocal_serviceplan.html.

[3] It should be highlighted, though, that our model only uses the degree distribution and neighbor degree distribution generated by the framework of Jackson and Rogers (2007); the specific network formation process in Jackson and Rogers (2007) does not enter our model. This point will be further clarified as we proceed with the model.

To describe the network game, we adopt the elegant linear-quadratic formulation presented by Ballester et al. (2006) where each customer's utility depends on her own consumption and that of her neighbors. The customer infers neighbors' consumption from their degrees (i.e. their numbers of neighbors); yet the customer does not know the degree of her neighbor except for information on its distribution as determined by the stochastic network formation process. This imposes the uncertainty in the neighbor consumption faced by the focal customer. In our paper we consider two alternative interpretations of how the payoff externalities are generated through the network formation: In the *out-neighbor* model, a customer's payoff is influenced by her out-neighbors' consumption levels. Whereas in the *in-neighbor* model, it is the in-neighbors' consumptions that influence an individual customer's payoff.

When the firm adopts nonlinear pricing to serve customers with heterogeneous and unobservable degrees, maximizing its profit resembles a principal–agent problem, in which customers' degrees become their private types. These types are correlated among agents in our setting, since one's neighbor's degree distribution is conditional on one's own degree in the network. In order to obtain the type distribution, we advance the results of Jackson and Rogers (2007) by providing the in-degree distributions of both the out-neighbors and the in-neighbors of a player of any given degree in the network. These results may prove useful in their own right.

For both the out-neighbor and in-neighbor models, we find that the nonlinear pricing is driven by a balance of rent extraction and sales promotion – On the one hand, the firm tends to price higher for higher degree customers to profit from their increased susceptibility to peer consumption; on the other hand, it also wants to discount the price in order to encourage higher degree customers' own consumptions. If the former dominates the latter, the optimal menu is characterized by a quantity premium. Inversely, it will exhibit a quantity discount. We also compare our results with those of linear pricing as in Fainmesser and Galeotti (2016), and demonstrate the advantage of nonlinear pricing in responding to the changes of network topology and economic factors.

The empirical literature on social networks has widely documented positive affiliations among neighbor degrees (also referred as assortativity). See Barclay et al. (2014), Newman and Park (2003), and the discussions and references in Jackson and Rogers (2007). In our model, this positive neighbor affiliation is crucial in two aspects. On the one hand, it makes the players' payoff functions satisfy the Spence-Mirrlees single-crossing condition, which allows us to apply standard approach to reduce the incentive constraints in mechanism design. On the other hand, positive neighbor affiliation is *not* featured in the networks based on which extant pricing models are drawn (under incomplete network information – Fainmesser and Galeotti

2016[4]), and thereby sets our paper apart from the literature.

Applying our results of out-neighbor model to the *Erdös and Rényi graph* (Callaway et al., 2001; Erdös and Rényi, 1960), we find that the pricing scheme should not screen network positions. As a consequence, the firm can optimally offer a simple *uniform* pricing scheme, under which every customer is induced to choose the same consumption level in equilibrium. This result echoes some prior work on social network pricing under complete information (such as Bloch and Quérou (2013) and Candogan et al. (2012)), which emphasizes the prevalence of uniform pricing (e.g., when the interaction matrix is symmetric). Our analysis reveals that even if the firm can implement more flexible (nonlinear) pricing schemes, in a Erdös and Rényi graph uniform pricing emerges as the optimal choice.

We robustness-checked our findings with two-way connections, and find the results in this new setting are well analogous to those of out-neighbor and in-neighbor models. In particular, we show the consumption under optimal pricing is proportional to the player degree, when the connections are two-way and the underlying network is scale free (Barabási and Albert, 1999).

The rest of our paper proceeds as follows. Section 2 reviews relevant literature. Section 3 presents the basic model. In Section 4 we characterize the degree distributions of both out-neighbors and in-neighbors for a player of any given degree in the network. Sections 5 and 6 study the out-neighbor model and the in-neighbor model, respectively. We conclude in Section 7. All the proofs are relegated to the appendix.

2 Literature review

Our paper is related to the literature on network games. This research stream started with discussions on the *aggregate* level of network externalities, where a player's utility depends on the aggregate consumption of all players (e.g., Rohlfs (1974), Katz and Shapiro (1985), Farrell and Saloner (1986)). Under this assumption, a number of papers have examined the monopoly pricing problem, including Dybvig and Spatt (1983), Cabral et al. (1999), and Ochs and Park (2010), and assortment planning (Wang and Wang (2016)). Another stream of papers, including ours, explicitly take into account the network structure and study *local* network effects among players. We refer interested readers to the following surveys of the growing literature on economics of networks: Jackson (2008); Ioannides (2012); Jackson and Zenou (2015); and Jackson et al. (2016).

Our utility function is adapted from Ballester et al. (2006). It features concavity in own con-

[4] The network structure studied in Fainmesser and Galeotti (2016) exhibits correlation between one's own in-degree and out-degree, but not across neighbor degrees. See Section 5.1 for further illustration.

sumption as well as strategic complementarity between neighbor consumption, so this function is appropriate for the product adoption problem we study. The linear quadratic form of this utility function has the further advantage of making a study of the equilibrium behavior of network games more tractable, as demonstrated in Ballester et al. (2006). Similarly to our paper, those of Bloch and Quérou (2013) and Candogan et al. (2012) also build upon the framework of Ballester et al. (2006) to study the pricing problem. These two papers concentrate on linear pricing, i.e., a constant unit price for each player, while the firm can observe the network structure and so can price discriminate based on players' network positions. It is interesting that, if the network effects are symmetric, then the optimal linear pricing turns out to be independent of players' network positions. Ehsani et al. (2012) and Swapna et al. (2012) employ a dynamic pricing approach; they consider again complete network information and assume that customers are not forward looking. Leduc et al. (2017) consider a two-period setting in which customers face uncertain product qualities and can learn it from neighbor's referrals. They examine the optimal referral and intertemporal pricing plan for the selling firm. Although Leduc et al. (2017) also consider incomplete network information, the authors study a very different problem than ours, where each customer has unit demand, and both product price and referral reward are linear. As already mentioned, our paper is closely related to Fainmesser and Galeotti (2016), which focuses also on linear pricing but allows the firm to have partial information about players' in-degrees and/or out-degrees.[5] Under the linear pricing scheme, the authors show that optimal pricing exhibits price discount for influence on others' consumption and price premium for susceptibility to others' consumption.

Our paper is distinguished from the above cited work in the following aspects. First, we study the nonlinear pricing problem and assume that the firm cannot observe individual players' network positions. This gives rise to a mechanism design (principal-agent) problem. Second, we work on a network topology that incorporates positive affiliation of neighbor degrees, as introduced in Jackson and Rogers (2007). The network model we employed nests classical models (Callaway et al. 2001, Erdös and Rényi 1960, Barabási and Albert 1999, Price 1976) as limiting cases. Third, because of these differences, the pricing implications of our paper are radically different. We demonstrate that both quantity premium and quantity discount can emerge as optimal schemes, and that, in a Erdös and Rényi graph (Callaway et al., 2001; Erdös and Rényi, 1960), uniform pricing becomes optimal. Noticeably, uniform pricing has been examined in a number of contexts different than ours, including Fazeli and Jadbabaie (2012),

[5]Incidentally, Jadbabaie and Kakhbod (2016), Chen et al. (2011) and Shen and Basar (2007) study the problem of pricing network goods when the customers' intrinsic product values (Chen et al., 2011; Shen and Basar, 2007) or strength of externalities (Jadbabaie and Kakhbod, 2016) are privately learned. These papers differ from ours in that the network structure is observable by the firm.

Banerji and Dutta (2009), Armstrong (2006), Rochet and Tirole (2006), and Caillaud and Jullien (2003).

Finally, our work is to related to the literature on contracting with externalities. That research stream considers a principal's problem when facing multiple agents whose payoffs depend on other agents' decisions. In examining this multilateral contracting problem, many papers allow the principal to make an individual agent's contract contingent on other agents' reports and selections. Thus, cross-checking agents' decisions helps the principal to reduce their information rents (especially when agents' types are correlated) and sometimes facilitates more efficient allocations. This idea has been adopted in, among others, Genicot and Ray (2006), Gomes (2005), Jehiel et al. (1996), Segal (1999), and Figueroa and Skreta (2011). In our model, the firm sets individual pricing schemes under which the price charged to one player does not depend on other players' choices. We view this as a natural restriction in a large stochastic network, because firms can hardly devise complicated pricing plans that specify how the entire profile of all customers' choices determines the payment of a single customer. This is aligned with the "publicly bilateral contracting" setting studied in Segal (1999). Csorba (2008) explores the contracting problem in a setting where one's consumption has impact on the entire population. Thus we could view the network effect reflected in Csorba (2008) as "global", which differs from the local network effect we study. The paper that is closest to ours in this stream of literature is Jadbabaie and Kakhbod (2016), which studies optimal contracting with local network externalities. Specifically, Jadbabaie and Kakhbod (2016) compares the performance of multilateral contracting with that of bilateral contracting upon various network structures. Unlike our work, however, Jadbabaie and Kakhbod (2016) assumes complete knowledge of network structure by both the firm and the customers, yet private information on the strength of externalities.

3 Model

We consider a model in which the firm intends to sell products to a group of N strategic customers in a social network. The customers have incomplete information on the network topology. For expositional convenience, we will navigate through each component of our model. We shall use "customer", "node", and "player" interchangeably.

Social networks. We adopt the elegant framework by Jackson and Rogers (2007) and for completeness we describe it in detail. The network evolves in discrete time via the following process. In each time epoch t, a new node t is born. Upon birth, the node links o_r extant nodes at random (called tier-1 neighbors), and further searches the neighborhoods of these nodes and randomly links o_n of their neighbors (called tier-2 neighbors). The initial network is assumed to be structurally supportive. The linked nodes are called *out-neighbor*s to the node in question, and the number of one's out-neighbors is one's *out-degree*, which is denoted by $m := o_r + o_n$.

Conversely, if a node gets linked by a new born node as just described, the node gains an *in-neighbor*, and the number of in-neighbors is one's *in-degree*. Therefore in expectation, nodes born earlier get higher in-degree, and one's in-neighbors have lower in-degree than oneself. Hereafter, we simply refer in-degree as *degree*, when no confusion could arise. Let $r := \frac{o_r}{o_n}$ measure the randomness of the network and at the same time capture the correlation of degrees among players.[6]

This framework includes several notable special cases. For example, when $o_r \to 0, o_n = 1$, the model degenerates to the one studied in Price (1976). When $r \to 0$, it becomes a variant of the preferential attachment model introduced by Barabási and Albert (1999).[7] At the other extreme, when $r \to \infty$, the model is a variant of the graph studied by Erdös and Rényi (1960) (as presented in Callaway et al. (2001)). We view customers as nodes, and their social relationships are described by the links in this network.

Degree distributions. The above network formation process admits a tractable mean field approximation with large enough N.[8] Under mean field approximation, Jackson and Rogers (2007) show that the degree distribution of network nodes is characterized by the following cdf. $F(\cdot)$:

$$F(k) = 1 - \left\{ \frac{rm}{k + rm} \right\}^{1+r}. \tag{1}$$

The pdf. is denoted by lower case $f(k)$ (a convention that is followed for all the distribution functions in this paper). Hence, the inverse hazard rate is $H(k) := \frac{1-F(k)}{f(k)} = \frac{k+rm}{1+r}$.[9]

Information structure. We amend the above model and introduce information asymmetry. Specifically, the global network characteristics r, m are common knowledge. However, each

[6] In the original model of Jackson and Rogers (2007), the node upon birth identifies m_r tier-1 neighbors and m_n tier-2 neighbors, and then link to the identified tier-1 [-2] neighbors with probability p_r [p_n]. The linked nodes are called out-neighbors to the focal node. We adopt a simplified version of this model by assuming a deterministic linking process (which is compatible with the mean field approximation). Our version of the model produces identical degree distribution and neighbor degree distributions as does Jackson and Rogers (2007), when we normalize $o_r = p_r m_r$ and $o_n = p_n m_n$.

[7] A strict resemblance to preferential attachment requires that the nodes be attached with d_0 in-degrees upon birth. In this paper we focus on the case $d_0 = 0$ (as in most of Jackson and Rogers (2007)).

[8] We require the network size to be large enough so that the difference of using N as the upper bound in degree distributions is negligible versus using infinity. This is illustrated in the proofs.

[9] Under large enough network size, we are able to approximate individual's degree as a continuous variable. This also allows for closed form expressions for degree distributions under the mean field approach. The continuous degree is a common setup in the complex network literature, e.g. Jackson and Rogers (2007), Callaway et al. (2001), Barabási and Albert (1999), among others.

customer privately learns her degree k. In that sense, both the customers and the firm have stochastic knowledge on the global network, whereas each customer has private (deterministic) knowledge of her local connectivity.

Payoffs. We describe the game by specifying customers' payoffs from interacting with each other. For that purpose, we embed the neat payoff formulation by Ballester et al. (2006), in which the uncertainty is induced by the social network. Given their *ex ante* symmetry, the players can be labeled by their degrees rather than their individual identities. For a degree-k customer, we use $x(k)$ to denote her consumption level [10], and $P(x(k))$ to denote the corresponding transfer payment made to the firm.

The customer's utility is given as follows:

$$\pi_k(x(\cdot), P(\cdot)) = ax(k) - bx^2(k) + x(k)\delta\mathbb{E}[\sum_{j\in N_k^z} x(j)|k] - P(x(k)), \tag{2}$$

where a measures the intrinsic marginal utility and b the degree of diminishing marginal returns. The parameter δ captures the strength of product externality: a higher value corresponds to a greater dependence on other customers' choices. Here N_k^z represents the set of in-neighbors (out-neighbors) of a degree-k player, when $z = i(o)$. Thus $\mathbb{E}[\sum_{j\in N_k^z} x(j)|k]$ is the expected consumption of all in-neighbors (out-neighbors) of the degree-k player when $z = i(o)$, based on the degree-k player's local knowledge.

Given consumption $x(\cdot)$ and payment scheme $P(\cdot)$, the firm's payoff is as follows:

$$\pi_0(x(\cdot), P(\cdot)) = N\int_0^N (P(x(k)) - cx(k))f(k)dk, \tag{3}$$

where we use index 0 to denote the firm. The parameter c is the per-unit cost of providing the products. The term $P(x(k)) - cx(k)$ represents the profit from selling to each degree-k player, and the total profit is collected over the network according to the degree distribution.

As inferred from (2) and (3), our model only uses the degree distribution and neighbor degree distribution derived from Jackson and Rogers (2007), but does not rely on how the network is formed. Thus in principle, one could specify our model without recourse to the specific network formation process, which increases the robustness of our results.

Time line. The sequence of events proceeds as follows. First, the firm announces the pricing scheme $P(x)$ for any consumption level x. Second, customers observe their own degrees and simultaneously choose their consumption levels conditional on their degrees, that is, $x(k)$ for degree k. Note that the firm cannot observe the customers' degrees. Using the revelation principle, it suffices to consider the truth-telling mechanism. Thus the game can be transformed as follows. First, the firm announces an allocation $\{x(k), P(x(k))\}$ for each degree k. Second, the customers report their degrees to the firm. Third, the firm implements the allocation in

[10]The pricing scheme in our model screens the customer's in-degree only, given that one's out-degree is constant.

accordance with the customers' reports. The payoff function of a degree-k customer reporting $\hat{k}$ when others report their own types is

$$\pi_k(\hat{k}, x(\cdot), P(\cdot)) = ax(\hat{k}) - bx^2(\hat{k}) + x(\hat{k})\delta\mathbb{E}[\sum_{j\in N_k^z} x(j)|k] - P(x(\hat{k})),$$

$z \in \{i, o\}$, while the mechanism ensures truth telling is optimal for each customer.

Optimization formulation. The firm's problem that we have described can be formally stated as

$$\max_{x(\cdot),P(\cdot)} \pi_0(x(\cdot), P(\cdot)) = N\int_0^N (P(x(k)) - cx(k))f(k)dk \tag{4}$$

$$\text{s.t.} \quad \pi_k(k, x(\cdot), P(\cdot)) \geq \pi_k(\hat{k}, x(\cdot), P(\cdot)), \forall k, \hat{k} \tag{5}$$

$$\pi_k(k, x(\cdot), P(\cdot)) \geq 0, \forall k \tag{6}$$

where (5) is the incentive compatibility (IC) constraint, and (6) is the individual rationality (IR) constraint. The IC constraint (5) ensures that, for a degree-k customer, reporting her true degree (type) yields a higher expected payoff than pretending to be any other type $\hat{k}$. At the same time, the IR constraint (6) ensures that upon accepting the offer, a degree-k customer receives a non-negative payoff that is at least weakly better than not buying at all. We impose the following conditions, where $\bar{k}$ denotes the highest possible degree in the network.

Assumption 1. *$a > c, \delta\bar{k} < 2b$.*

We assume that $\delta\bar{k} < 2b$ in order to prevent an unbounded equilibrium for the consumption game that exhibits strategic complementarity. This assumption is a direct analogy of that employed in Fainmesser and Galeotti (2016), and is analogous also to Assumption 1 of Candogan et al. (2012) in the environment of incomplete network information. To see this, consider a variant game in which the customers produce the goods themselves. Obviously, the player consumption level in this variant game should be higher than that in the presence of double marginalization.

The payoff function of this variant game is

$$\bar{\pi}_k(x(\cdot)) = (a - c)x(k) - bx^2(k) + x(k)\delta\mathbb{E}[\sum_{j\in N_k^z} x(j)|k], \tag{7}$$

$z \in \{i, o\}$. The best response to given neighbor strategy, $\mathbb{E}[\sum_{j\in N_k^z} x(j)|k]$, can be written as $\bar{x}(k) = \frac{1}{2b}\{a - c + \delta\mathbb{E}[\sum_{j\in N_k^z} x(j)|k]\}$. Then $\delta\bar{k} < 2b$ implies that the effect of one's own action on payoff dominates that of an average action of one's neighbors; as a result, the best response dynamics constitute a contraction mapping and the equilibrium of the variant game will exist uniquely and be finite. That sets a finite upper bound for the equilibrium of the original game. The (omitted) proof is the same as that for part 1 of Proposition 1 in Fainmesser and Galeotti

(2016), except that their problem uses $b = 1/2$. Assumption 1 applies in the rest of our paper unless otherwise noted.

In the section that follows, we present an analysis of the degree distributions of one's neighbors in the network, which will be useful for characterizing the consumption game. Since the underlying network formation process does not determine the influence structure (i.e. who exerts influence and who is influenced), we explore two possible models in this paper. In the *out-neighbor* model in Section 5, the customers are influenced by the consumption of their out-neighbors; that is, $\mathbb{E}[\sum_{j\in N_k^z} x(j)|k]$ in (2) refers to the expected total out-neighbor consumption ($z = o$). The *in-neighbor* model analyzed in Section 6 handles the case where $\mathbb{E}[\sum_{j\in N_k^z} x(j)|k]$ represents the expected total in-neighbor consumption ($z = i$).

4 Neighbors' degree distributions and consumption

Suppose the firm adopts nonlinear pricing to serve customers with heterogeneous and unobservable in-degrees. Under nonlinear pricing the firm faces a principal-agent problem: customers are the agents, and their in-degrees are the agents' private types. The type distribution corresponds to the (unconditional) in-degree distribution (1), as characterized by Jackson and Rogers (2007). However, another critical aspect of the firm's problem in this case is the in-degree distribution of the customer's *neighbor*s. By (2), neighbors' in-degree distribution directly affects the neighbors' expected consumption, which in turn influences the individual customer's willingness to pay.

This section advances the results of Jackson and Rogers (2007) to better serve our purpose. We show that the network modelled by Jackson and Rogers exhibits the following degree distributions of a player's neighbors:

Proposition 1. *The cdf. of in-degree distribution for* out-*neighbors of an indegree-k player is as follows:*

For tier-1 out-neighbors,

$$F_k^{o1}(d) = 1 - \left(\frac{k+rm}{d+rm}\right)^{r+1}, \forall d > k. \tag{8}$$

For tier-2 out-neighbors,

$$F_k^{o2}(d) = 1 - \left(\frac{k+rm}{d+rm}\right)^{r}\left(\frac{r(d-k)}{d+rm}+1\right), \forall d > k. \tag{9}$$

The cdf. of in-degree distribution for in-*neighbors of an indegree-k player is as follows:*

$$F_k(d) = 1 - \frac{(k+rm)^{\frac{rm}{d+rm}} - rm}{k}, \forall d < k \tag{10}$$

Both distributions $F_k^{o1}(\cdot), F_k^{o2}(\cdot)$ first-order stochastically increase with m. The distribution $F_k(\cdot)$ first-order stochastically increases with r and m.

All proofs are given in the appendix. We can use (8)-(10) to verify the following remark.

Remark 1 (Positive neighbor affiliation)**.** *The network exhibits* positive out-neighbor affiliation*:* $1 - F_k^{o1}(d)$ *and* $1 - F_k^{o2}(d)$ *increase with* k *for all* $d > k$*, and* positive in-neighbor affiliation*:* $1 - F_k(d)$ *increases with* k *for all* $d < k$*.*

The concept of positive neighbor affiliation is adapted from Galeotti et al. (2010). It implies the positive correlation between neighbor degrees – a property suggested by a variety of empirical research (e.g., Barclay et al. (2014) and Newman and Park (2003)). As we will see, the presence of positive neighbor affiliation will help qualify the single-crossing condition (together with the monotonicity of $x(\cdot)$), thus laying the foundation for the pricing mechanism design.

Besides that, notice $1 - F_k^{o1}(d) < 1 - F_k^{o2}(d)\ \forall d > k$. Therefore, the in-degree of tier-2 out-neighbors first-order stochastically dominates (FOSD) that of tier-1 out-neighbors. This outcome reflects the idea that searching through neighborhoods yields neighbors of higher degree than does making random connections. Not surprisingly, the degree distributions of all three kinds of neighbors are FOSD shifted by the average degree m (Proposition 1). In addition, the in-neighbor degree distribution is positively FOSD shifted by r. The intuition here is that a low value of r corresponds to high degree correlation, which makes it difficult for late-born players to accumulate in-degrees. Hence for each degree type player, the in-degree distribution of her in-neighbors (who are born later than she was) stochastically increases with r.

Next, we build upon Proposition 1 to characterize the expected consumption of all neighbors of the degree-k player ($\mathbb{E}[\sum_{j \in N_k^z} x(j)|k]$). This critical term is imperative for the consumption equilibrium among the customers. Recall that N_k^z, with $z \in \{i, o\}$, represents (respectively) the in- and out-neighborhood of a degree-k player.

Proposition 2. *The expected sums of neighbor consumption for a degree-k customer are as follows.*

- *For* out-*neighbor consumption*

$$\mathbb{E}[\sum_{j \in N_k^o} x(j)|k] = (k + rm)^r\, rm \int_k^N \frac{x(y)}{(y + rm)^{r+1}} dy, \tag{11}$$

  which increases in k *if* $x(\cdot)$ *is increasing.*

- *For* in-*neighbor consumption*

$$\mathbb{E}[\sum_{j \in N_k^i} x(j)|k] = (k + rm)rm \int_0^k \frac{x(y)}{(y + rm)^2} dy, \tag{12}$$

  which increases in k*.*

The intuition for Proposition 2 is as follows. Consider a specific player in the network and suppose that the player's degree k increases. By Remark 1, that will lead the FOSD shift of both out- and in-neighbor degree distributions $(F_k^{o1}(\cdot), F_k^{o2}(\cdot), F_k(\cdot))$. If the consumption level $x(\cdot)$ increases with the degree, it follows that the expected consumption of a (out- and in-) neighbor to the focal player, $\mathbb{E}[x(j)|_{j\in N_k^z}k]$, increases in k ($z \in \{o, i\}$). Then the aggregate neighbor consumption, $\mathbb{E}[\sum_{j\in N_k^z} x(j)|k]$, also rises with k ($z \in \{o, i\}$). [11] This gives rise to a boost of the positive product externality and thereby increases the value of marginal consumption to the player. The mathematical implication is that the compensation rate for the indifference curve is uniformly higher under higher player degree, thereby introducing the Spence-Mirrlees single-crossing condition as required for standard mechanism design (see e.g. Fudenberg and Tirole 1991, Theorem 7.3). These results pave the way for further derivations of optimal pricing and consumption.

5 Out-neighbor model

In this section we focus on the out-neighbor model, where the customers are influenced by the consumption of their out-neighbors. We will show that, in the out-neighbor model, the firm can implement consumption equilibria both dependent and independent of agent's degree, and that the degree-dependent implementation is superior to the degree-independent one from both perspectives of the firm and the customers. Hence our focus is given to degree-dependent implementation (which involves the screening of degrees in the contract).

Assumption 2. *$r \geq 1, b > \delta m$.*

Assumption 2 is required by the out-neighbor model in this section. Recall that $r := \frac{o_r}{o_n}$ represents the correlation of degrees among players. Thus, the condition $r \geq 1$ suggests that the degrees among neighbors do not exhibit strong correlation. Notably, Assumption 2 applies to an important special case – the Erdös and Rényi graph ($r \to \infty$). The second condition in the assumption can be interpreted as a bound imposed on the strength of the payoff externality δ. This is in line with Assumption 1, although there the upper bound on δ was motivated by ensuring the existence of an equilibrium.

[11] In the case of out-neighbor consumption, we have $\mathbb{E}[\sum_{j\in N_k^o} x(j)|k] = m\mathbb{E}[x(j)|_{j\in N_k^o}k]$; therefore $\mathbb{E}[\sum_{j\in N_k^o} x(j)|k]$ rises in k as long as $\mathbb{E}[x(j)|_{j\in N_k^o}k]$ increases in k. In the in-neighbor case, $\mathbb{E}[\sum_{j\in N_k^i} x(j)|k] = k\mathbb{E}[x(j)|_{j\in N_k^i}k]$, which implies that $\mathbb{E}[x(j)|_{j\in N_k^i}k]$ inclines in k is no longer necessary for $\mathbb{E}[\sum_{j\in N_k^i} x(j)|k]$ to increase in k. In fact, simple algebra reveals that $x(\cdot)$ being increasing is not required for $\mathbb{E}[\sum_{j\in N_k^i} x(j)|k]$ to increase in k. That said however, the monotonicity constraint $x'(\cdot) > 0$ remains present in the in-neighbor model in Section 6, because it is a standard requirement for local concavity of truth reporting in the agent's problem. See, for example, step (i) in the proof of Theorem 1.

5.1 Pricing with degree-screening

In the out-neighbor model, the firm can price the goods via screening the agents' degrees. That makes the optimal agent consumption contingent on their degrees, as characterized in the following theorem. For the ease of presentation, we refer to a degree-screening payment scheme as a *DS scheme*. Let $\underline{x}^o := \frac{a-c}{2(b-\frac{\delta rm}{r+1})}$, which is shown in the proof as a lower bound of consumption (i.e. $x(k) \geq \underline{x}^o$ for all k) and is independent of degree.

Theorem 1. *In the out-neighbor model, suppose that Assumptions 1 and 2 hold. The optimal solution to the firm's problem (4)-(6), $\{x^*(\cdot), P^*(\cdot)\}$, is such that:*

The induced consumption $x^*(\cdot)$ solves the following equation:

$$\frac{\delta(r-1)}{2(b+br-\delta rm)}\mathbb{E}[\sum_{j\in N_k^o} x^*(j)|k] + x^*(k) - (k+rm)x^{*\prime}(k) - \underline{x}^o = 0 \tag{13}$$

The corresponding payment scheme is

$$\begin{aligned} P^*(x^*(k)) = ax^*(k) - b(x^*(k))^2 + x^*(k)\delta\mathbb{E}[\textstyle\sum_{j\in N_k^o} x^*(j)|k] \\ - \textstyle\int_0^k x^*(u)\delta\frac{d}{du}\mathbb{E}[\sum_{j\in N_u^o} x^*(j)|u]du. \end{aligned} \tag{14}$$

Although the consumers are not influenced by their in-neighbors, the resulting consumptions, as revealed by Theorem 1, varies with their in-degrees. To understand this phenomenon, note when the player's degree k increases, $\mathbb{E}[\sum_{j\in N_k^o} x(j)|k]$ increases assuming increasing neighbor strategy (Proposition 2), as a result of positive neighbor affiliation (Remark 1). Given strategic complementarity, the focal player will raise his consumption $x(k)$ in response. That reinforces an equilibrium increasing in player degree. As the reader may suspect, the firm can also implement a consumption profile that is independent of agent's degree. However, as we argue in Section 5.2, the degree-dependent implementation presented in Theorem 1 is superior to the degree-independent one (c.f. Proposition 4), from both the perspective of the customers and that of the firm. Last but not least, although $P^*(\cdot)$ in Theorem 1 appears to involve one's neighbor consumption, the contract remains bilateral because the expected neighbor consumption $\mathbb{E}[\sum_{j\in N_k^o} x^*(j)|k]$ is specified in the contract according to the optimal $x^*(\cdot)$ instead of depending on the neighbors' actual inputs. This feature also resolves the coordination issue in the consumption game (by imposing a fixed belief), and ensures that the equilibrium $x^*(\cdot)$ (or truth reporting) is uniquely implemented by the scheme $P^*(\cdot)$.

To understand Theorem 1, one can uncover the marginal price from (14) as follows. [12]

$$P^{*\prime}(x^*(k)) = a - 2bx^*(k) + \delta\mathbb{E}[\sum_{j\in N_k^o} x^*(j)|k], \tag{15}$$

[12]The marginal price (15) can be obtained by differentiating (14) and applying $P^{*\prime}(x^*(k)) = \frac{dP^*(x^*(k))/dk}{x^{*\prime}(k)}$.

Observe that the marginal price $P^{*\prime}(x^*(k))$ increases in $\mathbb{E}[\sum_{j\in N_k^o} x^*(j)|k]$, while declining in $x^*(k)$. The former suggests the firm to capitalize one's susceptibility to peer consumption, while the latter characterizes the discount offered to motivate one's own consumption. When $\delta = 0$, the marginal price reduces to $P^{*\prime}(x^*(k)) = a - 2bx^*(k)$, which resembles the standard pricing result without network effect. The pricing strategy, as elaborated above, results in the consumption equilibrium characterized by (13), which can be obtained by calculus of variation (as illustrated in the next paragraph). Note that the term $\int_0^k x^*(u)\delta \frac{d}{du}\mathbb{E}[\sum_{j\in N_u^o} x^*(j)|u]du$ captures the information rent that a degree-k agent reserves in equilibrium, and $ax^*(k) - b(x^*(k))^2 + x^*(k)\delta\mathbb{E}[\sum_{j\in N_k^o} x^*(j)|k]$ corresponds to the maximal revenue associated with the consumption level $x^*(k)$. Thus the optimal menu (14) reflects that, while the firm attempts to maximize the revenue, it must subject itself to the customer's manipulation enabled by the private network information.

Because the payoff externalities render our problem nonstandard in mechanism design, we shall explicate our solution approach to Theorem 1. First, we show that the positive neighbor affiliation joint with the consumption monotonicity leads the player payoff to satisfy the single-crossing condition. Hence we can retain the classical approach to focus on the local incentive compatibility constraints, given that the single-crossing property will implicitly ensure the global incentive compatibility. That enables us to rewrite the firm's objective as a function of product consumption only:

$$\max_{x(\cdot)} \ \pi_0(x(\cdot)) = N\int_{\underline{k}}^{N} \left[\begin{array}{c} (a-c)x(k) - bx^2(k) \\ +x(k)\delta \left\{ \begin{array}{c} \mathbb{E}[\sum_{j\in N_k^o} x(j)|k] \\ -H(k)\frac{d}{dk}\mathbb{E}[\sum_{j\in N_k^o} x(j)|k] \end{array} \right\} \end{array} \right] f(k)dk. \tag{16}$$

The term $H(k)\frac{d}{dk}\mathbb{E}[\sum_{j\in N_k^o} x(j)|k]$ parallels the classical distortion that arises from information asymmetry. Here $\underline{k}$ denotes the lowest degree type of players to which the firm sells.

Then, we find the transformed problem can be tackled by calculus of variations, since (16) contains $\mathbb{E}[\sum_{j\in N_k^o} x(j)|k]$ as well as some variant of its derivative (corresponding to the term associated with $x(k)$).[13] Finally, we show that the candidate solution from the Euler equation admits monotonicity (i.e., an individual's consumption level increasing in her degree), which in turn implies that the local incentive compatibility is satisfied. Similar solution approaches are adopted for deriving Theorems 2 (in-neighbor model) and 3 (two-way influences).

Comparative statics. Having characterized the solution for optimal nonlinear pricing (in Theorem 1), we now explore its structural properties.

[13] In basic mechanism design problems, the actions from different types of agent are separable in the principal's (transformed) objective function, which makes pointwise optimization sufficient for solving the problem. In our case, however, the decision vector $x(\cdot)$ is nonseparable and coupled through the integral term that represents expected neighbor consumption. That necessitates the solution approach by calculus of variation.

Proposition 3. *In the out-neighbor model, and under the optimal nonlinear pricing scheme $P^*(\cdot)$ with induced consumption $x^*(\cdot)$, the following statements hold.*

- *If* $\frac{b}{b+br-\delta rm} > \frac{r}{r-1}$,
  - *the optimal payment scheme charges a lower marginal price per unit of goods for higher-degree customers;*
  - *the optimal payment scheme exhibits quantity discount.*
- *If*

$$\begin{aligned} 2b(r-1)\left[1-\frac{b}{b+br-\delta rm}\right]-\delta rm &> 0, \\ -\delta^2 r^2 m^2 (r-1) - 2b^2(r+1)^2 + b\delta rm(r^2+3) &> 0, \end{aligned}$$

  - *the optimal payment scheme charges a higher marginal price per unit of goods for higher-degree customers.*
  - *the optimal payment scheme exhibits quantity premium.*

When the customer degree increases, the firm faces the tradeoff between raising the price to extract the rent of customers attributed to increasing peer consumption, versus lowering the price to offset the increasing disutility from the customer's own consumption. Given higher degree types consume more, if the former strategy dominates the latter, the pricing scheme will exhibit quantity premium. In opposite, it will feature a quantity discount. To be specific, recall the optimal marginal price

$$P^{*\prime}(x^*(k)) = a - 2bx^*(k) + \delta \mathbb{E}[\sum_{j\in N_k^o} x^*(j)|k], \tag{17}$$

where the term $-2bx^*(k)$ captures the decline in marginal utility in the usage of the product, and $\delta\mathbb{E}[\sum_{j\in N_k^o} x^*(j)|k]$ the social utility derived from neighbor usage of the product. Then we have

$$\begin{aligned} P''(x^*(k)) &= \frac{\mathrm{d}P'(x^*(k))}{\mathrm{d}k}/x^{*\prime}(k) \\ &= -2b + \delta\frac{\mathrm{d}}{\mathrm{d}k}\mathbb{E}[\sum_{j\in N_k^o} x^*(j)|k]/x^{*\prime}(k). \end{aligned} \tag{18}$$

Thus, whether the optimal contract exhibits quantity premium or quantity discount ($P''(x^*(k)) > 0$ or < 0) depends on whether the increment of return from one's local network ($\delta\frac{\mathrm{d}}{\mathrm{d}k}\mathbb{E}[\sum_{j\in N_k^o} x^*(j)|k]$) relative to the increment of one's own consumption ($x^{*\prime}(k)$) dominates the associated reduction of marginal standalone utility ($2b$), or vice versa. Further algebra (substituting (52) and (13) to remove $x^{*\prime}(k)$) leads to the sufficient conditions for quantity-discount and -premium,

$\frac{b}{b+br-\delta rm} > \frac{r}{r-1}$ and $2b(r-1)\left[1-\frac{b}{b+br-\delta rm}\right]-\delta rm > 0$ & $-\delta^2 r^2 m^2(r-1)-2b^2(r+1)^2+b\delta rm(r^2+3)>0$, respectively given in Proposition 3.

When the firm is informed of the customer degrees and uses linear pricing, Fainmesser and Galeotti (2016) show that it is optimal for the firm to offer price discount [price premium] with respect to the customer's in-degree [out-degree]. In doing so, the firm makes profit by facilitating one's influence on [exploiting one's susceptibility to] others' consumption. When no such topological information is available, the firm is not able to discriminate with linear pricing; and as suggested in Fainmesser and Galeotti (2016), uniform pricing becomes optimal. However as shown in our paper, allowing for nonlinear pricing in this case leads to substantially different strategy recommendations. The firm could impose a price premium or price discount with respect to the self selected quantity by the customer, which conveys her private knowledge on ego-network. Hence the firm is able to conduct second-degree price discrimination when the consumer network position is unobservable. In the out-neighbor model we currently study, both quantity premium and quantity discount can arise at optimum. Since higher-degree customers choose higher consumption levels in equilibrium, a quantity premium implies that the firm discriminates against higher-degree customers, whereas a quantity discount implies that the firm favors them.

We work on a network architecture that exhibits positive neighbor affiliation, which reflects an important regularity (assortativity) in real social and economic networks (e.g. Barclay et al. 2014; Newman and Park 2003). Positive neighbor affiliation qualifies the single crossing condition for the mechanism design, while it also makes our network structure fundamentally different from the one examined in Fainmesser and Galeotti (2016). To be specific, positive neighbor affiliation in our model regards the correlation across *neighbors'* in-degrees. On the contrary, Fainmesser and Galeotti (2016) refers to the correlation between *one's own* in-degree and out-degree (in their Section 5). As such, positive neighbor affiliation is not featured in Fainmesser and Galeotti (2016). Therefore in summary, our setup differs from that of Fainmesser and Galeotti in that, 1) we allow for *nonlinear* pricing, and we contrast with the results of linear pricing in details in Appendix B; 2) we adopt a network topology that embraces *neighbor* degree correlation.

We now conduct some comparative statics regarding the consumption equilibrium and the firm's profit.

Corollary 1. *In the out-neighbor model:*

- *the firm reaps more profit from higher-degree customers;*
- *the induced consumption level $x^*(\cdot)$ is downward distorted from the first-best consumption;*
- *$x^*(\cdot)$ increases with customer degree k, and increases with the average degree m;*

- *the firm's profit at optimum also increases with m.*

These results are intuitive and affirms that our modeling and analysis are appropriate. The first observation from Corollary 1 indicates that the firm finds higher-degree customers to be more profitable. The second observation states that asymmetric network information leads consumption to become distorted, as each customer unambiguously lowers her consumption in equilibrium in comparison to when the firm has complete information on customer degrees (referred as *first-best* scenario). We also find that a more connected network leads to higher consumption levels and greater firm profit at the optimum.

5.2 Pricing without degree-screening

In the current section we assume players are influenced by their out-neighbors, whereas it is the number of in-neighbors that varies across players. As such, one may wonder whether it is necessary for the pricing scheme to screen players' in-degrees in the out-neighbor model. In this section we will show that, a non-degree-screening scheme (called a *NDS scheme*) will necessarily degenerate to uniform pricing and therefore, cannot improve the firm's profit.

Proposition 4. *The optimal NDS scheme in the out-neighbor model*

- *induces the consumption level* $x^* = \frac{a-c}{2(b-\delta m)}$ *for every customer independent of his or her degree, where* x^* *increases with* m*;*
- *charges everyone a payment of* $P^* = \frac{a^2-c^2}{4(b-\delta m)}$*, which increases with* m *;*
- *generates the maximum profit of* $N\frac{(a-c)^2}{4(b-\delta m)}$ *for the firm, which increases with* m *and is lower than the profit earned under optimal DS scheme;*
- *leaves every customer with no information rent.*

In Proposition 4, r does not affect the induced consumption x^*, because the equilibrium is independent of player degree and thereby not affected by how the degrees are correlated. m enters the expression of x^*, due to the fact that one's number of out-neighbors matters for her consumption in the out-neighbor model. We see that x^* is increasing in m. In other words, if the average degree is higher, then the induced consumption level is also higher for each individual customer. It follows that both the equilibrium price and firm's profit are increasing in m.

Unlike the DS contract analyzed in Section 5.1, NDS contract extracts all the rent of the customers at optimum. This happens because the customers do not utilize their private information in making the consumption. In contrast, the customers are able to seize a positive information rent under degree-screening scheme (Section 5.1). Therefore, the customers would favor the degree-screening scheme over the non-degree-screening one. For the firm, the inefficiency of NDS pricing lies in the fact that it cannot solicit the social network information, which matters for the agent's incentive of consumption. Consequently, the performance of NDS pricing is dominated by that of DS pricing. The firm would thus prefer the DS pricing to NDS pricing, too.

Erdös and Rényi graph. Although the NDS scheme is in general disfavored by both the firm and the customers, it is relevant for pricing in an important special case of our model. In our framework, if $r \to \infty$ then the limiting case becomes a variant of the graph studied by Erdös and Rényi (1960) (see Callaway et al. (2001)). It turns out that, NDS scheme performs equally well as DS scheme in the Erdös and Rényi graph. To understand this result, note that in the Erdös and Rényi network everyone is influenced by a same number of uniformly randomly sampled out-neighbors. [14] That means customers of different degree types should have the same level of susceptibility to consumption. Therefore they are induced to consume independent of their degrees by the pricing mechanism. As a consequence, the pricing results with and without screening customer degrees coincide in the Erdös and Rényi graph.

Proposition 5. *In the out-neighbor model, the NDS pricing specified in Proposition 4 is optimal for the Erdös and Rényi graph.*

As suggested by Proposition 5, uniform pricing emerges as the optimal choice in the Erdös and Rényi graph, even when the firm could implement more sophisticated pricing schemes which involve the screening of the network structure. This result echoes some prior work on social network pricing, in which the optimality of uniform pricing is driven by the symmetry of the interaction matrix under complete information (see Corollary 1 of Candogan et al. 2012, as well as Bloch and Quérou 2013), or by symmetry of the degree prior under incomplete network information (Fainmesser and Galeotti 2016, Proposition 2). Also shown in Proposition 5, customers' information rent is completely reaped by the firm; that indicates higher connectivity of the network benefits the firm but not the customers in an Erdös and Rényi graph.

6 In-neighbor model

This section discusses our model's alternative specification, where customers are influenced by their in-neighbors. We term it as the *in-neighbor model*. The firm's problem is unchanged from (4)-(6) except that now $\mathbb{E}[\sum_{j \in N_k^i} x(j)|k]$ represents the expected total neighbor consumption. The following assumption applies in the present section.

Assumption 3. $r < 1$.

[14]This stems from the fact that Erdös and Rényi network is formed by uniformly random connections (i.e. pure tier-1 linking) since $r \to \infty$.

Assumption 3 means that the degrees have to be enough correlated for the in-neighbor model to well behave. The essence of this assumption is to limit one's in-neighbor degree distribution (which, recall by Proposition 1, stochastically increases in r), so that one's own consumption remains higher than that of her in-neighbor. This enforces the desired monotonicity of consumption in degree, given that one's in-neighbor's degree is lower than that of oneself. Note that with respect to r, Assumption 3 and Assumption 2 are mutually exclusive. In a flavor similar to the out-neighbor model, the next theorem shows that the consumption level induced at optimality can be characterized by a simple differential equation. For the ease of presentation, let $\underline{x}^i := \frac{a-c}{2\left(b+\delta\frac{rm}{1+r}\right)}$, which is shown in the proof as a lower bound of consumption (i.e. $x(k) \geq \underline{x}^i$ for all k) and is independent of degree.

Theorem 2. *Suppose that Assumptions 1 and 3 hold for the in-neighbor model. The optimal solution to the firm's problem (4)-(6), $\{x^*(\cdot), P^*(\cdot)\}$, is such that:*

The induced consumption $x^*(\cdot)$ solves the following equation:

$$\frac{\delta(1-r)r}{2(b+br+\delta rm)}\mathbb{E}[\sum_{j\in N_k^i} x^*(j)|k] + r(x^*(k) - \underline{x}^i) - (k+rm)x^{*\prime}(k) = 0. \tag{19}$$

The corresponding payment scheme is

$$P^*(x^*(k)) = ax^*(k) - b(x^*(k))^2 + x^*(k)\delta\mathbb{E}[\sum_{j\in N_k^i} x^*(j)|k] - \int_0^k x^*(u)\delta\frac{d}{du}\mathbb{E}[\sum_{j\in N_u^i} x^*(j)|u]du. \tag{20}$$

In the in-neighbor model in this section (and the two-way influence model in Appendix A), it is clear that the optimal contract will screen the agents' in-degrees because one's consumption is influenced by his in-neighbors. The optimal marginal price for the firm, as implied from Theorem 2, is

$$P^{*\prime}(x^*(k)) = a - 2bx^*(k) + \delta\mathbb{E}[\sum_{j\in N_k^i} x^*(j)|k]. \tag{21}$$

Therefore, the firm also faces the tradeoff between capitalizing one's willingness-to-pay induced by neighbor consumption, and incentivizing one's own consumption in presence of declining marginal return. The optimal consumption (19) can be too worked out by calculus of variation. We shall now build on Theorem 2 to examine some structural properties of the pricing menu.

Proposition 6. *In the in-neighbor model, and under the optimal nonlinear pricing scheme $P^*(\cdot)$ with induced consumption $x^*(\cdot)$:*

- *the firm reaps more profit from higher-degree customers;*
- *the optimal payment scheme charges a higher marginal price per unit of goods for higher-degree customers;*

- *the optimal payment scheme exhibits quantity premium.*

In the in-neighbor model, higher-degree customers are more profitable from the firm's standpoint – but in a way that differs from the out-neighbor model. As shown in Proposition 6, the firm sets the price higher for higher degree customers to exploit their surplus accumulated from neighbor consumption. In other words, as the player degree increases, the tendency of raising the price to extract the rent from increased social susceptibility unambiguously outweighs that of discounting to stimulate the customer's self consumption. That shapes the quantity premium in the optimal payment scheme.

We then describe the comparative statics of the consumption equilibrium. This exercise shows, inter alia, the presence of information asymmetry reduces the sales to customers of each degree type.

Corollary 2. *In the in-neighbor model, the induced consumption $x^*(\cdot)$ is downward distorted from the first-best consumption. Moreover, $x^*(\cdot)$ increases with customer degree k.*

7 Conclusion and discussion

In this paper, we investigate the optimal nonlinear pricing of products and services in social networks, where customers are strategic and their consumptions exhibit local externality. Our model features information asymmetry: customers know about their local network characteristics (which are positively correlated across neighbors), while the selling firm has only aggregate network information. The firm may adopt nonlinear pricing to serve customers with heterogeneous and unobservable network positions. The firm's profit maximization resembles a principal-agent problem, complicated with the consumption externalities.

We consider two configurations of product externalities: the out-neighbor model and the in-neighbor model, both of which are solved by an approach based on calculus of variations. We show that the optimal pricing balances the extraction of rent due to neighbor consumption with the incentivization of one's own consumption. This can give rise to either a quantity premium or a quantity discount menu. Applying the results to Erdös and Rényi graphs, we show that the optimal scheme does not screen network positions, thereby offering a uniform price for all customers. As robustness checks, we extend the model to a setting of two-way connections (Appendix A), where we find the results are analogous to those in the main models. In addition, we compare our results to those of linear pricing as in Fainmesser and Galeotti (2016), and demonstrate the advantage of nonlinear pricing in responding to the changes of topological and economic factors (Appendix B).

As implied from Galeotti et al. (2010), the warranty of single crossing condition in the network game extends to the more general setting where the utility function features strategic complementarity between neighbor consumptions while the network displays positive neighbor affiliation (joint with the increasing-ness of consumption in degree). While that may suggest a possibility of generalizing our present model, there exists a few technical challenges in doing so: The derivation of optimal menu requires specific forms of degree distribution and neighbor degree distribution.[15] The qualification of a candidate solution (i.e. nonnegativity and monotonicity) and the subsequent comparative statics also rely on a specific form of utility function. That said, we focused on a parameterized family of network configurations based on Jackson and Rogers (2007), because of its generality (covering a full topological spectrum ranged from Erdös and Rényi graphs to scale free networks) and reality (empowered to fit real-life social and economic networks as explained in their Section III).

When degrees are correlated in the network, there is correlation among players' private information. Thus, a multi-lateral contract could be possibly designed to exploit players' surplus as indicated by Cremer and McLean (1988), McAfee and Reny (1992), and Riordan and Sappington (1988), among others. However, the implementation of such contracts requires the principal to cross-check the reports by (potentially) all players to detect an individual's deviation. This can be daunting when the network size is large. A future research could focus on some partial implementation of multi-lateral contracting. For example, a firm could employ a group buying mechanism to encourage referrals and coordinated purchases from neighbors, and determine the individual payment according to the joint purchase outcome (Jing and Xie, 2011). See Leduc et al. (2017) and Lobel et al. (2016) for the studies of strategic referral programs in the network context.

As another direction of future research, one might incorporate network formation process so as to consider dynamic pricing in transient states of the network. A firm in this setup may, over time, learn the network structure from customers' responses; simultaneously, individual customers could use the firm's price adjustments to infer the network's structure outside their neighborhoods. This dynamic setup is intriguing and challenging to handle, and it remains a research priority.

[15]While there seem to be opportunities to develop a model by directly making assumptions on the general forms of degree distribution and conditional neighbor degree distributions, it is generally hard to justify these assumptions in terms of topological consistency (i.e. whether these assumptions can be met in a well defined stochastic graph). This issue does not exist with our setup, as the network formation process employed in our model specifies how the graph is generated in a consistent manner.

Acknowledgement

The authors have benefited from the discussions with Ozan Candogan, Jiangtao Li, Peng Sun, Junjie Zhou, and the seminar/conference participants in UNIST, HKUST, POMS-HK 2017, AMES 2017, INFORMS 2016, POMS 2016, National University of Singapore and University of Science & Technology of China. This work was supported by National Natural Science Foundation of China (NSFC) under Grant no. 71501108, Beijing Natural Science Foundation under Grant no. 9164030, and NSFC Grant no. 91646118. Yang Zhang is grateful for the research assistance provided by Naixin Zhang. All remaining errors are our own.

References

Armstrong, M. (2006). Competition in two-sided markets. *The RAND Journal of Economics 37*(3), 668–691.

Ballester, C., A. Calvó-Armengol, and Y. Zenou (2006). Who's who in networks. wanted: the key player. *Econometrica 74*(5), 1403–1417.

Banerji, A. and B. Dutta (2009). Local network externalities and market segmentation. *International Journal of Industrial Organization 27*(5), 605–614.

Barabási, A.-L. and R. Albert (1999). Emergence of scaling in random networks. *Science 286*(5439), 509–512.

Barclay, V. C., T. Smieszek, J. He, G. Cao, J. J. Rainey, H. Gao, A. Uzicanin, and M. Salathé (2014). Positive network assortativity of influenza vaccination at a high school: implications for outbreak risk and herd immunity. *PloS one 9*(2), e87042.

Bloch, F. and N. Quérou (2013). Pricing in social networks. *Games and Economic Behavior 80*, 243–261.

Cabral, L., D. J. Salant, and G. A. Woroch (1999). Monopoly pricing with network externalities. *International Journal of Industrial Organization 17*(2), 199–214.

Caillaud, B. and B. Jullien (2003). Chicken & egg: Competition among intermediation service providers. *RAND Journal of Economics 34*(2), 309–328.

Callaway, D. S., J. E. Hopcroft, J. M. Kleinberg, M. E. Newman, and S. H. Strogatz (2001). Are randomly grown graphs really random? *Physical Review E 64*(4), 041902.

Candogan, O., K. Bimpikis, and A. Ozdaglar (2012). Optimal pricing in networks with externalities. *Operations Research 60*(4), 883–905.

Chen, W., P. Lu, X. Sun, B. Tang, Y. Wang, and Z. A. Zhu (2011). Optimal pricing in social networks with incomplete information. In *Internet and Network Economics: 7th International Workshop, WINE 2011, Singapore, December 11-14, 2011, Proceedings*, Volume 7090, pp. 49–60. Springer Science & Business Media.

Cremer, J. and R. P. McLean (1988). Full extraction of the surplus in Bayesian and dominant strategy auctions. *Econometrica 56*, 1247–1257.

Csorba, G. (2008). Screening contracts in the presence of positive network effects. *International Journal of Industrial Organization 26*, 213–226.

Dybvig, P. H. and C. S. Spatt (1983). Adoption externalities as public goods. *Journal of Public Economics 20*(2), 231–247.

Ehsani, S., M. Ghodsi, A. Khajenezhad, H. Mahini, and A. Nikzad (2012). Optimal online pricing with network externalities. *Information Processing Letters 112*(4), 118–123.

Erdös, P. and A. Rényi (1960). On the evolution of random graphs. *Publication of the Mathematical Institute of the Hungarian Academy of Sciences 5*, 17–61.

Fainmesser, I. P. and A. Galeotti (2016). Pricing network effects. *The Review of Economic Studies 83*(1), 165–198.

Farrell, J. and G. Saloner (1986). Installed base and compatibility: Innovation, product preannouncement, and predation. *American Economic Review 76*, 940–955.

Fazeli, A. and A. Jadbabaie (2012). Duopoly pricing game in networks with local coordination effects. Working paper, University of Pennsylvania.

Figueroa, N. and V. Skreta (2011). Optimal allocation mechanisms with single-dimensional private information. *Review of Economic Design 15*(3), 213–243.

Fudenberg, D. and J. Tirole (1991). *Game Theory*. Cambridge, Massachussetts: The MIT Press.

Galeotti, A., S. Goyal, M. O. Jackson, F. Vega-Redondo, and L. Yariv (2010). Network games. *The Review of Economic Studies 77*(1), 218–244.

Genicot, G. and D. Ray (2006). Contracts and externalities: How things fall apart. *Journal of Economic Theory 131*(1), 71–100.

Gomes, A. (2005). Multilateral contracting with externalities. *Econometrica 73*(4), 1329–1350.

Ioannides, Y. (2012). *From Neighborhoods to Nations: The Economics of Social Interactions*. Princeton: Princeton University Press.

Jackson, M. (2008). *Social and Economic Networks*. Princeton University Press, Princeton, NJ.

Jackson, M., B. Rogers, and Y. Zenou (2016). The economic consequences of social network structure. *Journal of Economic Literature*, forthcoming.

Jackson, M. and Y. Zenou (2015). Games on networks. In P. Young and S. Zamir (Eds.), *Handbook of Game Theory, Vol. 4*, pp. 34–61. Amsterdam: Elsevier Publisher.

Jackson, M. O. and B. W. Rogers (2007). Meeting strangers and friends of friends: How random are social networks? *American Economic Review 97*(3), 890–915.

Jadbabaie, A. and A. Kakhbod (2016). Optimal contracting in networks. Working Paper, Massachusetts Institute of Technology.

Jehiel, P., B. Moldovanu, and E. Stacchetti (1996). How (not) to sell nuclear weapons. *The American Economic Review*, 814–829.

Jing, X. and J. Xie (2011). Group buying a new mechanism for selling through social interactions. *Management Science 57*(8), 1354–1372.

Katz, M. and C. Shapiro (1985). Network externalities, competition, and compatibility. *The American Economic Review 75*(3), 424–440.

Leduc, M. V., M. O. Jackson, and R. Johari (2017). Pricing and referrals in diffusion on networks. *Games and Economic Behavior* (104), 568–594.

Lobel, I., E. Sadler, and L. R. Varshney (2016). Customer referral incentives and social media. *Management Science 63*(10), 3147–3529.

McAfee, R. P. and P. J. Reny (1992). Correlated information and mechanism design. *Econometrica 60*, 395–421.

Newman, M. E. and J. Park (2003). Why social networks are different from other types of networks. *Physical Review E 68*(3), 036122.

Ochs, J. and I.-U. Park (2010). Overcoming the coordination problem: Dynamic formation of networks. *Journal of Economic Theory 145*(2), 689–720.

Price, D. d. S. (1976). A general theory of bibliometric and other cumulative advantage processes. *Journal of the American Society for Information Science 27*(5), 292–306.

Riordan, M. H. and D. E. Sappington (1988). Optimal contracts with public ex post information. *Journal of Economic Theory 45*(1), 189–199.

Rochet, J.-C. and J. Tirole (2006). Two-sided markets: A progress report. *The RAND Journal of Economics 37*(3), 645–667.

Rohlfs, J. (1974). A theory of interdependent demand for a communications service. *The Bell Journal of Economics and Management Science 5*(1), 16–37.

Segal, I. (1999). Contracting with externalities. *The Quarterly Journal of Economics 114*(2), 337–388.

Shen, H. and T. Basar (2007). Optimal nonlinear pricing for a monopolistic network service provider with complete and incomplete information. *IEEE Journal on Selected Areas in Communications 25*(6), 1216–1223.

Swapna, B., A. Eryilmaz, and N. B. Shroff (2012). Dynamic pricing strategies for social networks in the presence of externalities. In *Information Theory and Applications Workshop (ITA), 2012*, pp. 374–381. IEEE.

Wang, R. and Z. Wang (2016). Consumer choice models with endogenous network effects. *Management Science 63*(11), 3531–3997.

Appendix for Online Publication

A Two-way Influences

The models in the main text feature separate social influences from one's out-neighbors (Section 5) and in-neighbors (Section 6). In some situations, the player can be simultaneously influenced by out-neighbors and in-neighbors. This section will explore the robustness of our results in the context of two-way social influences. We employ the famous model of scale free network developed by Barabási and Albert (1999) in this section. In this model, every moment there is a player arriving at the system, who connects to M existing players. The connection follows the rule of *preferential attachment* – such that the probability that an extant node i with degree-k_i [16] gets connected by a single link from the arriving node equals $\frac{k_i}{\sum_j k_j}$, of which the denominator is the total degree of all existing nodes at that moment. Consistent with the foregoing terminology, we refer to the neighbors that a node obtains at its arrival by preferential attachment as its *out-neighbor*s, and the arriving node as an *in-neighbor* to the node reached in this way. As shown in Barabási and Albert (1999), the player degree distribution $f(k) = 2M^2/k^3$ and $F(k) = 1 - \frac{M^2}{k^2}$ for $k \in (M, \infty)$. Thereby the mean degree $m = 2M$, and the inverse hazard rate $H(k) = \frac{k}{2}$. The problems faced by the firm and each individual customer are the same as in the main text, except that the customers are influenced by both their out- and in-neighbors.

Proposition 7. *For a degree-k player in the scale free network, the cdf. of degree distribution of her out-neighbors is*

$$F_k^o(d) = 1 - \frac{k}{d}; \ \forall d \in (k, \infty) \tag{22}$$

that of her in-neighbors is

$$F_k^i(d) = \frac{k(d-M)}{d(k-M)}, \ \forall d \in (M, k) \tag{23}$$

which first-order stochastically increases in M.

Observe that both $1 - F_k^o(d)$ and $1 - F_k^i(d)$ increase in k. Therefore, both positive out- and in-neighbor affiliations exist for the scale free network. Furthermore, $1 - F_k^i(d)$ increases in M, which means $F_k^i(\cdot)$ first-order stochastically increases in M. To depict two-way interaction, we denote by λ and $1 - \lambda$ the weight on out-neighbors' and in-neighbors' influences on one's own consumption, respectively. In the extreme case, $\lambda = 1[0]$ means that the social influence comes exclusively from one's out-neighbors [in-neighbors], which reduces to the out-neighbor [in-neighbor] model studied in Section 5 [6].

[16] including the initial M degrees

Proposition 8. *In the scale free network: For a degree-k player, the expected weighted neighbor consumption is*

$$\mathbb{E}[\sum_{j\in N_k} x(j)|k] = kM\left(\lambda\int_k^N \frac{x(d)}{d^2}\mathrm{d}d + (1-\lambda)\int_M^k \frac{x(d)}{d^2}\mathrm{d}d\right), \tag{24}$$

which increases in k if $x'(\cdot) > 0$.

Using the neighbor degree distributions laid out in Proposition 7, one can obtain the expected weighted neighbor consumption as in (24). Following similar arguments in the main text (below Proposition 2), positive neighbor affiliation together with increasing consumption renders $\mathbb{E}[\sum_{j\in N_k} x(j)|k]$ increasing in k, which warrants the single crossing condition required in mechanism design.

Theorem 3. (Analogous to Theorems 1 and 2) *Under Assumption 1, the firm-optimal consumption is linear in the degree,*

$$x^*(k) = \theta k + \frac{a-c}{2b+\delta(1-2\lambda)M}, \tag{25}$$

for $\theta > 0$, and the corresponding payment is

$$P(x^*(k)) = ax^*(k) - bx^{*2}(k) + x^*(k)\delta\mathbb{E}[\sum_{j\in N_k} x^*(j)|k] - \int_{\underline{k}}^k x^*(u)\delta\frac{d}{du}\mathbb{E}[\sum_{j\in N_u} x^*(j)|u]du. \tag{26}$$

According to Theorem 3, the optimal unit price offered for degree-k is

$$P^{*\prime}(x^*(k)) = a - 2bx^*(k) + \delta\mathbb{E}[\sum_{j\in N_k} x^*(j)|k], \tag{27}$$

which decreases in own consumption and increases in neighbor consumption. Therefore, Theorem 3 bears a similar interpretation to that of Theorems 1 and 2, namely, the compromise of raising price to exploit network externality and discounting to compensate decreasing return. Moreover, in the case of scale free network, we find that the resultant consumption is rather simple – it is proportional to the agent's degree.

Proposition 9. (Analogous to Propositions 3 and 6) *Under the optimal nonlinear pricing scheme $P^*(\cdot)$ with induced consumption $x^*(\cdot)$, the following statements hold.*

- *If $\lambda > 1/2$ and $M(1-2\lambda) + \frac{a-c}{\theta(2b+\delta(1-2\lambda)M)} < 0$,*
  - *the optimal payment scheme charges a lower marginal price per unit of goods for higher-degree customers;*
  - *the optimal payment scheme exhibits quantity discount.*
- *If $\lambda < 1/2$,*
  - *the optimal payment scheme charges a higher marginal price per unit of goods for higher-degree customers.*

  - *the optimal payment scheme exhibits quantity premium.*

When $\lambda < 1/2$, one's consumption is predominately influenced by those guys of lower degree than oneself (in-neighbors). Proposition 9 implies that the optimal contract should charge a premium for the purchase. That echoes our results in the in-neighbor model in Section 6, where quantity premium contracts emerge as optimal. When $\lambda > 1/2$, one is primarily influenced by higher degree neighbors (out-neighbors). Analogous to the result of the out-neighbor model in Section 5, Proposition 9 suggests that the firm may favor quantity discount schemes conditionally. Like that in the main models, the emergence of quantity premium or discount is leveraged by the relative strengths of susceptibility exploitation and sales promotion, and how those counter forces vary with degree types (c.f. (27)).

Corollary 3. (Analogous to Corollaries 1 and 2) *In scale free networks with two-way influences,*

- *the firm reaps more profit from higher-degree customers;*
- *the induced consumption level $x^*(\cdot)$ is downward distorted from the first-best consumption;*
- *$x^*(\cdot)$ increases with customer degree k, and increases [decreases] with the network density M if $1 - 2\lambda < [>]0$;*
- *the firm's profit at optimum increases with M if $1 - 2\lambda < 0$.*

The downward distortion and monotonicity in consumption as revealed in Corollary 3 parallel those of Corollary 1 and 2. That the firm's profit increases in the agent's degree confirms the findings of both out-neighbor model (Corollary 1) and in-neighbor model (Proposition 6). Besides, Corollary 3 also extends the existing results regarding the effect of network topology: It shows that, when the influence from out-neighbors is adequately dominant ($\lambda > 1/2$), both the equilibrium consumption and the maximum profit of the firm increase in the network density (analogous to Corollary 1 for the out-neighbor model). In contrast, the induced consumption decreases in M when $\lambda < 1/2$. To comprehend this finding, note a player of degree k is influenced by M out-neighbors and $k - M$ in-neighbors, and $\lambda < 1/2$ indicates that the in-neighbors' influence dominates. Therefore when M increases, the aggregate influence declines and so does the consumption of the degree-k player in question (by strategic complementarity).

B Difference from linear pricing

Featuring nonlinear pricing, our model is by construction different from Fainmesser and Galeotti (2016) which concentrates on linear pricing. To elaborate the differences, we apply the linear pricing framework of Fainmesser and Galeotti (2016) to our setting, and compare the results with those derived under nonlinear pricing. Denote by p the unit price that the customer pays under linear pricing. Given others play strategy $x(\cdot)$, suppose degree-k type chooses a consumption level $x_{\hat{k}}$.

$$\frac{\partial \pi_k(x_{\hat{k}}, x(\cdot), P(\cdot))}{\partial x_{\hat{k}}} = a - 2bx(\hat{k}) + \delta \mathbb{E}[\sum_{j \in N_k^z} x(j)|k] - P'(x(\hat{k})) \quad (28)$$

$$\underbrace{=}_{P(x) \equiv p\,x} a - 2bx(\hat{k}) + \delta \mathbb{E}[\sum_{j \in N_k^z} x(j)|k] - p, \quad (29)$$

$z \in \{i, o\}$. Since $\frac{\partial \pi_k(x_{\hat{k}}, x(\cdot), P(\cdot))}{\partial x_{\hat{k}}}$ decreases in $x_{\hat{k}}$ regardless of the price p, the first order condition is sufficient for the customer's optimization. Hence for any given p, the incentive compatible $x(\cdot)$ is determined by the following differential equation,

$$a - 2bx(k) + \delta \mathbb{E}[\sum_{j \in N_k^z} x(j)|k] = p, \quad (30)$$

$z \in \{i, o\}$. Comparing (30) with (15) reveals that, when committed to a constant marginal price, the firm can implement fewer equilibria of the customer game than it does when choosing different marginal prices for different purchase quantities. This explains why nonlinear pricing increases the profit.

To proceed, substitute (30) into the firm's problem to remove transfer payment:

$$\pi_0(x(\cdot), P(\cdot)) = N \int_{\underline{k}}^{N} (p - c)x(k)f(k)dk \quad (31)$$

$$= N \int_{\underline{k}}^{N} \left(a - c - 2bx(k) + \delta \mathbb{E}[\sum_{j \in N_k^z} x(j)|k] \right) x(k)f(k)dk, \quad (32)$$

for $z \in \{i, o\}$, and let

$$\Pi(k) := \left(a - c - 2bx(k) + \delta \mathbb{E}[\sum_{j \in N_k^z} x(j)|k] \right) x(k) \quad (33)$$

$$= (a - c)x(k) - 2bx^2(k) + \delta x(k) \mathbb{E}[\sum_{j \in N_k^z} x(j)|k]. \quad (34)$$

Out-neighbor model. Recall the profit from degree-k under nonlinear pricing:

$$\Pi_0(k) = \begin{array}{c} (a - c)x(k) - \left(b - \frac{\delta rm}{1+r} \right) x^2(k) \\ + x(k)\delta \left\{ \frac{1}{1+r} \mathbb{E}[\sum_{j \in N_k^o} x(j)|k] \right\} \end{array}$$

For given $x(\cdot) > 0$,

$$\Pi_0(k) - \Pi(k) = x(k) \left[\begin{array}{c} \left(b + \frac{\delta rm}{1+r} \right) x(k) \\ - \frac{r}{1+r} \delta \mathbb{E}[\sum_{j \in N_k^o} x(j)|k] \end{array} \right] \quad (35)$$

$$= x(k) \left[bx(k) - \frac{\delta r}{1+r} \left(\mathbb{E}[\sum_{j \in N_k^o} x(j)|k] - mx(k) \right) \right], \quad (36)$$

which captures the *additional incentive* for increasing consumption under nonlinear pricing, relative to that with linear pricing. When m decreases, $\mathbb{E}[\sum_{j\in N_k^o} x(j)|k] - mx(k)$ decreases (which stems from the fact that the consumption of an additional out-neighbor is greater than that of the focal player [17]). Then the bracketed multiplier of $x(k)$ increases. When δ decreases, the bracketed multiplier of $x(k)$ also rises since $\mathbb{E}[\sum_{j\in N_k^o} x(j)|k] - mx(k) > 0$. Therefore, nonlinear pricing shifts *more* emphasis to sales generation than does linear pricing, when the network becomes sparser or the social interaction becomes weaker (both hindering the consumption).

In-neighbor model. For the in-neighbor model and under nonlinear pricing, recall

$$\Pi_0(k) = \begin{array}{c} (a-c)x(k) - bx^2(k) \\ +x(k)\delta\left\{\mathbb{E}[\sum_{j\in N_k^i} x(j)|k]\left(1-\frac{1}{1+r}\right) - H(k)\left(\frac{rm}{k+rm}\right)x(k)\right\} \end{array} \tag{37}$$

$$= \begin{array}{c} (a-c)x(k) - \left(b+\frac{\delta rm}{1+r}\right)x^2(k) \\ +x(k)\delta\left(1-\frac{1}{1+r}\right)\mathbb{E}[\sum_{j\in N_k^i} x(j)|k] \end{array} \tag{38}$$

For given $x(\cdot) > 0$,

$$\Pi_0(k) - \Pi(k) = \left(b - \frac{\delta rm}{1+r}\right)x^2(k) - \frac{1}{1+r}\delta x(k)\mathbb{E}[\sum_{j\in N_k^i} x(j)|k] \tag{39}$$

$$= \left[\left(b - \frac{\delta rm}{1+r}\right)x(k) - \frac{1}{1+r}\delta\mathbb{E}[\sum_{j\in N_k^i} x(j)|k]\right]x(k) \tag{40}$$

When m decreases, $b - \frac{\delta rm}{1+r}$ increases and $\mathbb{E}[\sum_{j\in N_k^i} x(j)|k]$ decreases (by Proposition 1 and increasing $x(\cdot)$); hence the bracketed multiplier of $x(k)$ increases. So does it when δ declines. In other words, nonlinear pricing adds *more* emphasis onto sales generation than does linear pricing, when the network becomes sparser or the interaction becomes weaker (which deters consumption).

Thus in both out- and in-neighbor models, the firm is able to respond more effectively with nonlinear pricing to the changes of network topology and economic factors, than it is with linear pricing.

C Proofs

Proof of Proposition 1. Jackson and Rogers (2007) provide valuable information regarding the degree distributions, but they do not explicitly lay out the relevant neighbor degree distribu-

[17] To see, note that the expected consumption of a single out-neighbor of degree-k, $\mathbb{E}[x(j)|_{j\in N_k^o} k] = \frac{o_r}{m}\int_k^N x(y) f_k^{o1}(y)dy + \frac{o_n}{m}\int_k^N x(y) f_k^{o2}(y)dy = \frac{r}{r+1}\int_k^N x(y) f_k^{o1}(y)dy + \frac{1}{r+1}\int_k^N x(y) f_k^{o2}(y)dy$. Since both F_k^{o1} and F_k^{o2} first order stochastically increase with m (Proposition 1) and $x(\cdot)$ is increasing, $\mathbb{E}[x(j)|_{j\in N_k^o} k]$ increases in m. Therefore, $\frac{d}{dm}\mathbb{E}[\sum_{j\in N_k^o} x(j)|k] = \mathbb{E}[x(j)|_{j\in N_k^o} k] + m\frac{d}{dm}\mathbb{E}[x(j)|_{j\in N_k^o} k] > \mathbb{E}[x(j)|_{j\in N_k^o} k] > x(k)$, given the out-neighbor's degree is greater than k.

tions. We now build upon their analysis and characterize the neighbor degree distributions in detail to serve our purpose. Consistent with Jackson and Rogers (2007), let $d_i(t)$ denote the in-degree of node i at time t. Under mean field approximation, the network formation process described in Section 3 leads to the following differential equation:

$$\frac{d}{dt}d_i(t) = \frac{o_n}{tm}d_i(t) + \frac{o_r}{t}, \tag{41}$$

where the initial in-degree of nodes upon birth is denoted by d_0. To understand (41), note when $o_n = p_n m_n, o_r = p_r m_r$ it is equivalent to equation (1) in Jackson and Rogers (2007), for which one should refer to the explanation in their paper. [18]

Equation (41) pins down the in-degree function:

$$d_i(t) = (d_0 + rm)\left(\frac{t}{i}\right)^{\frac{1}{1+r}} - rm. \tag{42}$$

For completeness we keep a general value of d_0, and at times derive the explicit formulas for the case $d_0 = 0$ (as does the majority of analysis in Jackson and Rogers (2007)). Under the mean field approximation, there is a unique mapping between node identity i, time t, and node in-degree k, so that one can use two of the variables to infer the third one. Let $i_t(d)$ be the birthdate (or the identity) of the node that has degree-d at time t. (42) implies:

$$i_t(d) = t\left\{\frac{d_0 + rm}{d + rm}\right\}^{1+r}. \tag{43}$$

In-neighbors. For now, denote by $F_i^t(\cdot)$ the cdf. of the in-degree distribution of in-neighbors of node i at time t, and we will later make the expression time-invariant (i.e. removing t). Recall (Jackson and Rogers, 2007, p.911, under the proof of Theorem 4):

$$1 - F_i^t(d) = \frac{d_i(i_t(d))}{d_i(t)},$$

Substituting the expression of $d_i(t)$, we obtain

$$d_i(i_t(d)) = (d_0 + rm)\left(\frac{i_t(d)}{i}\right)^{\frac{1}{1+r}} - rm = (d_0 + rm)\left(\frac{t}{i}\right)^{\frac{1}{1+r}}\left\{\frac{d_0 + rm}{d + rm}\right\} - rm,$$

so that

$$1 - F_i^t(d) = \frac{d_i(i_t(d))}{d_i(t)} = \frac{(d_0 + rm)\left(\frac{t}{i}\right)^{\frac{1}{1+r}}\left\{\frac{d_0+rm}{d+rm}\right\} - rm}{(d_0 + rm)\left(\frac{t}{i}\right)^{\frac{1}{1+r}} - rm}, \tag{44}$$

for $d < d_i(t)$ (one's in-neighbor's in-degree must be lower than the in-degree of oneself).

[18] Same as (1) of Jackson and Rogers (2007), (41) is not an exact calculation since it ignores the potential overlapping of search ranges. However, it remains a reasonable approximation for large networks ($N >> o_r, o_n$) and o_n is small relative to $o_r m$. See footnote 20 of Jackson and Rogers (2007).

As an important step, we will transform the degree distribution to be independent of time. For this purpose, observe by (42) that for a degree-k player at time t,

$$\left(\frac{t}{i}\right)^{\frac{1}{1+r}} = \frac{k+rm}{d_0+rm}, \tag{45}$$

and hereby rewrite (44):

$$1 - F_i^t(d) = \frac{(k+rm)\frac{d_0+rm}{d+rm} - rm}{k}. \tag{46}$$

Since the distribution is now time-invariant, we denote it by $F_k(\cdot)$, indicating the in-neighbor degree distribution of a degree-k node.

When $d_0 = 0$, we obtain from (46) that:

$$1 - F_k(d) = \frac{(k+rm)\frac{rm}{d+rm} - rm}{k},$$

which was included in the presentation of Proposition 1 in the main text.

Out-neighbors. Above we have solved for the degree distribution of in-neighbors. Next we study that of out-neighbors. First note that the foregoing approach of calculating neighbor degree distribution based on the equation

$$1 - F_i^t(d) = \frac{d_i(i_t(d))}{d_i(t)}, \tag{47}$$

is no longer useful, because one's out-degree does not cumulate in time; in other words, one's out-neighbors are all formed at one shot upon arrival at the system (so that the fraction of neighbors whose degree is higher than d cannot be derived by comparing the birth dates).

To proceed, we need to analyze the formation of out-neighbors upon a *now*-degree-k player's arrival at the system. There are two types of out-neighbors obtained at that moment: tier-1 out-neighbors who are connected by the player in question by random connection, and tier-2 out-neighbors who are reached via the out-degree links from tier-1 out-neighbors. Moreover, we can infer from equation (1) of Jackson and Rogers (2007) the probability of being a tier-1 or tier-2 out-neighbor of the node conditional on the node's own degree is k at time t as follows.

$$\begin{aligned} Pr\{\text{tier-1 out-neighbor}|\text{degree} = d\} &= \frac{o_r}{i_t(k)} \\ &\underbrace{=}_{(43)} \frac{o_r}{t\left\{\frac{d_0+rm}{k+rm}\right\}^{1+r}} \\ &\underbrace{=}_{d_0=0} \frac{o_r}{t\left\{\frac{rm}{k+rm}\right\}^{1+r}} \end{aligned} \tag{48}$$

where (48) is the probability that the degree-d node is found at random by the new born node who has now degree k at time t, under mean field approximation. Note this probability is independent of d, given that the selection of tier-1 out-neighbors is random rather than degree-based.

As for the tier-2 out-neighbors,

$$\begin{aligned}
& Pr\{\text{tier-2 out-neighbor}|\text{degree} = d\} \\
= \; & \left(\frac{o_r d_j(i_t(k))}{i_t(k)}\right)\left(\frac{o_n}{o_r m}\right) \qquad (49)\\
= \; & \frac{d_j(i_t(k))}{i_t(k)}\frac{o_n}{m} \\
\underbrace{=}_{j \text{ is such that } d_j(t)=d,(43),(42)} \; & \left(\frac{\frac{d+rm}{k+rm}(d_0+rm)-rm}{t\left(\frac{d_0+rm}{k+rm}\right)^{1+r}}\right)\left(\frac{o_n}{m}\right) \\
\underbrace{=}_{d_0=0} \; & \left(\frac{\frac{d-k}{k+rm}rm}{t\left(\frac{rm}{k+rm}\right)^{1+r}}\right)\left(\frac{o_n}{m}\right) \\
= \; & \frac{ro_n(k+rm)^r}{t(rm)^{1+r}}(d-k)
\end{aligned}$$

To understand (49), one should refer to the explanation of equation (1) in Jackson and Rogers (2007), and note the changes we made to equation (1) of Jackson and Rogers (2007) to adapt it for our purpose. Here we will briefly describe the intuition: Refer to node i the node in question, who has degree-k at time t, and we trace back to his birthdate, $i_t(k)$, to investigate the likelihood that he gets another node j, who has degree-d at time t, as his tier-2 out-neighbor. Note 1) The term $\frac{o_r d_j(i_t(k))}{i_t(k)}$ is the probability that some node with a link to j, is reached by node i as tier-1 out-neighbor, so that j has the potential of being met in this way. [19] 2) The term $\frac{o_n}{o_r m}$ is then the probability that j is found, given that some in-neighbor of him has been met randomly in 1)[20]. [21]

Now, what we have obtained is the probability of connecting to a node with certain degree, yet the concept of neighbor degree distribution is about, given the connection, the probability of the connected node having certain degree. This gap can be filled by applying the Bayes'

[19]The probability that *one* link reaches an in-neighbor of j at the time $i_t(k)$ is $\frac{d_j(i_t(k))}{i_t(k)}$, given there were $i_t(k)$ players in the system and $d_j(i_t(k))$ of them were j's in-neighbors. Since there are o_r such links, the probability that any of these links reaches an in-neighbor of j is approximated by o_r multiplied by the above probability.

[20]Since some in-neighbor of hers has been linked in 1), node j must lie among the $o_r m$ nodes which node i could possibly reach as tier-2 out-neighbors. Then the probability that node j is reached by the neighborhood search with o_n links is $\frac{o_n}{o_r m}$.

[21]Same as (1) of Jackson and Rogers (2007), (49) is not an exact calculation since it ignores the potential overlapping of search ranges. However, it remains a reasonable approximation for large networks ($N >> o_r, o_n$) and o_n is small relative to $o_r m$. See footnote 20 of Jackson and Rogers (2007).

theorem:

$$\begin{aligned} f_k^q(d) &= \frac{Pr\{\text{type-}q\text{ out-neighbor}|\text{degree}=d\}f(d|d>k)}{\int_k^{\bar{k}} Pr\{\text{type-}q\text{ out-neighbor}|\text{degree}=d\}f(d|d>k)\mathrm{d}d}, q \in \{o1,o2\} \\ &= \frac{Pr\{\text{type-}q\text{ out-neighbor}|\text{degree}=d\}f(d)/(1-F(k))}{\int_k^{\bar{k}} Pr\{\text{type-}q\text{ out-neighbor}|\text{degree}=d\}f(d)/(1-F(k))\mathrm{d}d}, q \in \{o1,o2\} \\ &= \frac{Pr\{\text{type-}q\text{ out-neighbor}|\text{degree}=d\}f(d)}{\int_k^{\bar{k}} Pr\{\text{type-}q\text{ out-neighbor}|\text{degree}=d\}f(d)\mathrm{d}d}, q \in \{o1,o2\} \end{aligned} \tag{50}$$

where the condition $d > k$ is imposed upon the original degree distribution for the discussion of out-neighbors (who are born prior to the now-degree-k player in question, thus having larger in-degrees).

Note then the term $1 - F(k)$ is cancelled off. This gives us

$$\begin{aligned} f_k^{o1}(d) &\underbrace{=}_{(50)} \frac{f(d)}{\int_k^{\bar{k}} f(d)\mathrm{d}d} \\ &= \frac{f(d)}{1-F(k)} \\ &\underbrace{=}_{(1)} \frac{(1+r)(k+rm)^{1+r}}{(d+rm)^{r+2}}. \end{aligned} \tag{51}$$

(51) results from the fact that $Pr\{\text{tier-1 out-neighbor}|\text{degree}=d\}$ is independent of d, so that it cancels off from the fraction. Integrating , we have the cdf.

$$F_k^{o1}(d) = 1 - \left(\frac{k+rm}{d+rm}\right)^{r+1}, \forall d > k.$$

For tier-2 out-neighbors,

$$\begin{aligned} f_k^{o2}(d) &\underbrace{=}_{(50)} \frac{(d-k)f(d)}{\int_k^{\bar{k}}(d-k)f(d)\mathrm{d}d} \\ &\underbrace{=}_{(90)} \frac{(d-k)/(d+rm)^{r+2}}{\int_k^{\bar{k}}(d-k)/(d+rm)^{r+2}\mathrm{d}d} \\ &= \frac{r(r+1)(d-k)(k+rm)^r}{(d+rm)^{r+2}}. \end{aligned}$$

Integrating, we get the cdf.

$$F_k^{o2}(d) = 1 - \left(\frac{k+rm}{d+rm}\right)^r \left(\frac{r(d-k)}{d+rm}+1\right), \forall d > k.$$

To see that $F_k^{o1}(\cdot), F_k^{o2}(\cdot)$ first-order stochastically increase with m, note that

$$\begin{aligned} \frac{\partial}{\partial m}\{1-F_k^{o1}(\cdot)\} &= \frac{(d-k)r(r+1)}{(d+rm)^2}\left(\frac{k+rm}{d+rm}\right)^r > 0, \\ \frac{\partial}{\partial m}\{1-F_k^{o2}(\cdot)\} &= \frac{(d-k)^2r^2(r+1)}{(d+rm)^3}\left(\frac{k+rm}{d+rm}\right)^{r-1} > 0, \end{aligned}$$

given $d > k$. To see $F_k(\cdot)$ first-order stochastically increases with r and m, rewrite

$$1 - F_k(d) = \frac{(k+rm)\frac{rm}{d+rm} - rm}{k} = \left(1 - \frac{d}{d+rm}\right)\frac{k-d}{k}.$$

Then observe that $1 - F_k(d)$ increases with r and m given $d < k$. □

Proof of Proposition 2. Out-neighbors. Let us first deal with out-neighbors. Recall that for tier-1 out-neighbors,

$$F_k^{o1}(d) = 1 - \left(\frac{k+rm}{d+rm}\right)^{r+1}, \forall d > k.$$

For tier-2 out-neighbors,

$$F_k^{o2}(d) = 1 - \left(\frac{k+rm}{d+rm}\right)^{r}\left(\frac{r(d-k)}{d+rm} + 1\right), \forall d > k.$$

The probability density functions are:

$$\begin{aligned} f_k^{o1}(d) &= \frac{r+1}{d+rm}\left(\frac{k+rm}{d+rm}\right)^{r+1}, \forall d > k, \\ f_k^{o2}(d) &= \frac{r(r+1)(d-k)}{(d+rm)^2}\left(\frac{k+rm}{d+rm}\right)^{r}, \forall d > k. \end{aligned}$$

For the out-neighbors, denote the set of degree-k's tier-1 and tier-2 out-neighbor degree types by N_k^{o1} and N_k^{o2}, respectively. Then we can calculate:

$$\begin{aligned} & \mathbb{E}[\sum_{j \in N_k^o} x(j)|k] \\ = \; & \mathbb{E}[\sum_{j \in N_k^{o1}} x(j) + \sum_{j \in N_k^{o2}} x(j)|k] \\ = \; & o_r \mathbb{E}[x(j)|_{j\in N_k^{o1}} k] + o_n \mathbb{E}[x(j)|_{j \in N_k^{o2}} k] \\ = \; & o_r \int_k^N x(y) f_k^{o1}(y) dy + o_n \int_k^N x(y) f_k^{o2}(y) dy. \end{aligned}$$

where the upper bound of integration is set to be the network size, given it being large enough and thus trivializing the differences from setting it to infinity or to $N-1$. Since the consumption is only determined by in-degree, one can further combine notations and express the above as (using $r = \frac{o_r}{o_n}$):

$$\mathbb{E}[\sum_{j \in N_k^o} x(j)|k] = o_n \int_k^N x(y)\left[r f_k^{o1}(y) + f_k^{o2}(y)\right] dy.$$

Plug in the definitions of $f_k^{o1}(d)$ and $f_k^{o2}(d)$:

$$\begin{aligned} r f_k^{o1}(y) + f_k^{o2}(y) &= r\frac{(r+1)}{y+rm}\left(\frac{k+rm}{y+rm}\right)^{r+1} + \frac{r(r+1)(y-k)}{(y+rm)^2}\left(\frac{k+rm}{y+rm}\right)^{r} \\ &= \frac{r(r+1)}{y+rm}\left(\frac{k+rm}{y+rm}\right)^{r}. \end{aligned}$$

Therefore, we have

$$\begin{aligned}
& \mathbb{E}[\sum_{j\in N_k^o} x(j)|k] \\
= \ & o_n \int_k^N x(y)\frac{r(r+1)}{y+rm}\left(\frac{k+rm}{y+rm}\right)^r dy \\
= \ & o_n r(r+1)\,(k+rm)^r \int_k^N x(y)\frac{1}{(y+rm)^{r+1}}dy \\
= \ & rm\,(k+rm)^r \int_k^N \frac{x(y)}{(y+rm)^{r+1}}dy,
\end{aligned}$$

Now suppose $x(\cdot)$ is increasing. Notice

$$\begin{aligned}
& \frac{d}{dk}\mathbb{E}[\sum_{j\in N_k^o} x(j)|k] \\
= \ & rm\left\{\begin{array}{c} r\,(k+rm)^{r-1}\int_k^N x(y)\frac{1}{(y+rm)^{r+1}}dy \\ -\,(k+rm)^r\,x(k)\frac{1}{(k+rm)^{r+1}} \end{array}\right\} \\
= \ & rm\left\{\begin{array}{c} r\,(k+rm)^{r-1}\int_k^N x(y)\frac{1}{(y+rm)^{r+1}}dy \\ -\frac{x(k)}{k+rm} \end{array}\right\} \\
= \ & \frac{r}{k+rm}\left\{\begin{array}{c} \mathbb{E}[\sum_{j\in N_k^o} x(j)|k] \\ -mx(k) \end{array}\right\} \\
= \ & \frac{r}{k+rm}\mathbb{E}[\sum_{j\in N_k^o} x(j)|k] - \frac{rm}{k+rm}x(k).
\end{aligned} \tag{52}$$

Since one's out-neighbors have higher in-degree than oneself, and that the consumption is assumed increasing in in-degree, each out-neighbor should have higher consumption than does oneself. That is, $\mathbb{E}[\sum_{j\in N_k^o} x(j)|k] > mx(k)$. Elaborated in detail,

$$\begin{aligned}
\mathbb{E}[\sum_{j\in N_k^o} x(j)|k] \quad & = \quad rm\,(k+rm)^r \int_k^N x(y)\frac{1}{(y+rm)^{r+1}}dy \\
& \underbrace{>}_{x(\cdot)\text{ increasing}} \quad rm\,(k+rm)^r\,x(k)\int_k^N \frac{1}{(y+rm)^{r+1}}dy \\
& \underbrace{=}_{N\text{ large enough}} \quad mx(k)
\end{aligned}$$

Therefore, $\frac{d}{dk}\mathbb{E}[\sum_{j\in N_k^o} x(j)|k] > 0$ if $x(k)$ increases in k.

In-neighbors. Now we consider the case with in-neighbors.

$$\begin{aligned}
\mathbb{E}[\sum_{j\in N_k^i} x(j)|k] &= \sum_{j\in N_k^i} \mathbb{E}[x(j)|k] \\
&= \sum_{j\in N_k^i} \left\{x(0) + \int_0^k \frac{dx(y)}{dy}\,[1-F_k(y)]\,dy\right\} \\
&= k\left\{x(0) + \int_0^k \frac{dx(y)}{dy}\frac{rm(k-y)}{k(y+rm)}dy\right\} \\
&= kx(0) + \int_0^k \frac{dx(y)}{dy}\frac{rm(k-y)}{y+rm}dy \\
&\underbrace{=}_{\text{integration by parts}} kx(0) + 0 - kx(0) - \int_0^k x(y)d\frac{rm(k-y)}{y+rm} \\
&= (k+rm)rm\int_0^k \frac{x(y)}{(y+rm)^2}dy
\end{aligned}$$

It easily follows that $\mathbb{E}[\sum_{j\in N_k^i} x(j)|k]$ is increasing in k. □

Proof of Theorem 1. The proof is composed of three steps. Step (i): We explore the structural properties of the optimization problem. Step (ii): We rewrite the firm objective as a function of consumption only. Step (iii): Using the change of variables, we show that the objective can be optimized by calculus of variations. We also verify that the candidate solution from calculus of variations indeed satisfies the remaining constraints.

Step (i): Structural properties of the optimization problem. As the standard mechanism design approach, we first explore some properties of (4)-(6) that are essential to reducing the firm's problem.

Recall the consumer's payoff when others report the truth

$$\pi_k(\hat{k}, x(\cdot), P(\cdot)) = ax(\hat{k}) - bx^2(\hat{k}) + x(\hat{k})\delta\mathbb{E}[\sum_{j\in N_k^o} x(j)|k] - P(x(\hat{k})).$$

The first-order condition for IC constraints implies:

$$\begin{aligned}
& d\pi_k(\hat{k}, x(\cdot), P(\cdot))/d\hat{k}\Big|_{\hat{k}=k} \\
= \;& ax'(\hat{k}) - 2bx(\hat{k})x'(\hat{k}) + x'(\hat{k})\delta\mathbb{E}[\sum_{j\in N_k^o} x(j)|k] - P'(x(\hat{k}))x'(\hat{k})\Big|_{\hat{k}=k} \\
= \;& ax'(k) - 2bx(k)x'(k) + \delta x'(k)\mathbb{E}[\sum_{j\in N_k^o} x(j)|k] - P'(x(k))x'(k) \\
= \;& \left[a - 2bx(k) + \delta\mathbb{E}[\sum_{j\in N_k^o} x(j)|k] - P'(x(k))\right]x'(k) \\
= \;& 0. \qquad (53)
\end{aligned}$$

The local second-order condition for IC constraints is

$$\begin{aligned}
& \frac{d^2}{d\hat{k}^2}\pi_k(\hat{k}, x(\cdot), P(\cdot))\Big|_{\hat{k}=k} \\
= & \left[-2bx'(\hat{k}) - \frac{d}{d\hat{k}}P'(x(\hat{k}))\right] x'(\hat{k}) + \left[a - 2bx(\hat{k}) + \delta\mathbb{E}[\sum_{j\in N_k^o} x(j)|k] - P'(x(\hat{k}))\right] x''(\hat{k})\Big|_{\hat{k}=k} \\
\underbrace{=}_{(53)} & \left[-2bx'(k) - \frac{d}{dk}P'(x(k))\right] x'(k) \\
\underbrace{=}_{(53)} & \left[-\delta\frac{d}{dk}\mathbb{E}[\sum_{j\in N_k^o} x(j)|k]\right] x'(k) \\
< & \ 0 \text{ if } x'(k) > 0
\end{aligned} \tag{54}$$

Therefore the local concavity for truth reporting requires the *monotonicity condition*, $x'(k) > 0$, which in our problem also leads to $\frac{d}{dk}\mathbb{E}[\sum_{j\in N_k^o} x(j)|k] > 0$ by Proposition 2. (53) and (54) constitute the local IC condition, under which the customer does not attempt to lie locally. We will soon show that the single crossing condition, justified in Proposition 2, extends local IC to global. Define the payoff in the truth telling equilibrium:

$$V(k) := \pi_k(k, x(\cdot), P(\cdot)) = ax(k) - bx^2(k) + x(k)\delta\mathbb{E}[\sum_{j\in N_k^o} x(j)|k] - P(x(k)), \tag{55}$$

and IR constraints (6) imply that $V(k) \geq 0, \forall k$.

$$\begin{aligned}
\frac{d}{dk}V(k) &= ax'(k) - 2bx(k)x'(k) + \frac{dx(k)\delta\mathbb{E}[\sum_{j\in N_k^o} x(j)|k]}{dk} - P'(x(k))x'(k) \\
&= ax'(k) - 2bx(k)x'(k) + x'(k)\delta\mathbb{E}[\sum_{j\in N_k^o} x(j)|k] + \\
&\quad x(k)\delta\frac{d}{dk}\mathbb{E}[\sum_{j\in N_k^o} x(j)|k] - P'(x(k))x'(k) \\
&\underbrace{=}_{(53)} x(k)\delta\frac{d}{dk}\mathbb{E}[\sum_{j\in N_k^o} x(j)|k],
\end{aligned} \tag{56}$$

which is positive if $x'(k) > 0$ (which leads to $\frac{d}{dk}\mathbb{E}[\sum_{j\in N_k^o} x(j)|k] > 0$ by Proposition 2).

The above result can be reached by the envelope theorem as well, in consideration of the IC constraints:

$$\frac{d}{dk}V(k) = \frac{\partial}{\partial k}\pi_k(\hat{k}^*, x(\cdot), P(\cdot))\Big|_{\hat{k}^*=k} = x(k)\delta\frac{d}{dk}\mathbb{E}[\sum_{j\in N_k^o} x(j)|k].$$

Accordingly, (from Fundamental Theorem of Calculus)

$$V(k) = V(\underline{k}) + \int_{\underline{k}}^{k} x(u)\delta\frac{d}{du}\mathbb{E}[\sum_{j\in N_u^o} x(j)|u]du,$$

where $\underline{k}$ is the lowest degree type considered. One can write

$$
\begin{aligned}
& V(k)-\pi_k(\hat{k},x(\cdot),P(\cdot)) \\
= \quad & V(k)-\left(V(\hat{k})-x(\hat{k})\delta\mathbb{E}[\sum_{j\in N^o_{\hat{k}}}x(j)|\hat{k}]+x(\hat{k})\delta\mathbb{E}[\sum_{j\in N^o_{k}}x(j)|k]\right) \\
= \quad & V(k)-V(\hat{k})+x(\hat{k})\delta\left(\mathbb{E}[\sum_{j\in N^o_{\hat{k}}}x(j)|\hat{k}]-\mathbb{E}[\sum_{j\in N^o_{k}}x(j)|k]\right) \\
= \quad & \int_{\hat{k}}^{k}x(u)\delta\frac{d}{du}\mathbb{E}[\sum_{j\in N^o_{u}}x(j)|u]du \\
& +x(\hat{k})\delta\left(\mathbb{E}[\sum_{j\in N^o_{\hat{k}}}x(j)|\hat{k}]-\mathbb{E}[\sum_{j\in N^o_{k}}x(j)|k]\right) \\
\underbrace{=}_{\text{Integration by parts}} \quad & (x(k)-x(\hat{k}))\delta\mathbb{E}[\sum_{j\in N^o_{k}}x(j)|k]-\int_{\hat{k}}^{k}x'(u)\delta\mathbb{E}[\sum_{j\in N^o_{u}}x(j)|u]du \\
= \quad & \int_{\hat{k}}^{k}x'(u)\delta\mathbb{E}[\sum_{j\in N^o_{k}}x(j)|k]du-\int_{\hat{k}}^{k}x'(u)\delta\mathbb{E}[\sum_{j\in N^o_{u}}x(j)|u]du \\
\geq \quad & 0,
\end{aligned}
$$

If $k>\hat{k}$, the above quantity is nonnegative given $x'(\cdot)>0$ (which also means $\mathbb{E}[\sum_{j\in N^o_k}x(j)|k]$ increases in k). If $k<\hat{k}$, rewrite the above as

$$\int_{k}^{\hat{k}}x'(u)\delta\mathbb{E}[\sum_{j\in N^o_{u}}x(j)|u]du-\int_{k}^{\hat{k}}x'(u)\delta\mathbb{E}[\sum_{j\in N^o_{k}}x(j)|k]du,$$

which is also nonnegative given $x'(\cdot)>0$ (which also means $\mathbb{E}[\sum_{j\in N^o_k}x(j)|k]$ increases in k). Thus IC is achieved globally. The payment $P(x(k))$ is

$$P(x(k))=ax(k)-bx^2(k)+x(k)\delta\mathbb{E}[\sum_{j\in N^o_{k}}x(j)|k]-V(\underline{k})-\int_{\underline{k}}^{k}x(u)\delta\frac{d}{du}\mathbb{E}[\sum_{j\in N^o_{u}}x(j)|u]du. \quad (57)$$

As a result, IC and IR constraints (5)-(6) can reduce to a single monotonicity constraint $x'(\cdot)>0$.

Step (ii): Rewriting the objective. Now we return to the firm's problem. Its objective can be rewritten as follows:

$$
\begin{aligned}
& N\int_{\underline{k}}^{N}(P(x(k))-cx(k))f(k)dk \quad (58) \\
= \quad & N\int_{\underline{k}}^{N}\left\{\begin{array}{c}(a-c)x(k)-bx^2(k)+x(k)\delta\mathbb{E}[\sum_{j\in N^o_k}x(j)|k]-V(\underline{k}) \\ -\int_{\underline{k}}^{k}x(u)\delta\frac{d}{du}\mathbb{E}[\sum_{j\in N^o_u}x(j)|u]du\end{array}\right\}f(k)dk.
\end{aligned}
$$

Note for any random variable $Y\in[\underline{y},\bar{y}]$ with pdf., cdf. f,F, and any generic function $g(\cdot)$, we have:

$$\begin{aligned}\mathbb{E}[g(Y)] &= \int_{\underline{y}}^{\bar{y}} g(y)f(y)dy \\ &= g(\underline{y}) + \int_{\underline{y}}^{\bar{y}} g^{'}(y)[1-F(y)]dy.\end{aligned}$$

In that way,

$$\begin{aligned}&\int_{\underline{k}}^{N}\left\{\int_{\underline{k}}^{k} x(u)\delta\frac{d}{du}\mathbb{E}[\sum_{j\in N_u^o} x(j)|u]du\right\} f(k)dk \\ =& \int_{\underline{k}}^{N} H(k)x(k)\delta\frac{d}{dk}\mathbb{E}[\sum_{j\in N_k^o} x(j)|k]f(k)dk,\end{aligned}$$

where in the last equality we recall the definition $H(k) = [1-F(k)]/f(k)$. Given the above, we rewrite the objective function as:

$$\begin{aligned}& \pi_0(x(\cdot)) \qquad (59)\\ =\;& N\int_{\underline{k}}^{N}\left\{\begin{array}{c}(a-c)x(k)-bx^2(k)+x(k)\delta\mathbb{E}[\sum_{j\in N_k^o}x(j)|k]-V(\underline{k}) \\ -\int_{\underline{k}}^{k}x(u)\delta\frac{d}{du}\mathbb{E}[\sum_{j\in N_u^o}x(j)|u]du\end{array}\right\}f(k)dk. \\ =\;& N\int_{\underline{k}}^{N}\left\{\begin{array}{c}(a-c)x(k)-bx^2(k)+x(k)\delta\mathbb{E}[\sum_{j\in N_k^o}x(j)|k]-V(\underline{k}) \\ -H(k)x(k)\delta\frac{d}{dk}\mathbb{E}[\sum_{j\in N_k^o}x(j)|k]\end{array}\right\}f(k)dk \\ \underbrace{=}_{V(\underline{k})=0 \text{ at optimum}}\;& N\int_{\underline{k}}^{N}\left[\begin{array}{c}(a-c)x(k)-bx^2(k) \\ +x(k)\delta\left\{\begin{array}{c}\mathbb{E}[\sum_{j\in N_k^o}x(j)|k] \\ -H(k)\frac{d}{dk}\mathbb{E}[\sum_{j\in N_k^o}x(j)|k]\end{array}\right\}\end{array}\right]f(k)dk\end{aligned}$$

The expression in (59) shows that the firm's expected payoff is decreasing in $V(\underline{k})$. Thus, at optimality $V(\underline{k}) = 0$.

The firm's (transformed) problem is

$$\max_{x(\cdot)} \quad N\int_{\underline{k}}^{N}\left[\begin{array}{c}(a-c)x(k)-bx^2(k) \\ +x(k)\delta\left\{\begin{array}{c}\mathbb{E}[\sum_{j\in N_k^o}x(j)|k] \\ -H(k)\frac{d}{dk}\mathbb{E}[\sum_{j\in N_k^o}x(j)|k]\end{array}\right\}\end{array}\right]f(k)dk \qquad (60)$$

$$\text{s.t.} \quad x(\cdot) \text{ is increasing.} \qquad (61)$$

For the moment, let us first ignore the constraint (61) and study the firm's objective (60).

Step (iii): Calculus of variations and verification of the candidate solution and its properties. From Proposition 2, the aggregate out-neighbor consumption of a degree-k customer is

$$\mathbb{E}[\sum_{j\in N_k^o} x(j)|k] = rm\,(k+rm)^r\int_k^N \frac{x(y)}{(y+rm)^{r+1}}dy,$$

and

$$
\begin{aligned}
&\frac{d}{dk}\mathbb{E}[\sum_{j\in N_k^o} x(j)|k] \\
= \quad & \frac{r}{k+rm}\mathbb{E}[\sum_{j\in N_k^o} x(j)|k] - \frac{rm}{k+rm}x(k). \qquad (62)
\end{aligned}
$$

Based on that, we rewrite the firm's transformed objective function $\pi_0(x(\cdot), P(\cdot))$ as

$$
\begin{aligned}
\pi_0(x(\cdot)) &= N\int_{\underline{k}}^{N}\left[\begin{array}{c}(a-c)x(k)-bx^2(k)\\ +x(k)\delta\left\{\begin{array}{c}\mathbb{E}[\sum_{j\in N_k^o} x(j)|k]\\ -H(k)\frac{d}{dk}\mathbb{E}[\sum_{j\in N_k^o} x(j)|k]\end{array}\right\}\end{array}\right]f(k)dk \qquad (63)\\
&= N\int_{\underline{k}}^{N}\left[\begin{array}{c}(a-c)x(k)-bx^2(k)\\ +x(k)\delta\left\{\begin{array}{c}\mathbb{E}[\sum_{j\in N_k^o} x(j)|k]\\ -H(k)\left(\begin{array}{c}\frac{r}{k+rm}\mathbb{E}[\sum_{j\in N_k^o} x(j)|k]\\ -\frac{rm}{k+rm}x(k)\end{array}\right)\end{array}\right\}\end{array}\right]f(k)dk\\
&= N\int_{\underline{k}}^{N}\left[\begin{array}{c}(a-c)x(k)-bx^2(k)\\ +x(k)\delta\left\{\begin{array}{c}\mathbb{E}[\sum_{j\in N_k^o} x(j)|k]\\ -\frac{r}{1+r}\left(\begin{array}{c}\mathbb{E}[\sum_{j\in N_k^o} x(j)|k]\\ -mx(k)\end{array}\right)\end{array}\right\}\end{array}\right]f(k)dk\\
&= N\int_{\underline{k}}^{N}\left[\begin{array}{c}(a-c)x(k)-bx^2(k)\\ +x(k)\delta\left\{\begin{array}{c}\frac{1}{1+r}\mathbb{E}[\sum_{j\in N_k^o} x(j)|k]\\ -\frac{r}{1+r}\left(-mx(k)\right)\end{array}\right\}\end{array}\right]f(k)dk\\
&= N\int_{\underline{k}}^{N}\left[\begin{array}{c}(a-c)x(k)-\left(b-\frac{\delta rm}{1+r}\right)x^2(k)\\ +x(k)\delta\left\{\frac{1}{1+r}\mathbb{E}[\sum_{j\in N_k^o} x(j)|k]\right\}\end{array}\right]f(k)dk \qquad (64)
\end{aligned}
$$

where we recall that $H(k) = \frac{k+rm}{1+r}$. Since $\frac{\delta}{1+r} > 0$, the presence of externality term $\mathbb{E}[\sum_{j\in N_k^o} x(j)|k]$ will positively shift the optimal consumption. That implies a lower bound on optimal consumption, denoted by $\underline{x}(\cdot)$, is the solution to the following objective that ignores the externality.

$$
\max_{x(\cdot)} \underline{\pi}_0(x(\cdot)) = N\int_{\underline{k}}^{N}\left[(a-c)x(k)-\left(b-\frac{\delta rm}{r+1}\right)x^2(k)\right]f(k)dk.
$$

Pointwise maximization of $\underline{\pi}_0(x(\cdot))$ (or by Euler equation approach) yields

$$
\underline{x}(k) \equiv \frac{a-c}{2\left(b-\frac{\delta rm}{r+1}\right)},
$$

for which the second-order condition holds obviously. It follows $\underline{x}(k) > 0$ given Assumptions 1, 2 ($a > c$ and $b > \delta m > \delta\frac{rm}{r+1}$). Thus the nonnegativity of $x(\cdot)$ is guaranteed. Denote $\underline{x}(k)$ by $\underline{x}^o$ for clarity.

Define

$$z^o(k) := \int_k^N \frac{x(y)}{(y+rm)^{r+1}} dy.$$

Given this definition, we have:

$$\begin{aligned} \mathbb{E}[\sum_{j\in N_k^o} x(j)|k] &= rm\,(k+rm)^r\, z^o(k), \\ x(k) &= -(k+rm)^{r+1}(z^o)'(k). \end{aligned}$$

Then one can rewrite the integrand within the objective function $\pi_0(x(\cdot))$, in terms of $z^o(k)$ and its derivative $(z^o)'(k)$:

$$\begin{aligned} G^o(k, z^o(k), (z^o)'(k)) &:= \left[\begin{array}{c} (a-c)x(k) - (b - \frac{\delta rm}{1+r})x^2(k) \\ +x(k)\delta \left\{ \frac{1}{1+r}\mathbb{E}[\sum_{j\in N_k^o} x(j)|k] \right\} \end{array} \right] f(k) \\ &= -\left[\begin{array}{c} (a-c)(k+rm)^{r+1}(z^o)'(k) + (b - \frac{\delta rm}{1+r}) \left((k+rm)^{r+1}(z^o)'(k) \right)^2 \\ +(k+rm)^{2r+1}(z^o)'(k)\frac{\delta rm}{r+1} z^o(k) \end{array} \right] f(k). \end{aligned}$$

From (1), one can derive:

$$\begin{aligned} f(k) &= \frac{(1+r)(rm)^{1+r}}{(k+rm)^{2+r}} \\ \Rightarrow\quad f'(k) &= \frac{-(1+r)(2+r)(rm)^{1+r}}{(k+rm)^{3+r}} = -\frac{r+2}{k+rm} f(k). \end{aligned}$$

Taking the partial derivatives of G^o w.r.t. $z^o(k)$ and $(z^o)'(k)$ gives

$$\begin{aligned} G^o_{z^o(k)}(k, z^o(k), (z^o)'(k)) &= -\delta \frac{m}{r+1} r(rm)^{r+1}(r+1)(k+rm)^{r-1}(z^o)'(k) \\ G^o_{(z^o)'(k)}(k, z^o(k), (z^o)'(k)) &= \frac{1}{k+rm} \left\{ \begin{array}{c} (rm)^{r+1}(r+1)\left[-(a-c) + (k+rm)^r \right. \\ \left. \left(-\delta \frac{m}{r+1} rz(k) + 2(k+rm)(-b + \delta \frac{m}{r+1} r)(z^o)'(k) \right) \right] \end{array} \right\} \end{aligned}$$

From the standard argument of calculus of variation, the optimal solution $z^{o*}(\cdot)$ for the firm is the solution to the following *Euler equation*:

$$G^o_{z^o(k)}(k, z^o(k), (z^o)'(k)) = \frac{d}{dk} G^o_{(z^o)'(k)}(k, z^o(k), (z^o)'(k)).$$

That is, $z^{o*}(\cdot)$ solves:

$$(a-c) + (k+rm)^r \left[-\frac{\delta rm}{r+1}(r-1)z^o(k) + 2(k+rm)\left(-b + \frac{\delta rm}{r+1} \right) \left(r(z^o)'(k) + (k+rm)(z^o)''(k) \right) \right] = 0\ \forall k. \tag{65}$$

(65) takes form of a second-order differential equation with regard to $z^o(\cdot)$, whose solution is explicit but cumbersome. Fortunately, we can verify the required monotonicity without recourse to the explicit solution (see below). The second-order condition, also called the Legendre

condition, is satisfied under Assumption 2:

$$G^{o}_{z^{o\prime}(k)z^{o\prime}(k)}(k,z^{o}(k),z^{o\prime}(k)) = -2(rm)^{r+1}(r+1)(k+rm)^{r}\left(b-\delta\frac{m}{r+1}r\right) \leq 0,\ \forall k.$$

To check for monotonicity, we put Euler equation into x-notation,

$$\begin{array}{r} (a-c)-\delta\frac{m}{r+1}(r-1)r(k+rm)^{r}\int_{k}^{N}\frac{x^{*}(y)}{(y+rm)^{r+1}}dy \\ +2\left(b-\delta\frac{m}{r+1}r\right)\left(-x^{*}(k)+(k+rm)x^{*\prime}(k)\right) \end{array} = 0,$$

and rearrange it as follows:

$$\begin{array}{r} -(a-c)+\delta\frac{m}{r+1}(r-1)r(k+rm)^{r}\int_{k}^{N}\frac{x^{*}(y)}{(y+rm)^{r+1}}dy \\ +2\left(b-\delta\frac{m}{r+1}r\right)x^{*}(k) \end{array} = 2\left(b-\delta\frac{m}{r+1}r\right)(k+rm)x^{*\prime}(k). \quad (66)$$

Under Assumption 2, $b-\delta\frac{m}{r+1}r > 0$. Observe that

$$\begin{aligned} 2\left(b-\delta\frac{m}{r+1}r\right)x^{*}(k) &> 2\left(b-\delta\frac{m}{r+1}r\right)\underline{x}^{o} \\ &= 2\left(b-\delta\frac{m}{r+1}r\right)\frac{a-c}{2\left(b-\frac{\delta rm}{r+1}\right)} \\ &= a-c. \end{aligned}$$

Thus, the right-hand side of (66) is greater than $\delta\frac{m}{r+1}(r-1)r(k+rm)^{r}\int_{k}^{N}\frac{x^{*}(y)}{(y+rm)^{r+1}}dy$, which suggests $x^{*\prime}(k)$ being positive if $r \geq 1$. Therefore, the monotonicity constraint is satisfied if $r \geq 1$ (Assumption 2).

Lastly, for the ease of presentation, we recollect the Eular equation using more existing notations:

$$\begin{aligned} & \begin{array}{r} -(a-c)+\frac{\delta rm}{r+1}(r-1)(k+rm)^{r}\int_{k}^{N}\frac{x^{*}(y)}{(y+rm)^{r+1}}dy \\ +2\left(b-\frac{\delta rm}{r+1}\right)x^{*}(k) \end{array} = 2\left(b-\frac{\delta rm}{r+1}\right)(k+rm)x^{*\prime}(k) \\ \Leftrightarrow\ & -(a-c)+\delta\frac{r-1}{r+1}\mathbb{E}[\sum_{j\in N_{k}^{o}}x^{*}(j)|k]+2\left(b-\frac{\delta rm}{r+1}\right)x^{*}(k) = 2\left(b-\frac{\delta rm}{r+1}\right)(k+rm)x^{*\prime}(k) \\ \Leftrightarrow\ & -\left(\frac{a-c}{b-\frac{\delta rm}{r+1}}\right)+\frac{\delta(r-1)}{b+br-\delta rm}\mathbb{E}[\sum_{j\in N_{k}^{o}}x^{*}(j)|k]+2x^{*}(k) = 2(k+rm)x^{*\prime}(k) \\ \Leftrightarrow\ & \frac{\delta(r-1)}{2(b+br-\delta rm)}\mathbb{E}[\sum_{j\in N_{k}^{o}}x^{*}(j)|k]+x^{*}(k)-(k+rm)x^{*\prime}(k)-\underline{x}^{o} = 0 \end{aligned}$$

This form of Eular equation is presented in Theorem 1. Plugging $x^{*}(\cdot)$ into the payment function (57) gives the optimal payment scheme $P^{*}(\cdot)$ below.[22]

$$P^{*}(x^{*}(k)) \quad = \quad ax^{*}(k)-b(x^{*}(k))^{2}+x^{*}(k)\delta\mathbb{E}[\sum_{j\in N_{k}^{o}}x^{*}(j)|k]-\int_{0}^{k}x^{*}(u)\delta\frac{d}{du}\mathbb{E}[\sum_{j\in N_{u}^{o}}x^{*}(j)|u]du$$

[22]Note the induced consumption below $\underline{k}$ is zero.

Given the nature of coordination in the consumption game, one may suspect whether a certain pricing scheme can lead to multiple consumption equilibria. To address this issue, we will show that the equilibrium optimal to the firm can be uniquely implemented. Consider a prescribed menu $\{x^*(\cdot), P^*(\cdot)\}$ determined optimally from Theorem 1 and agents report their degrees under this menu. Denote by $\tilde{k}(\cdot)$ the strategy that the focal customer perceives that his neighbors will play (which maps true degree to a reported one). Then let $\tilde{x}(k) := x^*(\tilde{k}(k))$ be the resulting consumption from the pattern of neighbors' misreporting as speculated by the focal customer. Note that $\tilde{x}(k)$ may not be consistent with the desired consumption $x^*(k)$. We will show that any belief $\tilde{x}(\cdot)$ other than $x^*(\cdot)$ will not get implemented by $P^*(\cdot)$. The degree-k customer's payoff when reporting $\hat{k}$ is

$$\pi_k(\hat{k}, \tilde{x}(\cdot), x^*(\cdot), P^*(\cdot)) = ax^*(\hat{k}) - bx^{*2}(\hat{k}) + x^*(\hat{k})\delta\mathbb{E}[\sum_{j\in N_k^o} \tilde{x}(j)|k] - P^*(x^*(\hat{k})) \tag{67}$$

$$\underbrace{=}_{(14)} x^*(\hat{k})\delta\mathbb{E}[\sum_{j\in N_k^o} \tilde{x}(j)|k] - x^*(\hat{k})\delta\mathbb{E}[\sum_{j\in N_{\hat{k}}^o} x^*(j)|\hat{k}] + \int_0^{\hat{k}} x^*(u)\delta\frac{d}{du}\mathbb{E}[\sum_{j\in N_u^o} x^*(j)|u]du \tag{68}$$

$$= x^*(\hat{k})\delta\left(\mathbb{E}[\sum_{j\in N_k^o} \tilde{x}(j)|k] - \mathbb{E}[\sum_{j\in N_{\hat{k}}^o} x^*(j)|\hat{k}]\right) + \int_0^{\hat{k}} x^*(u)\delta\frac{d}{du}\mathbb{E}[\sum_{j\in N_u^o} x^*(j)|u]du \tag{69}$$

Define k_0 such that $\mathbb{E}[\sum_{j\in N_k^o} \tilde{x}(j)|k] = \mathbb{E}[\sum_{j\in N_{k_0}^o} x^*(j)|k_0]$. Note that $k = k_0$ if $\tilde{x}(\cdot)$ coincides with $x^*(\cdot)$ (or truth reporting $\tilde{k}(k) = k$). Thus

$$\begin{aligned}\pi_k(k_0, \tilde{x}(\cdot), x^*(\cdot), P^*(\cdot)) &= x^*(k_0)\delta\left(\mathbb{E}[\sum_{j\in N_k^o} \tilde{x}(j)|k] - \mathbb{E}[\sum_{j\in N_{k_0}^o} x^*(j)|k_0]\right) \\ &\quad + \int_0^{k_0} x^*(u)\delta\frac{d}{du}\mathbb{E}[\sum_{j\in N_u^o} x^*(j)|u]du \qquad (70) \\ &= \int_0^{k_0} x^*(u)\delta\frac{d}{du}\mathbb{E}[\sum_{j\in N_u^o} x^*(j)|u]du \qquad (71)\end{aligned}$$

$$\begin{aligned}\pi_k(k_0, \tilde{x}(\cdot), x^*(\cdot), P^*(\cdot)) - \pi_k(\hat{k}, \tilde{x}(\cdot), x^*(\cdot), P^*(\cdot)) &= \int_{\hat{k}}^{k_0} x^*(u)\delta\frac{d}{du}\mathbb{E}[\sum_{j\in N_u^o} x^*(j)|u]du \\ &\quad -x^*(\hat{k})\delta\left(\mathbb{E}[\sum_{j\in N_{k_0}^o} x^*(j)|k_0] - \mathbb{E}[\sum_{j\in N_{\hat{k}}^o} x^*(j)|\hat{k}]\right) \qquad (72) \\ &> 0 \text{ given } x^*(\cdot) \text{ increasing} \qquad (73)\end{aligned}$$

Therefore under $P^*(\cdot)$, the degree-k player anticipating others to play $\tilde{x}(\cdot)$ will optimally declare k_0. The resulting consumption $x^*(k_0)$ generally does not equal $\tilde{x}(k)$, by definition of k_0, unless $\tilde{x}(\cdot)$ coincides with $x^*(\cdot)$. That means $x^*(\cdot)$ (or truth reporting) will be the *only* equilibrium implemented by the scheme $P^*(\cdot)$. □

Proof of Proposition 3. To study the marginal price in variation with degree, note that incentive compatibility (53) implies

$$P'(x(k)) = a - 2bx(k) + \delta\mathbb{E}[\sum_{j\in N_k^o} x(j)|k] \quad (74)$$

$$\frac{dP'(x(k))}{dk} = -2bx'(k) + \frac{r}{k+rm}\left[\delta\mathbb{E}[\sum_{j\in N_k^o} x(j)|k] - \delta m x(k)\right] \quad (75)$$

Substituting the optimal solution into the above expression,

$$\begin{aligned}\frac{dP'(x^*(k))}{dk} &= -2bx^{*\prime}(k) + \frac{r}{k+rm}\left[\delta\mathbb{E}[\sum_{j\in N_k^o} x^*(j)|k] - \delta m x^*(k)\right] \\ \underbrace{<}_{(13),\, x^*(\cdot)\geq \underline{x}} & -2bx^{*\prime}(k) + \frac{r}{k+rm}\left[\frac{r+1}{r-1}2\left(b-\frac{\delta rm}{r+1}\right)(k+rm)x^{*\prime}(k) - \delta m x^*(k)\right]\end{aligned}$$

The above quantity is negative if the multiplier of $x^{*\prime}(k)$ is negative, which can reduce to $\frac{b}{b+br-\delta rm} > \frac{r}{r-1}$. In this case, the scheme charges a lower marginal price for higher degree customers at optimum. Then,

$$P''(x^*(k)) = \frac{dP'(x^*(k))}{dk} / x^{*\prime}(k) < 0.$$

which gives rise to a quantity discount menu.

For the out-neighbor model, recall the upper bound of consumption, derived from the variant game, $\bar{x}(k) = \frac{a-c+\delta\mathbb{E}[\sum_{j\in N_k^o} x(j)|k]}{2b}$, or equivalently,

$$\mathbb{E}[\sum_{j\in N_k^o} x(j)|k] > \frac{2bx(k)-(a-c)}{\delta}. \quad (76)$$

In seeking for conditions for quantity premium menu, we obtain

$$
\begin{aligned}
\frac{dP'(x^*(k))}{dk} &= -2bx^{*\prime}(k) + \frac{r}{k+rm}\left[\delta\mathbb{E}[\sum_{j\in N_k^o} x^*(j)|k] - \delta m x^*(k)\right] \\
&= \frac{1}{k+rm}\left[-2b(k+rm)x^{*\prime}(k) + \delta r\left(\mathbb{E}[\sum_{j\in N_k^o} x^*(j)|k] - mx^*(k)\right)\right] \\
&\underbrace{=}_{(13)} \frac{1}{k+rm}\Big[-b\Big(-\Big(\tfrac{a-c}{b-\frac{\delta rm}{r+1}}\Big) + \tfrac{\delta}{b-\frac{\delta rm}{r+1}}\tfrac{r-1}{r+1}\mathbb{E}[\textstyle\sum_{j\in N_k^o} x^*(j)|k] + 2x^*(k)\Big) \\
&\quad +\delta r\left(\mathbb{E}[\sum_{j\in N_k^o} x^*(j)|k] - mx^*(k)\right)\Big] \\
&> \frac{\delta}{k+rm}\Big[-\Big(\tfrac{b}{b-\frac{\delta rm}{r+1}}\tfrac{r-1}{r+1}\mathbb{E}[\textstyle\sum_{j\in N_k^o} x^*(j)|k] + 2\tfrac{b}{\delta}x^*(k)\Big) \\
&\quad + \left(r\mathbb{E}[\sum_{j\in N_k^o} x^*(j)|k] - rmx^*(k)\right)\Big] \\
&= \frac{\delta}{k+rm}\left[\left(\left(r - \frac{b}{b-\frac{\delta rm}{r+1}}\frac{r-1}{r+1}\right)\mathbb{E}[\sum_{j\in N_k^o} x^*(j)|k] - \left(rm+\frac{2b}{\delta}\right)x^*(k)\right)\right] \\
&\underbrace{>}_{(76)} \frac{\delta}{k+rm}\left[\left(\left(r-\frac{b(r-1)}{b+br-\delta rm}\right)\left(\frac{2bx^*(k)-(a-c)}{\delta}\right) - \left(rm+\frac{2b}{\delta}\right)x^*(k)\right)\right] \\
&= \frac{1}{k+rm}\left[\left(\left(r-\frac{b(r-1)}{b+br-\delta rm}\right)(2bx^*(k)-(a-c)) - (\delta rm+2b)\,x^*(k)\right)\right] \\
&\underbrace{>}_{x^*(k)>\underline{x}^o,(77)} \frac{1}{k+rm}\left[\left(\left(r-\frac{b(r-1)}{b+br-\delta rm}\right)(2b\underline{x}^o-(a-c)) - (\delta rm+2b)\,\underline{x}^o\right)\right] \\
&= \frac{a-c}{k+rm}\left[\frac{-\delta^2r^2m^2(r-1)-2b^2(r+1)^2+b\delta rm(r^2+3)}{2(b+br-\delta rm)^2}\right] \\
&\underbrace{>}_{(78)} 0
\end{aligned}
$$

where the sufficient conditions include

$$
\begin{aligned}
2b(r-1)\left[1-\frac{b}{b+br-\delta rm}\right] - \delta rm &> 0, && (77)\\
-\delta^2r^2m^2(r-1)-2b^2(r+1)^2+b\delta rm(r^2+3) &> 0. && (78)
\end{aligned}
$$

In this case, the firm charges a higher marginal price for higher degree customers at optimum. This implies quantity premium, since $P''(x^*(k)) = \frac{dP'(x^*(k))}{dk}/x^{*\prime}(k) > 0$. □

Proof of Corollary 1

Denote by $\Pi^o(k) := P(x(k)) - cx(k)$ the profit earned from a single degree-k customer.

Therefore we obtain

$$\begin{aligned}\frac{d}{dk}\Pi^o(k) &= x'(k)\left(P'(x(k)) - c\right) \\ \underbrace{=}_{(53)} & x'(k)\left(a - c - 2bx(k) + \delta\mathbb{E}[\sum_{j\in N_k^o} x(j)|k]\right)\end{aligned}$$

At optimality, $x^{*\prime}(k) > 0$, and $a - c - 2bx^*(k) + \delta\mathbb{E}[\sum_{j\in N_k^o} x^*(j)|k] > 0$ (as $x^*(k) < \bar{x}(k)$). Thus $\frac{d}{dk}\Pi^{*o}(k) > 0$. That is, the firm grasps higher profit from higher degree customers at optimum.

Comparing the firm's second-best objective to that of the first-best, denoted as $\pi_0^{oFB}(x(\cdot))$,

$$\begin{aligned}\max_{x(\cdot)} \pi_0^o(x(\cdot)) &= N\int_{\underline{k}}^{N}\left[\begin{array}{c}(a-c)x(k) - (b - \frac{\delta rm}{1+r})x^2(k) \\ +x(k)\delta\left\{\frac{1}{1+r}\mathbb{E}[\sum_{j\in N_k^o} x(j)|k]\right\}\end{array}\right] f(k)dk, \\ &= N\int_{\underline{k}}^{N}\left[\begin{array}{c}(a-c)x(k) - bx^2(k) \\ +x(k)\delta\mathbb{E}[\sum_{j\in N_k^o} x(j)|k] \\ -x(k)\delta\frac{r}{1+r}\mathbb{E}[\sum_{j\in N_k^o} x(j)|k] + \frac{\delta rm}{1+r}x^2(k)\end{array}\right] f(k)dk, \\ &= N\int_{\underline{k}}^{N}\left[\begin{array}{c}(a-c)x(k) - bx^2(k) + x(k)\delta\mathbb{E}[\sum_{j\in N_k^o} x(j)|k] \\ -x(k)\frac{\delta r}{1+r}\left(\mathbb{E}[\sum_{j\in N_k^o} x(j)|k] - mx(k)\right)\end{array}\right] f(k)dk \\ \max_{x(\cdot)} \pi_0^{oFB}(x(\cdot)) &= N\int_{\underline{k}}^{N}\left[\begin{array}{c}(a-c)x(k) - bx^2(k) \\ +x(k)\delta\mathbb{E}[\sum_{j\in N_k^o} x(j)|k]\end{array}\right] f(k)dk\end{aligned}$$

Since any feasible second-best solution will satisfy $\mathbb{E}[\sum_{j\in N_k^o} x(j)|k] - mx(k) > 0$ (given the consumption must be increasing in degree), the difference between first-best and second-best objectives,

$$x(k)\frac{\delta r}{1+r}\left(\mathbb{E}[\sum_{j\in N_k^o} x(j)|k] - mx(k)\right),$$

consists of $x(k)$ and a positive multiplier. Thus the incentive for consumption in second-best is less than that in first-best, under any feasible choice of $x(\cdot)$ in the second-best case. That suggests downward distortion in consumption in the second-best scenario.[23]

That $x^*(\cdot)$ increases by k is a direct consequence of being a feasible solution to the firm's

[23]Note that the value of $\underline{k}$ is lower in the first-best profit function than that of the second-best case, as the firm earns more in FB from each degree type-k customers (extracting the information rent), thus selling to more customers down the list of types in FB. Therefore, the downward distortion result still holds in the range of types between $\underline{k}$ of FB and $\underline{k}$ of SB, where the induced consumption is positive for FB and 0 for SB. For types lower than $\underline{k}$ of FB, consumptions in both FB and SB are 0. (As shown before:) In types higher than $\underline{k}$ of SB, consumption in FB dominates that in SB (both nonzero). Hence, downward distortion holds for all degree types.

problem (4)-(6). Now we show that $x^*(\cdot)$ also increases in m. Recall the firm's profit

$$\pi_0^o(x(\cdot)) = N\int_{\underline{k}}^{N}\left[\begin{array}{c}(a-c)x(k)-\left(b-\frac{\delta rm}{1+r}\right)x^2(k)\\ +x(k)\delta\left\{\ \frac{1}{1+r}\mathbb{E}[\sum_{j\in N_k^o}x(j)|k]\ \right\}\end{array}\right]f(k)dk,$$

where the bracketed term

$$(a-c)x(k)-\left(b-\frac{\delta rm}{1+r}\right)x^2(k)+x(k)\delta\left\{\ \frac{1}{1+r}\mathbb{E}[\sum_{j\in N_k^o}x(j)|k]\ \right\} \tag{79}$$

captures the profit gathered from a single degree-k customer under consumption $x(\cdot)$. We start with the hypothesis that $\mathbb{E}[\sum_{j\in N_k^o}x(j)|k]$ increases in m, and will show that this can be reinforced at optimum. Now suppose m increases. Then the disutility coefficient $b-\frac{\delta rm}{1+r}$ in (79) is reduced, and $\mathbb{E}[\sum_{j\in N_k^o}x(j)|k]$ increases (by hypothesis). Therefore, the induced incentive for consumption for the customer in question increases. The optimal consumption level thus increases in m. Given that $x(k)$ increases in k (at optimum) and that increasing m triggers a first-order stochastic increase of the out-neighbor degree distributions F_k^{o1}, F_k^{o2} (Proposition 1), it follows that $\mathbb{E}[\sum_{j\in N_k^o}x(j)|k]$ increases in m,[24] which reinforces the foregoing hypothesis and concludes the proof.

It remains to show that the optimal profit of the firm also increases in m. Given that the coefficient $b-\frac{\delta rm}{1+r}$ declines in m, and $\mathbb{E}[\sum_{j\in N_k^o}x(j)|k]$ increase in m at optimum (as shown above), it follows that the profit earned from one single degree-k player, (79), also increases in m at optimum. Since the degree distribution $F(\cdot)$ first-order stochastically increases in m (c.f. Theorem 7 of Jackson and Rogers (2007)), and (79) is at optimum an increasing function of degree k (c.f. the first bullet point of Corollary 1), we conclude that the firm's overall profit $\pi_0^o(x(\cdot))$ (aggregated over all degree types) increases in m at optimum. □

Proof of Proposition 4. To allow for arbitrary price discrimination, suppose that the firm draws a random number $\eta\in[0,1]$ for each agent from cdf. $Q(\cdot)$, based on which the allocation $\{x(\eta), P(\eta)\}$ is made to that agent. The agent will then decide to accept or reject. Notice that, since the firm does not know who is connected to whom, it cannot enforce any correlation between random numbers of neighbors. So everyone will perceive his neighbor's number as samely randomly distributed according to $Q(\cdot)$. Note that the topological correlation is not explored here, exactly because the pricing does not solicit the degree information. If it does, the firm can then utilize the knowledge of neighbor degree distribution to refine individual allocations.

For any player with number η, the expected total out-neighbor consumption when neighbors accept the contract is $\mathbb{E}[\sum_{j\in N_\eta^o}x(j)|\eta] = m\int_0^1 x(j)q(j)dj$. Furthermore, the expected out-neighbor consumption does not change with one's own number because neighbors' numbers

[24]Recall that $\mathbb{E}[\sum_{j\in N_k^o}x(j)|k] = o_r\int_k^N x(y)f_k^{o1}(y)dy + o_n\int_k^N x(y)f_k^{o2}(y)dy$, where $o_r=\frac{rm}{r+1}$ and $o_n=\frac{m}{r+1}$ both increase in m.

are independent, i.e. $\frac{d}{d\eta}\mathbb{E}[\sum_{j\in N_\eta^o} x(j)|\eta] = 0$. For given $Q(\cdot)$, the firm's optimization resembles a first-best pricing problem, where the agent's type – his number – is observable to the firm.[25] If accepting the firm's offer, the type-η agent's payoff, while others accepting the contract, is given by

$$\pi_\eta(x(\cdot), P(\cdot)) \quad = \quad ax(\eta) - bx^2(\eta) + x(\eta)\delta\mathbb{E}[\sum_{j\in N_\eta^o} x(j)|\eta] - P(\eta),$$

Under first-best, we should have $\pi_\eta(x(\cdot), P(\cdot)) = 0$ (no information rent for the agent). That leads to the pricing scheme is $P^*(\eta) = ax(\eta) - bx^2(\eta) + x(\eta)\delta\mathbb{E}[\sum_{j\in N_\eta^o} x(j)|\eta]$. The firm then obtains the maximum social welfare as follows (where $\mathbb{E}, \mathbb{V}$ respectively represents expectation and variance).

$$\max \pi_0(x, P(\cdot)) = N\mathbb{E}\left\{(a-c)x(\eta) - bx^2(\eta) + x(\eta)\delta\mathbb{E}[\sum_{j\in N_\eta^o} x(j)|\eta]\right\} \tag{80}$$

$$= N\left\{(a-c)\mathbb{E}[x(\eta)] - b\mathbb{E}[x^2(\eta)] + \mathbb{E}[x(\eta)]\delta m\mathbb{E}[x(j)|_{j\in N_\eta^o}\eta]\right\} \tag{81}$$

$$= N\left\{(a-c)\mathbb{E}[x(\eta)] - b\mathbb{E}[x^2(\eta)] + \delta m\mathbb{E}[x(\eta)]^2\right\} \tag{82}$$

$$= N\left\{(a-c)\mathbb{E}[x(\eta)] - b\left(\mathbb{V}[x(\eta)] + \mathbb{E}[x(\eta)]^2\right) + \delta m\mathbb{E}[x(\eta)]^2\right\} \tag{83}$$

$$= N\left\{(a-c)\mathbb{E}[x(\eta)] - b\mathbb{V}[x(\eta)] - (b-\delta m)\mathbb{E}[x(\eta)]^2\right\} \tag{84}$$

Thus the firm at optimum wants to minimize the variance of $x(\eta)$ by providing a menu that contains only a single consumption level x, i.e. $\mathbb{V}[x(\eta)] = 0, \mathbb{E}[x(\eta)] = x$. So the firm's problem reduces to

$$\max \pi_0(x, P(\cdot)) \quad = \quad N\left[(a-c)x - (b-\delta m)x^2\right] \tag{85}$$

Provided $b > \delta m$ (Assumption 2), the optimal consumption level $x^*(\eta) = \frac{a-c}{2(b-\delta m)}$; and the optimal payment is $P^*(\eta) = ax^*(\eta) - bx^{*2}(\eta) + x^*(\eta)\delta\mathbb{E}[\sum_{j\in N_\eta^o} x^*(j)|\eta] = \frac{a^2-c^2}{4(b-\delta m)}$. Substituting x^* into (85), the firm's profit at the optimum is $N\frac{(a-c)^2}{4(b-\delta m)}$.

Observe the firm's profit with the DS pricing, (87), and that under NDS, (85),

$$\pi_0(x(\cdot)) \underbrace{=}_{(16)} N\int_{\underline{k}}^{N}\left[\begin{array}{c}(a-c)x(k) - bx^2(k) \\ +x(k)\delta\left\{\begin{array}{c}\mathbb{E}[\sum_{j\in N_k^o} x(j)|k] \\ -H(k)\frac{d}{dk}\mathbb{E}[\sum_{j\in N_k^o} x(j)|k]\end{array}\right\}\end{array}\right] f(k)dk \tag{86}$$

$$\underbrace{=}_{(64)} N\int_{\underline{k}}^{N}\left[\begin{array}{c}(a-c)x(k) - \left(b - \frac{\delta rm}{1+r}\right)x^2(k) \\ +x(k)\delta\left\{\frac{1}{1+r}\mathbb{E}[\sum_{j\in N_k^o} x(j)|k]\right\}\end{array}\right] f(k)dk \tag{87}$$

[25]The decision of $Q(\cdot)$ is cosmetic since η merely serves a proxy for the firm to arbitrarily price-discriminate the customers. As shown in Proposition 4, the induced consumption and optimal pricing are independent of $Q(\cdot)$.

and note that (87) reduces to (85) if $x(k) \equiv x$ and $\underline{k} \to 0$ (i.e. all customers served in a homogenous equilibrium). Since the firm is strictly better off with more flexible control of $x(k)$ over all k (as in (87)) than restricting to a single x (as in (85)), [26] the profit earned through NDS pricing is dominated by that under DS pricing. □

Proof of Proposition 5. In case of the Erdös and Rényi graph, $r \to \infty$, $\lim_{r\to\infty} H(k) := \lim_{r\to\infty} \frac{k+rm}{1+r} = m$. We obtain:

$$\begin{aligned} & \lim_{r\to\infty} \mathbb{E}[\sum_{j\in N_k^o} x(j)|k] \\ = & \lim_{r\to\infty} \left\{ rm\,(k+rm)^r \int_k^N \frac{x(y)}{(y+rm)^{r+1}} dy \right\} \\ = & \int_k^N x(y) \left\{ \lim_{r\to\infty} \frac{rm\,(k+rm)^r}{(y+rm)^{r+1}} \right\} dy \\ = & \int_k^N x(y) \exp\left(\frac{k-y}{m}\right) dy. \end{aligned}$$

Accordingly, its derivative is:

$$\begin{aligned} & \frac{d}{dk}\mathbb{E}[\sum_{j\in N_k^o} x(j)|k]\Big|_{r\to\infty} \\ = & \frac{r}{k+rm}\mathbb{E}[\sum_{j\in N_k^o} x(j)|k]\Big|_{r\to\infty} - \frac{rm}{k+rm}x(k) \\ = & \frac{1}{m}\mathbb{E}[\sum_{j\in N_k^o} x(j)|k]\Big|_{r\to\infty} - x(k). \end{aligned}$$

Recall that in the firm's objective (16); the profit earned from a single degree-k player in the Erdös and Rényi graph can be rewritten as:

$$\begin{aligned} \Pi_o^R(x(\cdot)) &= (a-c)x(k) - bx^2(k) + x(k)\delta \left\{ \begin{array}{c} \mathbb{E}[\sum_{j\in N_k^o} x(j)|k] \\ -H(k)\frac{d}{dk}\mathbb{E}[\sum_{j\in N_k^o} x(j)|k] \end{array} \right\} \\ &= (a-c)x(k) - bx^2(k) + x(k)\delta \left\{ \begin{array}{c} \mathbb{E}[\sum_{j\in N_k^o} x(j)|k] \\ -m\left[\frac{1}{m}\mathbb{E}[\sum_{j\in N_k^o} x(j)|k] - x(k)\right] \end{array} \right\} \\ &= (a-c)x(k) - (b-\delta m)x^2(k). \end{aligned}$$

Notice that the externality term is cancelled off in the deduction. The resulting objective function of the firm becomes identical to (85). The proof follows from that of Proposition 4. □

Proof of Theorem 2. The proof is similar to that of Theorem 1, except for the modification of expected neighbor consumption. In steps (i) and (ii), the analysis is identical and therefore

[26] Note that maximizing (87) will necessarily result in degree-heterogeneous consumption that differs from the solution to (85), provided that the degree distribution and neighbor degree distributions are not symmetric across degrees. Only in the Erdös and Rényi graph where the network externalities are cancelled off, the solutions of DS pricing and NDS pricing will become identical. See Proposition 5.

we omit it. We start directly with step (iii). We decompose it into two sub-steps: (a) Calculus of variations and (b) Verification of the candidate solution and its properties.

(a) Calculus of variations. Let

$$z(k) \equiv \int_0^k \frac{x(y)}{(y+rm)^2}dy = \frac{1}{(k+rm)rm}\mathbb{E}[\sum_{j\in N_k^i} x(j)|k].$$

Hence we have

$$\begin{aligned} x(k) &= z'(k)(k+rm)^2, \\ z'(k) &= \frac{x(k)}{(k+rm)^2}, \\ z''(k) &= \frac{x'(k)}{(k+rm)^2} - \frac{2x(k)}{(k+rm)^3}. \end{aligned}$$

Using this definition, we convert the firm's profit (16) as a function of $z(k)$ and $z'(k)$ below.

$$\begin{aligned} &\pi_0(x(\cdot)) \\ &= N\int_{\underline{k}}^N \left\{ \begin{array}{c} (a-c)z'(k)(k+rm)^2 - b[z'(k)]^2(k+rm)^4 \\ +z'(k)(k+rm)^2\delta \left[\begin{array}{c} rm(k+rm)z(k) \\ -H(k)\frac{d}{dk}\{rm(k+rm)z(k)\} \end{array} \right] \end{array} \right\} f(k)dk \\ &= N\int_{\underline{k}}^N \left\{ \begin{array}{c} (a-c)z'(k)(k+rm)^2 - b[z'(k)]^2(k+rm)^4 \\ +z'(k)(k+rm)^2\delta \left[\begin{array}{c} rm(k+rm)z(k) \\ -H(k)rmz(k) \\ -H(k)rm(k+rm)z'(k) \end{array} \right] \end{array} \right\} f(k)dk. \end{aligned} \tag{88}$$

Then (88) exhibits the structure of calculus of variations. Let

$$G^i(k,z(k),z'(k)) := \left\{ \begin{array}{c} (a-c)z'(k)(k+rm)^2 - b[z'(k)]^2(k+rm)^4 \\ +z'(k)(k+rm)^2\delta \left[\begin{array}{c} rm(k+rm)z(k) - H(k)rmz(k) \\ -H(k)rm(k+rm)z'(k) \end{array} \right] \end{array} \right\} f(k).$$

We obtain its partial derivatives as follows:

$$\begin{aligned} G^i_{z(k)}(k,z(k),z'(k)) &= \{z'(k)(k+rm)^2\delta[rm(k+rm) - H(k)rm]\}f(k) \\ G^i_{z'(k)}(k,z(k),z'(k)) &= f(k)\left\{ \begin{array}{c} (a-c)(k+rm)^2 - 2bz'(k)(k+rm)^4 \\ +\delta(k+rm)^2 \left[\begin{array}{c} rm(k+rm)z(k) \\ -2H(k)rm(z(k)+(k+rm)z'(k)) \end{array} \right] \end{array} \right\}. \end{aligned}$$

The optimal solution $z^*(\cdot)$ is identified by the first-order condition below (referred as the Euler equation):

$$G^i_{z(k)}(k,z(k),z'(k)) = dG^i_{z'(k)}(k,z(k),z'(k))/dk. \tag{89}$$

Note that (89) takes form of a second-order differential equation with regard to $z(\cdot)$, whose solution is explicit but cumbersome. In our case, we can verify the required monotonicity without writing down the explicit form of the solution (see later Step (b) of the proof).

Recall that

$$\begin{aligned} H(k) &= \frac{1-F(k)}{f(k)} = \frac{k+rm}{1+r} \Rightarrow H^{'}(k) = \frac{1}{r+1}, \\ f(k) &= \frac{(1+r)(rm)^{1+r}}{(k+rm)^{2+r}} \Rightarrow f^{'}(k) = \frac{-(r+1)(r+2)(rm)^{r+1}}{(k+rm)^{r+3}}. \end{aligned} \tag{90}$$

Putting the Euler equation in x-notation, we have

$$\begin{aligned} &(a-c)r(1+r) + \delta m(r-1)r^2(k+rm)\int_0^k \frac{x^*(y)}{(y+rm)^2}dy \\ &+2(b+br+\delta rm)(-rx^*(k)+(k+rm)x^{*\prime}(k)) = 0. \end{aligned} \tag{91}$$

The second-order condition for optimality (Legendre condition)

$$G^i_{z'(k)z'(k)}(k, z(k), z^{'}(k)) \leq 0, \ \forall k,$$

can be easily shown satisfied. Thus, the Euler equation (89) identifies the maximum point of the firm's objective function.

(b) Verification of the candidate solution and its properties. Now we proceed to show the solution to Euler equation (91) does qualify the monotonic constraint (61). First, we need to derive some lower bound on the induced consumption. As suggested in Proposition 2,

$$\mathbb{E}[\sum_{j\in N_k^i} x(j)|k] = (k+rm)rm\int_0^k \frac{x(y)}{(y+rm)^2}dy.$$

Its derivative is

$$\frac{d}{dk}\mathbb{E}[\sum_{j\in N_k^i} x(j)|k] = rm\left(\int_0^k \frac{x_j(y)}{(y+rm)^2}dy + \frac{x_j(k)}{k+rm}\right) = \frac{\mathbb{E}[\sum_{j\in N_k^i} x(j)|k]}{k+rm} + \frac{rmx(k)}{k+rm}. \tag{92}$$

Rewrite the firm's transformed objective function by substituting $\frac{d}{dk}\mathbb{E}[\sum_{j\in N_k^i} x(j)|k]$ above:

$$\begin{aligned}
&\max_{x(\cdot)} \pi_0(x(\cdot)) \\
= \;& N\int_{\underline{k}}^{N}\left[\begin{array}{c}(a-c)x(k)-bx^2(k)\\ +x(k)\delta\left\{\mathbb{E}[\sum_{j\in N_k^i} x(j)|k]-H(k)\frac{d}{dk}\mathbb{E}[\sum_{j\in N_k^i} x(j)|k]\right\}\end{array}\right]f(k)dk \\
= \;& N\int_{\underline{k}}^{N}\left[\begin{array}{c}(a-c)x(k)-bx^2(k)\\ +x(k)\delta\left\{\mathbb{E}[\sum_{j\in N_k^i} x(j)|k]-H(k)\left(\frac{\mathbb{E}[\sum_{j\in N_k^i} x(j)|k]}{k+rm}+\frac{rmx(k)}{k+rm}\right)\right\}\end{array}\right]f(k)dk \\
= \;& N\int_{\underline{k}}^{N}\left[\begin{array}{c}(a-c)x(k)-bx^2(k)\\ +x(k)\delta\left\{\mathbb{E}[\sum_{j\in N_k^i} x(j)|k]\left(1-\frac{H(k)}{k+rm}\right)-H(k)\left(\frac{rm}{k+rm}\right)x(k)\right\}\end{array}\right]f(k)dk \\
= \;& N\int_{\underline{k}}^{N}\left[\begin{array}{c}(a-c)x(k)-bx^2(k)\\ +x(k)\delta\left\{\mathbb{E}[\sum_{j\in N_k^i} x(j)|k]\left(1-\frac{1}{1+r}\right)-H(k)\left(\frac{rm}{k+rm}\right)x(k)\right\}\end{array}\right]f(k)dk.
\end{aligned}$$

Since $1-\frac{1}{1+r}>0$, the presence of externality term $\mathbb{E}[\sum_{j\in N_k^i} x(j)|k]$ will positively shift the optimal consumption. That implies a lower bound on optimal consumption, denoted by $\underline{x}(\cdot)$, is the solution to the following objective that ignores the externality.

$$\max_{x(\cdot)} \underline{\pi}_0(x(\cdot)) = N\int_{\underline{k}}^{N}\left[\begin{array}{c}(a-c)x(k)-bx^2(k)\\ +x^2(k)\delta\left\{-H(k)\left(\frac{rm}{k+rm}\right)\right\}\end{array}\right]f(k)dk$$

Pointwise maximization of $\underline{\pi}_0(x(\cdot))$ (or by Euler equation approach equivalently) yields

$$\underline{x}(k)\equiv\frac{a-c}{2\left(b+\delta H(k)\frac{rm}{k+rm}\right)}=\frac{a-c}{2\left(b+\delta\frac{rm}{1+r}\right)}=\frac{(a-c)(1+r)}{2(b+br+\delta rm)},\tag{93}$$

for which the second-order condition holds obviously. $\underline{x}(k)>0$ given $a>c$. Thus the nonnegativity of $x(\cdot)$ is guaranteed. Denote $\underline{x}(k)$ by $\underline{x}^i$ for clarity.

Now note the Euler equation

$$(a-c)r(1+r)+\delta m(r-1)r^2(k+rm)\int_0^k\frac{x^*(y)}{(y+rm)^2}dy+2(b+br+\delta rm)(-rx^*(k)+(k+rm)x^{*\prime}(k))=0$$

implies

$$\begin{aligned}
&(k+rm)x^{*\prime}(k)\\
= \;& rx^*(k)-\frac{(a-c)r(1+r)}{2(b+br+\delta rm)}-\frac{\delta m(r-1)r^2(k+rm)}{2(b+br+\delta rm)}\int_0^k\frac{x^*(y)}{(y+rm)^2}dy\\
\underbrace{=}_{(93)} \;& r(x^*(k)-\underline{x}^i)+\frac{\delta m(1-r)r^2(k+rm)}{2(b+br+\delta rm)}\int_0^k\frac{x^*(y)}{(y+rm)^2}dy\\
> \;& \frac{\delta m(1-r)r^2(k+rm)}{2(b+br+\delta rm)}\int_0^k\frac{x^*(y)}{(y+rm)^2}dy\\
> \;& 0\;\text{ if } r<1
\end{aligned}$$

This suggests that the monotonicity constraint, $x^{*\prime}(\cdot) > 0$, is satisfied if $r < 1$ (Assumption 3).

For the ease of presentation, we rewrite Eular equation using more existing notations:

$$\begin{aligned}(k+rm)x^{*\prime}(k) &= r(x^*(k)-\underline{x}^i) + \frac{\delta m(1-r)r^2(k+rm)}{2(b+br+\delta rm)}\int_0^k \frac{x^*(y)}{(y+rm)^2}dy \\ &= r(x^*(k)-\underline{x}^i) + \frac{\delta(1-r)r}{2(b+br+\delta rm)}\mathbb{E}[\sum_{j\in N_k^i} x^*(j)|k], \qquad (94)\end{aligned}$$

which then gives us the optimality condition presented in Theorem 2:

$$\frac{\delta(1-r)r}{2(b+br+\delta rm)}\mathbb{E}[\sum_{j\in N_k^i} x^*(j)|k] + r(x^*(k)-\underline{x}^i) - (k+rm)x^{*\prime}(k) = 0.$$

Finally, one can substitute $x^*(\cdot)$ back to (57) (and note the consumption of degree types below $\underline{k}$ is zero) to obtain the payment function $P^*(\cdot)$ at optimum. Similar to Theorem 1, it can also be shown that the equilibrium $x^*(\cdot)$ can be uniquely implemented under $P^*(\cdot)$. □

Proof of Proposition 6. The first-order condition of IC constraints (53) implies that, at optimum,

$$\begin{aligned}P'(x(k)) &= a - 2bx(k) + \delta\mathbb{E}[\sum_{j\in N_k^i} x(j)|k], \\ \frac{dP'(x(k))}{dk} &= -2bx'(k) + \delta\frac{d\mathbb{E}[\sum_{j\in N_k^i} x(j)|k]}{dk} \\ &\underbrace{=}_{(92)} -2bx'(k) + \delta\left(\frac{\mathbb{E}[\sum_{j\in N_k^i} x(j)|k]}{k+rm} + \frac{rmx(k)}{k+rm}\right)\end{aligned}$$

Substituted into the derivative of incentive-compatible marginal price,

$$\begin{aligned}
\frac{dP'(x^*(k))}{dk} &= -2bx^{*\prime}(k) + \delta\left(\frac{\mathbb{E}[\sum_{j\in N_k^i} x^*(j)|k]}{k+rm} + \frac{rmx^*(k)}{k+rm}\right) \qquad (95)\\
&= \frac{1}{k+rm}\left(-2b(k+rm)x^{*\prime}(k) + \delta\left(\mathbb{E}[\sum_{j\in N_k^i} x^*(j)|k] + rmx^*(k)\right)\right)\\
&\underbrace{=}_{(19)} \frac{1}{k+rm}\left[-2b\left(\begin{array}{c} r(x^*(k)-\underline{x}^i) \\ +\frac{\delta(1-r)r}{2(b+br+\delta rm)}\mathbb{E}[\sum_{j\in N_k^i} x^*(j)|k]\end{array}\right) + \delta\left(\mathbb{E}[\sum_{j\in N_k^i} x^*(j)|k] + rmx^*(k)\right)\right]\\
&= \frac{1}{k+rm}\left[\begin{array}{c}\delta rmx^*(k) - 2br(x^*(k)-\underline{x}^i)\\ +\left(\delta - \frac{b\delta(1-r)r}{b+br+\delta rm}\right)\mathbb{E}[\sum_{j\in N_k^i} x^*(j)|k]\end{array}\right]\\
&> \frac{1}{k+rm}\left[\begin{array}{c}\delta rmx^*(k) - 2br(\bar{x}(k)-\underline{x}^i)\\ +\left(\delta - \frac{b\delta(1-r)r}{b+br+\delta rm}\right)\mathbb{E}[\sum_{j\in N_k^i} x^*(j)|k]\end{array}\right]\\
&= \frac{1}{k+rm}\left[\begin{array}{c}\delta rmx^*(k) + 2br\underline{x}^i - r(a-c+\delta\mathbb{E}[\sum_{j\in N_k^i} x^*(j)|k])\\ +\left(\delta - \frac{b\delta(1-r)r}{b+br+\delta rm}\right)\mathbb{E}[\sum_{j\in N_k^i} x^*(j)|k]\end{array}\right]\\
&= \frac{1}{k+rm}\left[\begin{array}{c}\delta rmx^*(k) + 2br\underline{x}^i - r(a-c)\\ +\left(\delta(1-r) - \frac{b\delta(1-r)r}{b+br+\delta rm}\right)\mathbb{E}[\sum_{j\in N_k^i} x^*(j)|k]\end{array}\right]\\
&> \frac{1}{k+rm}\left[\begin{array}{c}(\delta rm + 2br)\underline{x}^i - r(a-c)\\ +\left(\delta(1-r) - \frac{b\delta(1-r)r}{b+br+\delta rm}\right)\mathbb{E}[\sum_{j\in N_k^i} x^*(j)|k]\end{array}\right]\\
&\underbrace{>}_{r<1} 0 \qquad (96)
\end{aligned}$$

Notice $\frac{br}{b+br+\delta rm} < 1$, and that $r < 1$ implies $\frac{2r\delta m}{1+r} < \delta m$, which then leads to $\underline{x}^i > \frac{a-c}{2b+\delta m}$. Altogether it gives (96). This indicates the optimal payment scheme charges higher marginal price at optimum for higher degree customers. In this case we also conclude

$$P''(x^*(k)) = \frac{dP'(x^*(k))}{dk}/x^{*\prime}(k) > 0.$$

which gives rise to a *quantity premium* menu.

Next we will show that the firm, at optimality, reaps more profit from higher degree customers. The profit grasped from a degree-k consumer is given by

$$\Pi(k) := P(x(k)) - cx(k).$$

Note

$$\begin{aligned}
\frac{d}{dk}P(x(k)) &= P'(x(k))x'(k)\\
&= \left(a - 2bx(k) + \delta\mathbb{E}[\sum_{j\in N_k^i} x(j)|k]\right)x'(k)
\end{aligned}$$

So

$$\frac{d}{dk}\Pi(k) = x'(k)\left(a-c-2bx(k)+\delta\mathbb{E}[\sum_{j\in N_k^i} x(j)|k]\right)$$

At optimality, $x^{*\prime}(k) > 0$, and $a-c-2bx^*(k)+\delta\mathbb{E}[\sum_{j\in N_k^i} x^*(j)|k] > 0$ (as $x^*(k) < \bar{x}(k)$). Thus $\frac{d}{dk}\Pi^*(k) > 0$. □

Proof of Corollary 2. We first show the induced consumption is downward shifted compared to the consumption if the firm had complete information on customer degrees (first-best scenario). To be brief, the firm's first-best objective is

$$\max_{x(\cdot)} \pi_0^{FB}(x(\cdot)) = N\int_{\underline{k}}^{N}\left[\begin{array}{c}(a-c)x(k)-bx^2(k)\\ +x(k)\delta\left\{\mathbb{E}[\sum_{j\in N_k^i} x(j)|k]\right\}\end{array}\right] f(k)dk, \tag{97}$$

while recall the second-best firm's objective as

$$\begin{aligned}\max_{x(\cdot)} \pi_0(x(\cdot)) &= N\int_{\underline{k}}^{N}\left[\begin{array}{c}(a-c)x(k)-bx^2(k)\\ +x(k)\delta\left\{\mathbb{E}[\sum_{j\in N_k^i} x(j)|k]-H(k)\frac{d}{dk}\mathbb{E}[\sum_{j\in N_k^i} x(j)|k]\right\}\end{array}\right] f(k)dk\\ &= N\int_{\underline{k}}^{N}\left[\begin{array}{c}(a-c)x(k)-bx^2(k)\\ +x(k)\delta\left\{\mathbb{E}[\sum_{j\in N_k^i} x(j)|k]-\frac{1}{1+r}\left(\mathbb{E}[\sum_{j\in N_k^i} x(j)|k]+rmx(k)\right)\right\}\end{array}\right] f(k)dk\\ &= N\int_{\underline{k}}^{N}\left[\begin{array}{c}(a-c)x(k)-\left(b+\frac{\delta rm}{1+r}\right)x^2(k)\\ +x(k)\delta\left\{\mathbb{E}[\sum_{j\in N_k^i} x(j)|k]\left(1-\frac{1}{1+r}\right)\right\}\end{array}\right] f(k)dk. \end{aligned}\tag{98}$$

Compared to that of first-best (97), the linear benefit of consumption is discounted in the second-best objective (98) since $1-\frac{1}{1+r} < 1$, while the quadratic disutility term is strengthened. Thus the resulting consumption $x^*(k)$ in second-best should be lower than that in first-best. In other words, the firm faces a downward distortion in the consumption when its information regarding the social network is incomplete.[27] Lastly, note that the monotonicity of $x^*(k)$ with regard to k follows from its feasibility to the firm's problem (4)-(6). □

Proof of Proposition 7. The derivation of $f_k^t(\cdot)$, i.e. the pdf of neighbor degree distribution for a degree-k player (he) at time t involves the following steps:

Under mean field approximation, the preferential attachment gives rise to $k_i(t) = M(t/i_t(k))^{1/2}$ (Barabási and Albert, 1999), where $i_t(k)$ is the birthdate of the focal player. Thereby we have $i_t(k) = M^2t/k^2$.

Observe (i) at his birthdate, the focal player already had M out of the k neighbors that he has now. If a now-degree-d player (referred as player j) was of degree d' at the moment $i_t(k)$, we must have $d' = dM/k$. [28]

[27] By the same argument as in footnote 23, it can be shown the downward distortion holds for the whole range of degree types, although the first-best and second-best profit functions possess different values of $\underline{k}$.

[28] which can be obtained by solving the simultaneous equations $k_j(t) = M(t/j)^{1/2} = d$ and $k_j(i_t(k)) = d'$.

Therefore, the probability that player j became a neighbor to the focal player is $\frac{Md'}{2Mi_t(k)} = \frac{d'}{2i_t(k)} = \frac{dM/k}{2i_t(k)} = \frac{dMk}{2M^2t} = \frac{dk}{2Mt}$.[29] Applying Bayes' theorem,

$$
\begin{aligned}
f_k^t(d) &= \frac{Pr\{\text{neighbor}|\text{degree} = d\}f(d|d > k)}{\int_k^{\bar{k}} Pr\{\text{neighbor}|\text{degree} = d\}f(d|d > k)\mathrm{d}d} \\
&= \frac{Pr\{\text{neighbor}|\text{degree} = d\}f(d)}{\int_k^{\bar{k}} Pr\{\text{neighbor}|\text{degree} = d\}f(d)\mathrm{d}d} \\
&= \frac{\frac{dk}{2Mt}f(d)}{\int_k^{\bar{k}} \frac{dk}{2Mt}f(d)\mathrm{d}d} \\
&= \frac{df(d)}{\int_k^{\bar{k}} df(d)\mathrm{d}d} \\
&= \frac{1/d^2}{\int_k^{\bar{k}} 1/d^2\mathrm{d}d} \\
&= \frac{1/d^2}{d^{-1}|_N^k} \\
&\underbrace{=}_{N\to\infty} \frac{k}{d^2}
\end{aligned} \tag{99}
$$

which is time-invariant. Redenote $f_k^t(d)$ by $f_k^o(d)$, and we have

$$
\begin{aligned}
F_k^o(d) &= \int_k^d f_k^o(y)dy \\
&= 1 - \frac{k}{d}
\end{aligned} \tag{100}
$$

One can verify the legitimacy of $F_k^o(\cdot)$ on (k, ∞) by $F_k^o(k) = 0, F_k^o(\infty) = 1$.

(ii) After his birth, the focal player i got his $k - M$ neighbors over time. The probability of a now-degree-d player being his neighbor is calculated as follows. At date $i_t(d) = M^2t/d^2$, suppose the focal player i's degree was k'. Then it must be that $k' = kM/d$. [30]

The now-degree-d player reaches the focal player w.p. $M\frac{k'}{2Mi_t(d)} = \frac{k'}{2i_t(d)} = \frac{kM/d}{2i_t(d)} = \frac{kM/d}{2M^2t/d^2} = \frac{kd}{2Mt}$. Applying Bayes' theorem,

[29] Note that the total degrees of all players at the moment is $2Mi_t(k)$, since each player brings in M links which raise $2M$ degrees systemwise. Thus the probability j is reached by any of the M links under preferential attachment is $\frac{Md'}{2Mi_t(k)}$.

[30] which can be obtained by solving the simultaneous equations $k = M(t/i)^{1/2}$ and $k' = M(i_t(d)/i)^{1/2}$.

$$
\begin{aligned}
f_k^t(d) &= \frac{Pr\{\text{neighbor}|\text{degree}=d\}f(d|d<k)}{\int_M^k Pr\{\text{neighbor}|\text{degree}=d\}f(d|d<k)\mathrm{d}d} \\
&= \frac{Pr\{\text{neighbor}|\text{degree}=d\}f(d)}{\int_M^k Pr\{\text{neighbor}|\text{degree}=d\}f(d)\mathrm{d}d} \\
&= \frac{\frac{dk}{2Mt}f(d)}{\int_M^k \frac{dk}{2Mt}f(d)\mathrm{d}d} \\
&= \frac{df(d)}{\int_M^k df(d)\mathrm{d}d} \\
&= \frac{1/d^2}{\int_M^k 1/d^2\mathrm{d}d} \\
&= \frac{1/d^2}{d^{-1}|_k^M} \\
&= \frac{1/d^2}{1/M-1/k} \\
&= \frac{kM}{d^2(k-M)}, \qquad (101)
\end{aligned}
$$

which is time-invariant. Redenote $f_k^t(d)$ by $f_k^i(d)$.

$$
\begin{aligned}
F_k^i(d) &= \int_M^d f_k^i(y)dy \\
&= \frac{k(d-M)}{d(k-M)} \qquad (102)
\end{aligned}
$$

One can verify that $F_k^i(M)=0$ and $F_k^i(k)=1$, so that $F_k^i(\cdot)$ is a legitimate distribution on $[M,k)$. Also notice that $1-F_k^o(d) > 1-F_k^i(d)\,\forall d$, which indicates that one's out-neighbors have a stochastically higher degree distribution than in-neighbors do. Note that $1-F_k^i(d)$ increases with M. That means one's in-neighbor's degree distribution first-order stochastically increases in M. □

Proof of Proposition 8. We have

$$
\begin{aligned}
\mathbb{E}[\sum_{j\in N_k} x(j)|k] &= \lambda M\int_k^N x(d)f_k^o(d)\mathrm{d}d + (1-\lambda)(k-M)\int_M^k x(d)f_k^i(d)\mathrm{d}d \\
&= \lambda M\int_k^N x(d)\frac{k}{d^2}\mathrm{d}d + (1-\lambda)(k-M)\int_M^k x(d)\frac{kM}{d^2(k-M)}\mathrm{d}d \\
&= \lambda kM\int_k^N \frac{x(d)}{d^2}\mathrm{d}d + (1-\lambda)kM\int_M^k \frac{x(d)}{d^2}\mathrm{d}d \\
&= kM\left(\lambda\int_k^N \frac{x(d)}{d^2}\mathrm{d}d + (1-\lambda)\int_M^k \frac{x(d)}{d^2}\mathrm{d}d\right) \qquad (103)
\end{aligned}
$$

In order for $\mathbb{E}[\sum_{j\in N_k} x(j)|k]$ increasing in k, it suffices to have

$$\begin{aligned}
\frac{d}{dk}\left\{kM\int_k^N \frac{x(d)}{d^2}\mathrm{d}d\right\} &= M\left(\int_k^N \frac{x(d)}{d^2}\mathrm{d}d - \frac{x(k)}{k}\right)\\
&= M\left(-\int_k^N x(d)\mathrm{d}\frac{1}{d} - \frac{x(k)}{k}\right)\\
&\underbrace{=}_{\text{Integration by part}} M\left(\int_k^N \frac{1}{d}x'(d)\mathrm{d}d - \frac{x(d)}{d}\Big|_k^N - \frac{x(k)}{k}\right)\\
&\underbrace{=}_{N\to\infty} M\int_k^N \frac{1}{d}x'(d)\mathrm{d}d\\
&> 0 \ \text{ if } x'(\cdot) > 0
\end{aligned}$$

Therefore, the single crossing condition is met if the consumption strategy is increasing in degree. □

Proof of Theorem 3. The proof is analogous to that of Theorem 1, and we recombine it into two steps. Step (i): We explore the structural properties of the optimization problem. Step (ii): We rewrite the firm objective as a function of consumption only. Using the change of variables, we show that the objective can be optimized by calculus of variations. We then verify that the candidate solution from calculus of variations indeed satisfies the remaining constraints.

Step (i): Structural properties of the optimization problem. As the standard mechanism design approach, we first reduce the constraints of incentive compatibility and individual rationality by single crossing property. Recall the payoff function of a degree-k customer reporting $\hat{k}$ when others report their own types is

$$\pi_k(\hat{k}, x(\cdot), P(\cdot)) = ax(\hat{k}) - bx^2(\hat{k}) + x(\hat{k})\delta\mathbb{E}[\sum_{j\in N_k} x(j)|k] - P(x(\hat{k})),$$

The first-order condition for IC constraints implies:

$$\begin{aligned}
& d\pi_k(\hat{k}, x(\cdot), P(\cdot))/d\hat{k}\Big|_{\hat{k}=k}\\
= \ & ax'(\hat{k}) - 2bx(\hat{k})x'(\hat{k}) + x'(\hat{k})\delta\mathbb{E}[\sum_{j\in N_k} x(j)|k] - P'(x(\hat{k}))x'(\hat{k})\Big|_{\hat{k}=k}\\
= \ & \left[a - 2bx(\hat{k}) + \delta\mathbb{E}[\sum_{j\in N_k} x(j)|k] - P'(x(\hat{k}))\right]x'(\hat{k})\Big|_{\hat{k}=k}\\
= \ & \left[a - 2bx(k) + \delta\mathbb{E}[\sum_{j\in N_k} x(j)|k] - P'(x(k))\right]x'(k)\\
= \ & 0.
\end{aligned} \tag{104}$$

So we have $a - 2bx(k) + \delta\mathbb{E}[\sum_{j\in N_k} x(j)|k] = P'(x(k))$, which implies

$$\frac{d}{dk}P'(x(k)) = -2bx'(k) + \delta\frac{d}{dk}\mathbb{E}[\sum_{j\in N_k} x(j)|k] \tag{105}$$

The local second-order condition for IC constraints is

$$\begin{aligned}
& \frac{d^2}{d\hat{k}^2}\pi_k(\hat{k}, x(\cdot), P(\cdot))\Big|_{\hat{k}=k} \\
= & \left[-2bx'(\hat{k}) - \frac{d}{d\hat{k}}P'(x(\hat{k}))\right]x'(\hat{k}) + \left[a - 2bx(\hat{k}) + \delta\mathbb{E}[\sum_{j\in N_k} x(j)|k] - P'(x(\hat{k}))\right]x''(\hat{k})\Big|_{\hat{k}=k} \\
\underbrace{=}_{(104)} & \left[-2bx'(k) - \frac{d}{dk}P'(x(k))\right]x'(k) \\
\underbrace{=}_{(105)} & \left[-\delta\frac{d}{dk}\mathbb{E}[\sum_{j\in N_k} x(j)|k]\right]x'(k) \\
< & \ 0 \text{ if } x'(k) > 0
\end{aligned} \tag{106}$$

Therefore the local concavity for truth reporting requires the monotonicity condition, $x'(k) > 0$, which in our problem also leads to $\frac{d}{dk}\mathbb{E}[\sum_{j\in N_k} x(j)|k] > 0$ by Proposition 8. (104) and (106) constitute the local IC condition, under which the customer does not attempt to lie locally. We will soon show that the single crossing condition, justified in Proposition 8, extends local IC to global. Define the payoff in the truth telling equilibrium:

$$V(k) := \pi_k(k, x(\cdot), P(\cdot)) = ax(k) - bx^2(k) + x(k)\delta\mathbb{E}[\sum_{j\in N_k} x(j)|k] - P(x(k)), \tag{107}$$

and IR constraints (6) imply that $V(k) \geq 0, \forall k$.

$$\begin{aligned}
\frac{d}{dk}V(k) &= ax'(k) - 2bx(k)x'(k) + \frac{dx(k)\delta\mathbb{E}[\sum_{j\in N_k} x(j)|k]}{dk} - P'(x(k))x'(k) \\
&= ax'(k) - 2bx(k)x'(k) + x'(k)\delta\mathbb{E}[\sum_{j\in N_k} x(j)|k] + \\
&\quad x(k)\delta\frac{d}{dk}\mathbb{E}[\sum_{j\in N_k} x(j)|k] - P'(x(k))x'(k) \\
&\underbrace{=}_{(104)} x(k)\delta\frac{d}{dk}\mathbb{E}[\sum_{j\in N_k} x(j)|k],
\end{aligned}$$

The above result can be reached by the envelope theorem as well, in consideration of the IC constraints:

$$\frac{d}{dk}V(k) = \frac{\partial}{\partial k}\pi_k(\hat{k}^*, x(\cdot), P(\cdot))\Big|_{\hat{k}^*=k} = x(k)\delta\frac{d}{dk}\mathbb{E}[\sum_{j\in N_k} x(j)|k].$$

Accordingly, (from Fundamental Theorem of Calculus)

$$V(k) = V(\underline{k}) + \int_{\underline{k}}^{k} x(u)\delta\frac{d}{du}\mathbb{E}[\sum_{j\in N_u} x(j)|u]du,$$

where $\underline{k}$ is the lowest degree type considered. One can write

$$
\begin{aligned}
& V(k) - \pi_k(\hat{k}, x(\cdot), P(\cdot)) \\
= \quad & V(k) - \left(V(\hat{k}) - x(\hat{k})\delta\mathbb{E}[\sum_{j\in N_{\hat{k}}} x(j)|\hat{k}] + x(\hat{k})\delta\mathbb{E}[\sum_{j\in N_k} x(j)|k] \right) \\
= \quad & V(k) - V(\hat{k}) + x(\hat{k})\delta \left(\mathbb{E}[\sum_{j\in N_{\hat{k}}} x(j)|\hat{k}] - \mathbb{E}[\sum_{j\in N_k} x(j)|k] \right) \\
= \quad & \int_{\hat{k}}^{k} x(u)\delta\frac{d}{du}\mathbb{E}[\sum_{j\in N_u} x(j)|u]du \\
& + x(\hat{k})\delta \left(\mathbb{E}[\sum_{j\in N_{\hat{k}}} x(j)|\hat{k}] - \mathbb{E}[\sum_{j\in N_k} x(j)|k] \right) \\
\underbrace{=}_{\text{Integration by parts}} \quad & x(k)\delta\mathbb{E}[\sum_{j\in N_k} x(j)|k] - \int_{\hat{k}}^{k} \mathbb{E}[\sum_{j\in N_u} x(j)|u]\delta x'(u)du \\
& + x(\hat{k})\delta \left(-\mathbb{E}[\sum_{j\in N_k} x(j)|k] \right) \\
= \quad & (x(k) - x(\hat{k}))\delta\mathbb{E}[\sum_{j\in N_k} x(j)|k] - \int_{\hat{k}}^{k} x'(u)\delta\mathbb{E}[\sum_{j\in N_u} x(j)|u]du \\
= \quad & \int_{\hat{k}}^{k} x'(u)\delta\mathbb{E}[\sum_{j\in N_k} x(j)|k]du - \int_{\hat{k}}^{k} x'(u)\delta\mathbb{E}[\sum_{j\in N_u} x(j)|u]du \\
\geq \quad & 0,
\end{aligned}
$$

if $\hat{k} < k$, since $x'(\cdot) > 0$ (which also means $\mathbb{E}[\sum_{j\in N_k} x(j)|k]$ increases in k). When $\hat{k} > k$, simply rewrite the above as

$$
\begin{aligned}
& V(k) - \pi_k(\hat{k}, x(\cdot), P(\cdot)) \\
= \; & \int_{k}^{\hat{k}} x'(u)\delta\mathbb{E}[\sum_{j\in N_u} x(j)|u]du - \int_{k}^{\hat{k}} x'(u)\delta\mathbb{E}[\sum_{j\in N_k} x(j)|k]du \\
\geq \; & 0,
\end{aligned}
$$

where the last inequality follows from $x'(\cdot) > 0$ (which also means $\mathbb{E}[\sum_{j\in N_k} x(j)|k]$ increases in k). Thus IC is achieved globally. Since $V(\underline{k}) = 0$, the payment $P(x(k))$ is

$$
P(x(k)) = ax(k) - bx^2(k) + x(k)\delta\mathbb{E}[\sum_{j\in N_k} x(j)|k] - \int_{\underline{k}}^{k} x(u)\delta\frac{d}{du}\mathbb{E}[\sum_{j\in N_u} x(j)|u]du. \tag{108}
$$

As a result, IC and IR constraints (5)-(6) can reduce to a single monotonicity constraint $x'(\cdot) > 0$.

Step (ii): **Rewriting the firm's objective**. The firm's problem can be transformed as

$$\max_{x(\cdot)} \quad N\int_{\underline{k}}^{N} \left[\begin{array}{c} (a-c)x(k) - bx^2(k) \\ +x(k)\delta \left\{ \begin{array}{c} \mathbb{E}[\sum_{j\in N_k} x(j)|k] \\ -H(k)\frac{d}{dk}\mathbb{E}[\sum_{j\in N_k} x(j)|k] \end{array} \right\} \end{array} \right] f(k)dk \tag{109}$$

$$\text{s.t.} \quad x(\cdot) \text{ is increasing.}$$

Observe

$$\begin{aligned} \frac{d}{dk}\mathbb{E}[\sum_{j\in N_k} x(j)|k] &= \frac{\mathbb{E}[\sum_{j\in N_k} x(j)|k]}{k} + kM\left(-\lambda\frac{x(k)}{k^2} + (1-\lambda)\frac{x(k)}{k^2}\right) \\ &= \frac{\mathbb{E}[\sum_{j\in N_k} x(j)|k]}{k} + M(1-2\lambda)\frac{x(k)}{k} \qquad (110) \\ \frac{d^2}{dk^2}\mathbb{E}[\sum_{j\in N_k} x(j)|k] &= \frac{M(1-2\lambda)}{k}x'(k) \end{aligned}$$

Therefore

$$k\frac{d}{dk}\mathbb{E}[\sum_{j\in N_k} x(j)|k] - \mathbb{E}[\sum_{j\in N_k} x(j)|k] = M(1-2\lambda)x(k) \tag{111}$$

$$x(k) = \frac{1}{M(1-2\lambda)}\left(k\frac{d}{dk}\mathbb{E}[\sum_{j\in N_k} x(j)|k] - \mathbb{E}[\sum_{j\in N_k} x(j)|k]\right) \tag{112}$$

Let

$$G := \left[\begin{array}{c} (a-c)x(k) - bx^2(k) \\ +x(k)\delta \left\{ \begin{array}{c} \mathbb{E}[\sum_{j\in N_k} x(j)|k] \\ -H(k)\frac{d}{dk}\mathbb{E}[\sum_{j\in N_k} x(j)|k] \end{array} \right\} \end{array} \right] f(k) \tag{113}$$

and $z(k) := \mathbb{E}[\sum_{j\in N_k} x(j)|k]$. Thus G can be expressed as a function of $z(k)$ and $z'(k)$, which makes the transformed objective (109) readily solvable by calculus of variation.

Substituted with (112), (113) becomes a function of $G(k, z(k), z'(k))$. The Eular equation

$$G_{z(k)}(k, z(k), (z)'(k)) = \frac{d}{dk}G_{(z)'(k)}(k, z(k), (z)'(k)).$$

yields a surprisingly simple formula (in x-terms):

$$x^*(k) - kx^{*\prime}(k) = \frac{a-c}{2b + \delta(1-2\lambda)M}. \tag{114}$$

This means that the induced consumption is linear in degree, i.e.

$$x^*(k) = \theta k + \frac{a-c}{2b + \delta(1-2\lambda)M}, \tag{115}$$

where θ is a positive real number (to satisfy the monotonicity constraint $x'(\cdot) > 0$). Also note that Assumption 1 implies $2b > \delta M$, which ensures the non-negativity of the consumption regardless of $\lambda \in [0, 1]$ and k.

The second order condition, known as Legendre condition, is satisfied:

$$G_{z'(k)z'(k)}(k, z(k), z'(k)) = -\frac{2(2b + \delta(1-2\lambda)M)}{k(1-2\lambda)^2} \leq 0, \ \forall k, z(k).$$

Plugging $x^*(\cdot)$ into the payment function (108) gives the optimal payment scheme $P^*(\cdot)$. Similar to Theorems 1 and 2, it can be shown that $x^*(\cdot)$ can be uniquely implemented by $P^*(\cdot)$. □

Proof of Proposition 9. Recall that

$$\begin{aligned} P^{*\prime\prime}(x^*(k)) &= \frac{\mathrm{d}P'(x^*(k))}{\mathrm{d}k}/x^{*\prime}(k) \\ &= -2b + \delta\frac{\mathrm{d}}{\mathrm{d}k}\mathbb{E}[\sum_{j\in N_k} x^*(j)|k]/x^{*\prime}(k). \end{aligned} \tag{116}$$

Substituted with $\frac{d}{dk}\mathbb{E}[\sum_{j\in N_k} x(j)|k] = \frac{\mathbb{E}[\sum_{j\in N_k} x(j)|k]}{k} + M(1-2\lambda)\frac{x(k)}{k} = \frac{1}{k}\left(\mathbb{E}[\sum_{j\in N_k} x(j)|k] + M(1-2\lambda)x(k)\right)$ and $x^{*\prime}(k) = \theta$, it yields

$$\begin{aligned} P^{*\prime\prime}(x^*(k)) &= -2b + \delta\frac{1}{k\theta}\left(\mathbb{E}[\sum_{j\in N_k} x^*(j)|k] + M(1-2\lambda)x^*(k)\right) \\ &= -2b + \frac{1}{k\theta}\left(\delta\mathbb{E}[\sum_{j\in N_k} x^*(j)|k] + \delta M(1-2\lambda)x^*(k)\right) \\ &> \frac{1}{k\theta}\left(2b(x^*(k) - k\theta) - (a-c) + \delta M(1-2\lambda)x^*(k)\right) \\ &\underbrace{=}_{(25)} \frac{1}{k\theta}\left(2b\left(\frac{a-c}{2b+\delta(1-2\lambda)M}\right) - (a-c) + \delta M(1-2\lambda)x^*(k)\right) \\ &\underbrace{=}_{(25)} \frac{\delta(1-2\lambda)M}{k\theta}\left(\left(\frac{-(a-c)}{2b+\delta(1-2\lambda)M}\right) + x^*(k)\right) \\ &= \delta(1-2\lambda)M \end{aligned}$$

The inequality “>” above stems from the fact that $x^*(k) < \frac{1}{2b}\{a - c + \delta\mathbb{E}[\sum_{j\in N_k} x^*(j)|k]\} \Leftrightarrow \delta\mathbb{E}[\sum_{j\in N_k} x^*(j)|k] > 2bx^*(k) - (a-c)$. Therefore, we can conclude that, if $\lambda < 1/2$, $P^{*\prime\prime}(x^*(k)) > 0$ (quantity premium). Since $x^{*\prime}(\cdot) > 0$, we also have $\frac{\mathrm{d}P^{*\prime}(x^*(k))}{\mathrm{d}k} > 0$.

When $\lambda > 1/2$,

$$\begin{aligned} P^{*\prime\prime}(x^*(k)) &= -2b + \delta\frac{1}{k\theta}\left(\mathbb{E}[\sum_{j\in N_k} x^*(j)|k] + M(1-2\lambda)x^*(k)\right) \\ &< -2b + \delta M(1-2\lambda) + \frac{1}{k\theta}\left(\delta\mathbb{E}[\sum_{j\in N_k} x^*(j)|k]\right) \\ &< -2b + \delta M(1-2\lambda) + \frac{1}{\theta}\delta x^*(\bar{k}) \\ &\underbrace{=}_{(25)} -2b + \delta M(1-2\lambda) + \delta\left(\bar{k} + \frac{a-c}{\theta(2b+\delta(1-2\lambda)M)}\right) \\ &= -2b + \delta\bar{k} + \delta M(1-2\lambda) + \delta\frac{a-c}{\theta(2b+\delta(1-2\lambda)M)} \end{aligned}$$

Given Assumption 1, $P^{*\prime\prime}(x^*(k)) < 0$ (quantity discount) holds so long as $M(1-2\lambda) + \frac{a-c}{\theta(2b+\delta(1-2\lambda)M)} < 0$. Since $x'(k) > 0$, we also have in this case $\frac{\mathrm{d}P^{*\prime}(x^*(k))}{\mathrm{d}k} < 0$. □

Proof of Corollary 3. Recall that of first best,

$$G^{FB} := \left[\begin{array}{c} (a-c)x(k) - bx^2(k) \\ +x(k)\delta\left\{\ \mathbb{E}[\sum_{j\in N_k} x(j)|k]\ \right\} \end{array}\right] f(k). \tag{117}$$

and G defined in (113), so that

$$\begin{aligned} G^{FB} - G &= \left[x(k)\delta H(k)\frac{d}{dk}\mathbb{E}[\sum_{j\in N_k} x(j)|k]\right] f(k) \\ &= x(k)\delta(1-F(k))\frac{d}{dk}\mathbb{E}[\sum_{j\in N_k} x(j)|k]. \end{aligned} \tag{118}$$

For the expression of $G^{FB} - G$, the coefficient for $x(k)$ is positive given that $\mathbb{E}[\sum_{j\in N_k} x(j)|k]$ increases in k (for any feasible $x'(\cdot) > 0$). Thus the incentive for consumption is lower in second best case than that in first best scenario. That implies downward distortion of consumption in the second best case.

To see that the profit earned from degree-k customer increases in k, recall

$$\begin{aligned} P(x(k)) &= ax(k) - bx^2(k) + x(k)\delta\mathbb{E}[\sum_{j\in N_k} x(j)|k] - \int_{\underline{k}}^{k} x(u)\delta\frac{d}{du}\mathbb{E}[\sum_{j\in N_u} x(j)|u]du \\ \frac{d}{dk}P(x(k)) &= \left(a - 2bx(k) + \delta\mathbb{E}[\sum_{j\in N_k} x(j)|k]\right) x'(k) \end{aligned}$$

Therefore for $\Pi(k) := P(x(k)) - cx(k)$,

$$\begin{aligned} \Pi'(k) &= \frac{d}{dk}P(x(k)) - cx'(k) \\ &= \left(a - c - 2bx(k) + \delta\mathbb{E}[\sum_{j\in N_k} x(j)|k]\right) x'(k) \end{aligned}$$

Notice that $x^{*\prime}(k) > 0$ and $a - c - 2bx^*(k) + \delta\mathbb{E}[\sum_{j\in N_k} x^*(j)|k] > 0$ (since $x^*(k) < \bar{x}(k)$). Thus $\Pi^{*\prime}(k) > 0$.

That $x^*(k)$ increases in k directly follows from the monotonicity constraint. Next we show how the consumption and the profit change with the network density. Since the average degree in the network is $2M$, the parameter M captures the density of network. Recall

$$x^*(k) = \theta k + \frac{a-c}{2b+\delta(1-2\lambda)M}, \tag{119}$$

If $1-2\lambda > 0$, the induced consumption $x^*(k)$ declines in M. If $1-2\lambda < 0$, $x^*(k)$ increases in M.

To show that the firm's maximum profit increases with M if $1-2\lambda < 0$, we need more notations. Let $\tilde{x}(k) := \theta k$; and note $\tilde{x}(k)$ is independent of M and $x^*(k) = \tilde{x}(k) + \frac{a-c}{2b+\delta(1-2\lambda)M}$. Define $\xi(k) := kM\frac{a-c}{2b+\delta(1-2\lambda)M}\left(\lambda\int_k^N \frac{1}{d^2}dd + (1-\lambda)\int_M^k \frac{1}{d^2}dd\right) = \frac{a-c}{2b+\delta(1-2\lambda)M}\left[(2\lambda-1)M+(1-\lambda)k\right]$, which increases in M given $1-2\lambda < 0$. It follows that $\mathbb{E}[\sum_{j\in N_k} x^*(j)|k] = \mathbb{E}[\sum_{j\in N_k}\tilde{x}(j)|k] + \xi(k)$. Observe that

$$\begin{aligned}
\frac{d}{dM}\mathbb{E}[\sum_{j\in N_k}\tilde{x}(j)|k] &= \frac{\mathbb{E}[\sum_{j\in N_k}\tilde{x}(j)|k]}{M} - k(1-\lambda)\frac{\tilde{x}(M)}{M} \\
&= \frac{k}{M}\left[\mathbb{E}[\tilde{x}(j)|_{j\in N_k}k] - (1-\lambda)\tilde{x}(M)\right] \\
&> 0, \text{ given } \tilde{x}(\cdot) \text{ increasing and } M \text{ the lowest degree}
\end{aligned}$$

Thus, $\frac{d}{dM}\mathbb{E}[\sum_{j\in N_k} x^*(j)|k] = \frac{d}{dM}\mathbb{E}[\sum_{j\in N_k}\tilde{x}(j)|k] + \frac{d}{dM}\xi(k) > 0$. Also recall that

$$\begin{aligned}
\delta\mathbb{E}[\sum_{j\in N_k} x^*(j)|k] &> 2bx^*(k) - (a-c) \\
&= 2b\left(\theta k + \frac{a-c}{2b+\delta(1-2\lambda)M}\right) - (a-c) \\
&> 2b\theta k \text{ since } 1-2\lambda < 0
\end{aligned} \tag{120}$$

$$\begin{aligned}
\Pi_0(k) &= (a-c)x(k) - bx^2(k) + x(k)\delta\left\{\begin{array}{c}\mathbb{E}[\sum_{j\in N_k} x(j)|k] \\ -H(k)\frac{d}{dk}\mathbb{E}[\sum_{j\in N_k} x(j)|k]\end{array}\right\} \\
&= (a-c)x(k) - bx^2(k) + x(k)\delta\left\{\begin{array}{c}\mathbb{E}[\sum_{j\in N_k} x(j)|k] \\ -\frac{1}{2}\left(\mathbb{E}[\sum_{j\in N_k} x(j)|k] + M(1-2\lambda)x(k)\right)\end{array}\right\} \\
&= (a-c)x(k) - bx^2(k) + \frac{1}{2}x(k)\delta\left(\mathbb{E}[\sum_{j\in N_k} x(j)|k] - M(1-2\lambda)x(k)\right)
\end{aligned}$$

At optimum,

$$\begin{aligned}
\Pi_0^*(k) &= (a-c)x^*(k) + x^*(k)\left[\frac{1}{2}\delta\mathbb{E}[\sum_{j\in N_k} x^*(j)|k] - \left(b + \frac{1}{2}\delta M(1-2\lambda)\right)x^*(k)\right] \\
&\underbrace{=}_{(119)} x^*(k)\left[\frac{1}{2}\delta\mathbb{E}[\sum_{j\in N_k} x^*(j)|k] - \left(b + \frac{1}{2}\delta M(1-2\lambda)\right)\theta k + \frac{1}{2}(a-c)\right] \\
&= \frac{1}{2}x^*(k)\left[\delta\mathbb{E}[\sum_{j\in N_k} x^*(j)|k] - (2b + \delta M(1-2\lambda))\,\theta k + (a-c)\right]
\end{aligned}$$

Note that both $x^*(k)$ and the bracketed term $\delta\mathbb{E}[\sum_{j\in N_k} x^*(j)|k] - (2b+\delta M(1-2\lambda))\,\theta k + (a-c)$ are positive (given (120)) and increase in M. Hence, $\Pi_0^*(k)$ increases in M. Furthermore, since $F(\cdot)$ first-order stochastically increases with M and $\Pi_0^*(k)$ increases in k (as shown above), the firm's overall profit rises with the network density M at optimum. □