Impact of Network Structure on New Service Pricing

Saed Alizamir  
Yale School of Management, New Haven, CT 06511, saed.alizamir@yale.edu

Ningyuan Chen  
Department of Industrial Engineering & Decision Analytics, The Hong Kong University of Science and Technology,  
Hong Kong, nychen@ust.hk

Sang-Hyun Kim  
Yale School of Management, New Haven, CT 06511, sang.kim@yale.edu

Vahideh Manshadi  
Yale School of Management, New Haven, CT 06511, vahideh.manshadi@yale.edu

We analyze a firm’s optimal pricing of a new service when consumers interact in a network and impose positive externality on one another. The firm initially provides its service for free, leveraging network externality to promote rapid service consumption growth. The firm raises the price and starts earning revenue only when a sufficient level of consumer interactions is established. Incorporating the local network effects in a non-stationary dynamic model, we study the impact of network structure on the firm’s revenue and optimal pricing decision. We find that the firm delays the timing of service monetization when it faces a more strongly connected network, despite the fact that such a network allows the firm to monetize the service sooner by resulting in faster consumption growth. We also find that the firm benefits from network imbalance, i.e., the firm prefers a network of consumers with varying degrees of connections to that with similar degrees of connections. We also study the value of knowing the network structure and discuss how such knowledge impacts the firm’s profitability.

Key words: local network effects, network externality, service pricing

1. Introduction

Aided by technologies that enable instant interpersonal communications, an increasing number of service providers generate revenue by monetizing consumer interactions on their platforms. For instance, the revenue of Dropbox, one of the leading providers of online file sharing services, is directly tied to the degree of consumer interactions on the company’s platform, as an individual customer’s consumption of the service increases with her collaborators’ consumption. The success of such interaction-based services rely critically on an existing social network of consumers, established through personal and professional connections. As such, it is important to understand how the structure of a social network impacts the firms’ decisions on service pricing. The pricing problem is
especially challenging for newly-introduced services, since network-driven consumption growth is highly dependent on initial pricing; a firm has to balance between short-term consumption growth and long-term revenue generation. In this paper, we study this problem of initial service pricing, developing a model based on the constructs of graph theory.

A well-known “chicken-and-egg” problem exists in this setting. That is, in order to develop and sustain a network-driven revenue generation, a firm has to build up a robust consumer interactions through its service, but individual consumers will increase their consumption only if there is already a significant level of consumption among their peers (Caillaud and Jullien 2003). To overcome this challenge, firms often employ the penetration pricing strategy: a new service is given away for free initially in order to encourage the consumers to develop the habit of using the service, before phasing into a revenue generation period (Dean 1950). Thus, a firm focuses sequentially on rapid consumption growth followed by monetization, effectively subsidizing interaction-based consumption first and then raising the price only when a critical level of consumption is attained. For example, a Chinese social media company Tencent offered its peer-to-peer money transfer service WeChat Pay free of charge for the first two years before price increase in 2016 (Brennan 2016).

While this pricing strategy is intuitive and widely adopted in practice as a rule of thumb, less understood is how to devise the optimal strategy for different types of consumer networks. Clearly, networks are not created equal; there are varying degrees of complexities in networks, and how consumers interact in these networks can have a significant effect on a firm’s ability to profit from it. A number of interesting research questions arise: Should a firm plan on a lengthy initial “giveaway” period if a network is concentrated around only a few consumers? At what price should a firm charge consumers to best exploit the network effects? Does a firm benefit from having complete visibility into the structure of consumer network?

To answer these questions, we develop a model that captures a firm’s intertemporal pricing decision while taking into account the service consumption growth through local network effects. We consider a monopolistic firm that offers a new service over an infinite time horizon to a network of consumers, who determine their optimal service consumption based on their interactions with other consumers. Consumers are heterogeneous in their level of interactions in the network, and as such, disparity among consumers plays an important role in consumption dynamics. Facing the evolution of service consumption trajectory over time, the firm solves an optimal stopping problem to determine both the timing and magnitude of price increase in order to maximize its total discounted revenue.

It is important to distinguish our work from other studies that examine adoption of durable goods under network effects (see Section 2 for references). The non-durable nature of services
studied in our paper introduces new dynamics that have received little attention in the literature and differ from those governing the diffusion of durable goods on networks (e.g., product adoption through word-of-mouth). In particular, a firm’s pricing strategy in models that focus on durable goods is aimed at converting non-users into users by influencing the consumers’ one-time purchase decision; once a consumer is converted into a user, any subsequent change in price will not impact that consumer’s decision. Hence, network interactions for durable goods act primarily as a medium for spreading awareness of the product. By contrast, in many common service settings, users can increase or decrease the level of consumption over time as long as they have access to the service, i.e., a consumer makes repeated consumption decisions. This implies that a price change affects all users simultaneously, influenced by the network externality that evolves over time. Our model captures this interplay in a non-stationary dynamic program framework.

Our analysis reveals the following insights. First, a nontrivial relationship exists between the strength of network connection and the timing of price increase. We find that it is optimal for a firm to delay price increase when it faces a more strongly connected network. This is despite the fact that a high level of consumer interactions on such networks expedites consumption growth and hence enables the firm to monetize its service sooner. Second, keeping the total level of interactions in the network fixed, the firm generates the least possible revenue when it faces a balanced network in which all consumers have equally strong network connections. Further, for a large class of networks, the more heterogeneous the consumers’ connection strengths become, the more revenue the firm can generate. Hence, the firm benefits from imbalance in the network: having a few “star consumers” who are strongly connected to the rest of the network is preferred to having many consumers with weak connections. Finally, we find that the firm’s ability to benefit from acquiring consumer-level network information and incorporating it in its pricing decision depends critically on the network imbalance; in a sufficiently balanced network, for instance, the firm will not suffer substantially from lack of such information.

From a methodological standpoint, we adopt several techniques from algebraic graph theory to solve the firm’s high-dimensional dynamic pricing problem and study the impact of network structure on firm’s pricing policy and revenue. Our analysis relates the firm’s pricing policy to the spectral characteristics of the network in which consumers are embedded. In particular, we develop a novel proof technique based on establishing a connection between firm’s revenue and the number/weight of paths of a given length on the underlying network.

We complement our theoretical findings with an extensive numerical study using a dataset obtained from Stack Overflow (an online knowledge-exchange platform for computer programmers) through which we construct a network that exhibits the heterogeneity of user interactions. In this proof-of-concept study, we demonstrate a monetizing opportunity for a platform like Stack
Overflow and quantify the revenue improvement that can be achieved when network effects are explicitly taken into account.

The rest of the paper is organized as follows. In Section 2 we review the related literature and highlight the differences between the current paper and previous works. In Section 3 we formally introduce our modeling framework that incorporates consumers’ consumption decision, the resulting consumption growth dynamics, and the firm’s pricing decision. Section 4 describes the firm’s optimal pricing policy and its structural properties. Section 5 focuses on the impact of network structure on firm’s pricing policy and revenue, discussing our main insights and their practical implications. Building on the analyses in the previous sections, in Section 6 we study the value of obtaining information on the network structure. Section 7 presents our numerical study based on the Stack Overflow dataset, followed by the conclusion in Section 8. Detailed proofs of all formal results are found in the appendix; in the main text of the paper, only the proof ideas of selected results are presented.

2. Related Literature

Managing penetration of a consumer product in the presence of network effects has been an active area of research in operations, marketing, and economics. The most prominent is a stream of literature started by Bass (1969) that considers diffusion of durable goods in the presence of global network effects. Representative articles in the operations and economics literature include Katz and Shapiro (1985), Cabral (2011), Jing (2011), Radner et al. (2014), Shen et al. (2014), Weber and Guérin (2014), Alizamir et al. (2016), Du et al. (2016), Hu et al. (2016), Wang and Wang (2016), Crapis et al. (2017), Papanastasiou and Savva (2017), and Chen and Chen (2017). Some of these papers focus specifically on word-of-mouth communication and learning, the dynamics of which can be interpreted as those arising from network effects. In comparison, we consider a dynamic model for non-durable services that explicitly accounts for the underlying network structure.

In recent years, the role of network structure has received attention as researchers started considering diffusion processes in the presence of local network effects. This line of research is particularly motivated by the emergence of communication tools over which agents interact only with their “neighbors” in social networks, and firm’s access to information about such interactions. Manshadi et al. (2018) consider a diffusion process in random networks and study the impact of degree distribution on diffusion trajectory and the optimal seeding strategy. Campbell (2013) and Ajourlou et al. (2017) study dynamics of word-of-mouth effects in random networks of myopic consumers, and a monopolist’s pricing policy (static and dynamic, respectively) that maximizes firm’s revenue. Hartline et al. (2008) studies a similar dynamic pricing problem for a general network and shows that computing the optimal dynamic pricing policy is NP-hard. More recently, Makhdoumi et al.
(2017) address the problem of dynamic pricing for a firm that sells a durable good to a network of forward-looking consumers. The works of Lobel et al. (2017) and Leduc et al. (2017) are concerned with designing referral incentives to attract consumers and increase revenue. In another direction, Sunar et al. (2015) study jointly optimizing product development, targeting, and pricing of a durable good in the presence of local network effects. Allon and Zhang (2017) investigate service differentiation polices in the presence of heterogenous social network influence among users. Further, Gur et al. (2018) study the competitive facility location problem which has application to the study of the diffusion of competing word-of-mouth campaigns in social networks.

Our work differs from the ones mentioned above in that we do not explicitly model diffusion dynamics, i.e., how the interaction between adopters and non-adopters through word-of-mouth and learning lead to new adoptions over time. Instead, we focus on how consumption of a non-durable service changes through externality in the existing network of users. In our model, a consumer’s utility depends on the current level of consumption by her neighbors, which gives rise to a network externality that evolves over time. For example, a consumer who has used Dropbox service for a few years may suddenly find little use for it if her collaborators stop using the same service. The externality-driven consumption in the network, in turn, influences the firm’s pricing decision.

With optimal pricing as a key feature of the model, our paper is closely related to those that examine static pricing for a firm that offers a divisible good to consumers who enjoy positive local network effects (Ballester et al. 2006, Candogan et al. 2012, Fainmesser and Galeotti 2016). These papers consider a two-stage sequential game in which a firm first chooses a price (possibly different for each consumer), and in response consumers simultaneously decide on their consumption levels. Candogan et al. (2012) assume a complete information setting and analyze the unique Nash equilibrium of this network game. Fainmesser and Galeotti (2016), on the other hand, focus on random networks and incomplete information setting, and explore the value of knowing the details of the random network. Cohen and Harsha (2017) investigate a closely related static pricing problem assuming a more general form of network effects (influence). While the insights gained from these static settings are valuable, they do not address temporal price adjustments commonly observed in practice. Further, a common assumption that consumers make optimal decisions with perfect knowledge of the entire network may not be realistic, due to privacy restrictions as well as consumers’ limited cognitive and computational capability. In this paper, we inject more realism by considering a dynamic setting in which each consumer is only aware of the identities of her immediate neighbors and bases her consumption decision on her observation of the neighbors’ past consumption levels. As such, consumers do not have perfect foresight in our model; similar forms of bounded rationality have been considered in the literature, including Radner et al. (2014). A
related issue of a firm’s incomplete visibility to the network structure is considered in Section 6, in line with the analyses by Candogan et al. (2012) and Fainmesser and Galeotti (2016).

The dynamics of service consumption growth considered in our paper resemble those of epidemics spread in networks (Borgs et al. 2010, Drakopoulos et al. 2014). The main difference is that the goal of epidemics control is to slow down or stop the spread (of a virus), whereas in our problem context a firm aims at accelerating service consumption.

In closing this section, we remark that local network effects need not to be positive. For instance, a luxury product may only be valuable to a consumer if she is the unique owner of such a product in her circle of friends. In this paper, we do not address issues arising from negative network effects; for analyses of such problems, see recent papers by Momot et al. (2016) and Belloni et al. (2017).

3. Model

We model the interaction between a monopolistic service-provider and a network of $n$ individual consumers that unfolds over an infinite horizon. At the beginning of the horizon, the firm (hereafter referred to as “he”) offers a new service, which is made available to all consumers for purchase. Without loss of generality, the firm’s cost of providing the service is constant and normalized to zero. The price of service, however, is time-dependent and allowed to vary as time progresses. Each individual consumer (hereafter referred to as “she”), in turn, determines her consumption amount over time to maximize her utility. Time is discrete and indexed by $t$, such that each period corresponds to a time window over which price and consumption decisions remain constant. The service is divisible and non-storable in the sense that consumption increments are infinitesimal and stockpiling across periods is infeasible. This represents, for instance, online services for which consumption is adjustable in a continuous fashion, and consumers gain their utilities in real time as consumption takes place (hence, utility is not transferred from one period to the next).

3.1. Consumer Utility

Before specifying a consumer’s utility function, we first describe its elements. The utility that consumer $i$ enjoys from the service in period $t$ is increasing and concave in her consumption level $x_{i,t}$, and quasi-linear in service price $p_t$. In line with the extant literature on network effects (e.g., Ballester et al. 2006, Candogan et al. 2012, Fainmesser and Galeotti 2016) and to enable analytical tractability, we adopt a quadratic functional form for the consumer’s utility function, with the negative sign for the quadratic term to induce concavity. The utility function in our model allows for cross-consumer and inter-temporal dependencies, which are respectively captured through network externality and habit persistence, which we describe next.
**Consumer Network Externality.** Consumers in our model are embedded in a network, where their connections represent the interdependency of their service consumptions. In particular, the consumption of service by each consumer imposes externality on the others, the magnitude of which is determined by the specifics of the network. To formalize this, we correspond each consumer to a node in the network, and define the weighted adjacency matrix \( M = [m_{ij}; 1 \leq i \leq n, 1 \leq j \leq n] \), where \( m_{ij} \) represents the weight of the edge connecting nodes \( i \) and \( j \). These weights are assumed to be nonnegative and symmetric, i.e., \( m_{ij} \geq 0 \), and \( m_{ij} = m_{ji} \) for all \( i \neq j \). Consumers \( i \) and \( j \) are said to be neighbors if \( m_{ij} > 0 \). In essence, \( m_{ij} \) is a measure of the influence that consumers \( i \) and \( j \) have on each other, reflecting the intensity of their mutual interaction. Without loss of generality, self-externalities are normalized to zero, i.e., \( m_{ii} = 0 \) for all \( i \). We refer to \( m_{ij} \) as externality weight throughout the paper. The pairwise externality weights are then summarized in matrix \( M \), which is exogenously given and remains constant over time. The magnitude of the externality that consumer \( i \) imposes on consumer \( j \) (and vice versa) is therefore proportional to \( m_{ij} \).

**Habit Persistence.** We adopt the well-known notion of habit persistence introduced by Pollak (1970) to capture the inter-temporal dependence in consumption utility. With habit persistence, each consumer anchors on her last period’s consumption to determine her consumption in the current period. Formally, consumer \( i \)'s utility from consuming \( x_{i,t} \) in period \( t \) depends on her previous consumption amount \( x_{i,t-1} \) through the term \( w_{i,t} = x_{i,t} - \mu x_{i,t-1} \), where \( \mu \in [0,1] \) is the habit coefficient. Parameter \( \mu \) measures the influence of last-period consumption or the strength of “habit,” and is assumed to be the same for all consumers. In the extreme case of \( \mu = 0 \), habit is absent and the consumer’s present utility becomes independent of her last-period consumption. In the other extreme where \( \mu = 1 \), the effect of habit is the strongest since utility is gained only from the additional consumption over that of the last period; maintaining the same level of consumption as in the last period therefore leads to zero utility in this case. Habit persistence can be viewed as the externality that a consumer imposes on herself.

Combining the elements described above, we specify consumer \( i \)'s utility in period \( t \) as a function of her consumption \( x_{i,t} \) as

\[
\begin{align*}
    u_{i,t}(x_{i,t}; x_{i,t-1}, x_{-i,t}) = \alpha \frac{w_{i,t}^2}{2\beta} + \frac{1}{\beta} x_{i,t} \sum_{j=1}^{n} m_{ij} x_{j,t} - p_t x_{i,t} ,
\end{align*}
\]

where \( \alpha \) and \( \beta \) are constant coefficients, \( w_{i,t} = x_{i,t} - \mu x_{i,t-1} \), and \( p_t \) is the unit price in period \( t \). The argument \( x_{-i,t} \) appearing in (1) represents the vector of consumptions by all consumers except for
consumer $i$ in period $t$. (Throughout the paper, bold letters denote vectors.) The first two terms in (1) represent the utility that consumer $i$ earns individually by consuming $x_{i,t}$ units of service under the habit persistence framework. The third term captures the externality between consumer $i$ and her neighbors. Note that the externality generated by consumer $j$ is multiplicative in $m_{ij}$, and grows proportionally with the product of individual consumptions, $x_{i,t}x_{j,t}$. As such, the externality is greater if, all else being equal, two neighbors interact more intensely (larger $m_{ij}$) and/or either of their consumptions goes up (larger $x_{i,t}$ or $x_{j,t}$). Finally, the last term represents consumer $i$’s disutility for paying the firm for service in period $t$. The utility function (1) is specified under the assumptions that the consumers are heterogeneous in their network interactions (different $m_{ij}$’s) but homogeneous in their valuations of service and price sensitivities (same $\alpha$ and $\beta$). We make this simplifying assumption in order to isolate the impact of the network structure.

Note that in evaluating utility function (1), the consumer does not need complete visibility into the structure of the entire network. This is because the network externality extends only to each consumer’s immediate neighbors, and hence the actions by a neighbor’s neighbors do not directly influence the consumer’s decision. This is consistent with the discussion in Section 2 that the consumers in our model have partial information about the network and do not observe the consumption level of non-neighbor consumers.

### 3.2. Consumer’s Consumption Decision

As specified in (1), consumer $i$’s utility in period $t$ depends on the consumption decisions of the other consumers in the same period. Since all consumption decisions are made simultaneously, consumer $i$ does not observe the consumption level of her neighbors in period $t$ before deciding on her own consumption. Thus, for every neighbor $j$, consumer $i$ has to form an expectation about $j$’s consumption decision in the present period, which we denote by $\hat{x}_{i,j,t}$ for all $j$’s with $m_{ij} > 0$. We assume that consumers are non-strategic in the sense that each consumer simply uses her observation from last period’s consumptions to form expectation about present period’s decisions, i.e., $\hat{x}_{j,t} = x_{j,t-1}$. In other words, consumers’ decisions in our framework are responsive rather than forward-looking; consumers rely on the best information available at the time of their decision-making, without attempting to predict others’ decisions. Similar assumptions are commonly found in the literature, especially in the context of consumer decisions under network effects; see, for example, Arthur (1989), Radner et al. (2014), and Hu et al. (2016). (Incorporating forward-looking consumer behavior requires each consumer having complete visibility to the entire network, an assumption inconsistent with our model framework.)
Having formed the expectation, each consumer then decides on her consumption level that maximizes her utility in (1). Differentiating (1) with respect to $x_{i,t}$ and setting it equal to zero with $\hat{x}_{j,t} = x_{j,t-1}$ yields the optimal consumption level

$$x_{i,t}^* = \mu x_{i,t-1} + \alpha - \beta p_t + \sum_{j=1}^{n} m_{ij} x_{j,t-1},$$

(2)

provided that $x_{i,t}^* > 0$. (Part (2) of Assumption 1 that appears below guarantees that the optimal price is less than $\alpha/\beta$ and thus $x_{i,t}^* > 0$, as we prove in Proposition 1.) According to (2), each consumer’s optimal consumption level is increasing in her last-period consumption as well as those of her immediate neighbors. This illustrates how habit persistence and local network externalities create interdependent utilities, resulting in an overall consumption growth over time. The firm, in turn, has to set its inter-temporal pricing policy in a way to optimally leverage on network-driven consumption growth and maximize its long-term revenue.

3.3. Firm’s Pricing Problem

Anticipating the individual consumption decisions as characterized in (2), the firm determines the service price in each period in order to maximize his total discounted revenue over the infinite horizon. In general, this may involve setting a different price for every period, which leads to infinitely many price adjustments over time. In practice, however, frequent price adjustments are not prevalent, especially for services with high price sensitivity. This “price stickiness” phenomenon is well-documented (Netessine 2006, Stokey 2009), and is attributed to factors such as fixed costs of announcing price changes and negative consumer reactions to frequent price fluctuations. Moreover, it has been shown in the literature (e.g., Shen et al. 2014, Ajarlou et al. 2017) that allowing for multiple price adjustments leads to non-monotone and complex pricing patterns, even if only global network effects are considered.

To succinctly capture the infrequency of price adjustments observed in practice and highlight the impact of local network effects, we consider a setting where the firm has only one opportunity to change the price over the entire horizon. The firm starts by charging a low price (in fact, zero) early on since his initial goal is to exploit network externalities and induce rapid consumption growth. We normalize the initial price to zero to simplify the derivations and be consistent with observed penetration pricing practices. Once the overall consumption level reaches a certain threshold and sufficient growth momentum is built, the firm starts monetizing his service by increasing the price. Thereafter, the price remains unchanged. The firm thus determines both timing and magnitude of this one-time price increase with the goal of maximizing his total discounted revenue.

Let $\Pi$ denote the class of all policies with only one price adjustment. Since a policy $\pi \in \Pi$ specifies when and by how much the price should increase, it is represented by the dual $(\tau, p)$ where $\tau$ is
the period in which the firm decides to end the free-offering upon observing the consumption level \( x_t \). Hence, \( p_1 = \ldots = p_r = 0 \) and \( p_{r+1} = p_{r+2} = \ldots = p \) under policy \( \pi \). The firm’s pricing problem is then reduced to

\[
J(x_0) = \max_{\pi \in \Pi} \left\{ \sum_{t=1}^{\infty} \delta^{t+1} p(x_{t+1}, 1) \right\},
\]

where \( \delta \) is the discount factor, and \( J(.) \) is the firm’s optimal value function. The term \( \langle x_t, 1 \rangle = \sum_{i=1}^{n} x_{i,t} \) appearing in (3) represents the total consumption of all consumers in period \( t \), with the symbol \( \langle \cdot, \cdot \rangle \) denoting the inner product of the consumption vector \( x_t \) and the unit vector \( 1 = (1, 1, \ldots, 1) \). We assume that the initial consumption is zero, i.e., \( x_0 = 0 \).

The above formulation implicitly assumes that the firm has perfect knowledge of the network structure, and hence, can infer the exact trajectory of the consumption growth over time. This is reasonable especially for providers of internet-based services, since they can track user activities by directly managing network platforms. In Section 6 we relax this assumption and examine the case in which a firm has imperfect knowledge of the network structure. Furthermore, we impose the following technical conditions on \( M \) and \( \mu \) throughout the paper:

**Assumption 1.** Let \( \lambda_1 \geq \ldots \geq \lambda_n \) be the ordered eigenvalues of adjacency matrix \( M \), and \( \delta \in (0, 1) \) be the firm’s discount factor. We assume:

1. \(-1 - \lambda_n < \mu < 1 - \lambda_1\).
2. \( \delta \geq 2 - \frac{1}{\lambda_1 + \mu} \).

Part (1) of the assumption excludes situations where the network effects are so strong that the firm’s discounted revenue can be unbounded. Part (2) guarantees that no consumer stops using the service under the optimal policy after price increases.

The firm’s optimal policy that solves problem (3) when consumption grows according to (2) is denoted by \( \pi^* = (\tau^*, p^*) \in \Pi \), which we characterize in the next section.

### 4. Optimal Pricing Policy

To characterize the firm’s optimal pricing policy, we first derive the closed-form expression of the per-period consumption vector using the evolution of consumption as specified in (2). Define *adjusted adjacency matrix* \( \tilde{M} = M + \mu I_n \), where \( I_n \) is the \( n \times n \) identity matrix. Note that \( \tilde{M} \) is the same as the adjacency matrix \( M \) defined in Section 3 except that its diagonal elements (zeros) are replaced by the habit coefficient \( \mu \). Then, in lieu of the recursive formula (2), the evolution trajectory of consumptions for any fixed price \( p \leq \alpha / \beta \) can be written succinctly in the following vector form:

\[
x_t | x_{t-1} = (\alpha - \beta p)1 + \tilde{M} x_{t-1},
\]
from which we obtain
\[ \mathbf{x}_t | \mathbf{x}_0 = (\alpha - \beta p)(\mathbf{I}_n - \hat{\mathbf{M}})^{-1}(\mathbf{I}_n - \hat{\mathbf{M}}') \mathbf{1} + \hat{\mathbf{M}}' \mathbf{x}_0. \] (4)

Next, we analyze a special case of problem (3) in which the price is exogenously set. After characterizing the optimal solution for this case, we then extend our analysis to account for the endogenous price and derive an algorithm that can achieve the optimal policy in finite iterations.

### 4.1. Exogenous Price

Suppose the price in (3) is exogenously given. Studying this special case enables us to gain valuable insights into the impact of network characteristics on the timing of price increase. (Although somewhat outside of the scope of our model, this case is relevant in certain practical settings where the price is ex-ante determined by outside factors such as regulations.) With the price fixed, the problem is reduced to an optimal stopping problem in which the only decision variable is the time of terminating the initial free-offering of the service.

With a fixed price over the entire horizon, (4) allows us to express the firm’s total discounted revenue in closed form. To see this, let \( \Gamma(\mathbf{x}; p) \) denote the firm’s total discounted revenue where \( \mathbf{x} = (x_1, \ldots, x_n) \) is the initial consumption vector and a fixed price \( p \) is charged at \( \tau = 0 \). That is,
\[ \Gamma(\mathbf{x}; p) = \sum_{t=1}^{\infty} \delta^t p(\mathbf{x}_t | \mathbf{x}_0 = \mathbf{x}, \mathbf{1}), \]
where \( \mathbf{x}_t \) evolves according to (2) (or equivalently, (4)). The following lemma characterizes this quantity in closed form.

**Lemma 1.** For any given price \( p \leq \alpha/\beta \) and initial consumption vector \( \mathbf{x} \),
\[ \Gamma(\mathbf{x}; p) = \sum_{i=1}^{n} \frac{p(\alpha - \beta p)\delta v_i, 1^2 + p\delta(1 - \delta)\hat{\lambda}_i (v_i, 1)(v_i, x)}{(1 - \delta)(1 - \delta \hat{\lambda}_i)}, \]
where \( \hat{\lambda}_i \) and \( v_i \) are the \( i^{th} \) eigenvalue and eigenvector of matrix \( \hat{\mathbf{M}} \), respectively.

The proof of this lemma relies on several techniques from algebraic graph theory, in particular, spectral decomposition of the adjacency matrix, its algebraic property, and geometric sums of symmetric matrices. Such decomposition enables us to express the discounted revenue \( \Gamma(\mathbf{x}; p) \) in a much simpler form as a summation of \( n \) decoupled parts, each one of which corresponding to one pair of eigenvalue and eigenvector. This simple and compact form of the revenue proves useful in characterizing the optimal policy and studying the impact of network structure. Interested readers are referred to Appendix EC.1 which is dedicated to reviewing preliminary results from algebraic graph theory. The proof of the lemma itself is presented in Appendix EC.2.1.
Using Lemma 1, we can recast the firm’s problem as a discrete-time infinite-horizon optimal stopping problem with \(n\)-dimensional state space that represents the current consumption vector. Let \(J(x; p)\) denote the firm’s optimal value function (revenue-to-go) when current consumption vector is \(x\) and the price has not increased to \(p\) yet. Then, we have the following Bellman equation:

\[
J(x; p) = \max \left\{ \Gamma(x; p), \delta J(\alpha 1 + \tilde{M}x; p) \right\},
\]

in which the first term corresponds to increasing the price now, and the second term corresponds to delaying it for at least one more period.

The state space of the above dynamic program is \(n\)-dimensional, and hence the problem may seem intractable at first glance. Nevertheless, we are able to show that in deriving the optimal policy, the state collapses to one dimension that is a function of current consumption vector as well as the network characteristics. The following theorem uses this approach to characterize the optimality condition for (5) in closed form.

**Theorem 1.** Suppose that price \(p \leq \alpha / \beta\) is exogenously given. Then:

(i) The optimal time to increase price from zero to \(p > 0\) is

\[
\tau^*(p) = \min \left\{ t \in \mathbb{Z}_+ \mid \langle \tilde{M}x_t, 1 \rangle \geq \frac{\sum_{i=1}^{n} \delta \beta p \hat{\lambda}_i (v_i, 1)^2}{1 - \delta \hat{\lambda}_i} - n(\alpha - \beta p) \right\}.
\]

(ii) \(\tau^*(p)\) increases in \(p\).

The characterization of \(\tau^*(p)\) as given by (6) has an intuitive interpretation; it is optimal for the firm to increase the price once the aggregate level of externality among the consumers during the free-offering episode reaches a certain threshold (To be precise, \(\langle \tilde{M}x_t, 1 \rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} m_{ij} x_{j,t} + \mu \sum_{i=1}^{n} x_{i,t}, i.e., it also contains the aggregate impact of habit.) The stopping condition in (6) implies that this aggregate level of externality is a sufficient statistics for the underlying networked dynamical system. In other words, the firm’s optimal timing for price increase remains the same for two different consumption vectors as long as they induce equal aggregate externalities in the network. Note that as long as price is zero, the consumption level for each consumer (and therefore the externality she imposes on her neighbors) follows an increasing trajectory (see Example 1 for an illustration of consumption trajectories). The threshold \(\sum_{i=1}^{n} \frac{\delta \beta p \hat{\lambda}_i (v_i, 1)^2}{1 - \delta \hat{\lambda}_i} - n(\alpha - \beta p)\) appearing in (6) depends on the network structure (through eigenvalues/eigenvectors of its adjusted adjacency matrix) as well as other model parameters.

The second part of Theorem 1 states that the firm delays price increase if a higher price is going to be charged eventually. By extending the duration of the initial free-offering, the firm waits longer and allows the consumers’ service consumption to reach a higher level. While by doing so the firm
forgoes revenue for a longer period of time, such a loss is more than compensated for by the higher
price charged and the larger quantity sold once the free-offering ends.

The following corollary establishes monotonicity of $\tau^*(p)$ with respect to the discount factor, $\delta$, and the price sensitivity, $\beta$.

**Corollary 1.** The optimal time to increase price, $\tau^*(p)$, is increasing in both $\delta$ and $\beta$.

Corollary 1 states that $\tau^*(p)$ increases in $\delta$ and $\beta$, which is intuitive; if the firm values future more (higher $\delta$), then he is willing to wait longer and induce a higher consumption level before increasing the price. Similarly, more price sensitive consumers (higher $\beta$) implies that consumers’ reaction to a price increase would be more severe. This, in turn, drives the firm to delay price increase in order to build up more consumption and consequently a higher level of externality.

In the following, we provide the sketch of the proof for Theorem 1 by outlining the main steps used in the proof.

**4.1.1. Proof Sketch of Theorem 1:** The main step in proving (6) is to show that the one-step look-ahead policy is optimal. The intuition behind the optimality of one-step look-ahead policy is as follows: First, the marginal value from stopping in period $t$ is increasing in consumption levels. To see this, let us revisit the consumption dynamic (given by (2)). In period $t$ and for each consumer $i$, the difference between her consumption under scenarios that the free-offering ends vs. continues is a constant $\beta p$. Therefore, the difference from the revenue that these two scenarios generate from time $t+1$ onwards remains a constant independent of consumption vector $x_t$. On the other hand, if the free-offering ends at time $t$, the firm generates revenue at time $t$ that is increasing in current consumption levels. This, in turn, implies that marginal revenue from stopping in period $t$ increases with the current consumption. These two, together with the fact that the consumption trajectory is increasing in the length of the free-offering episode, imply that if it is optimal to stop the free-offering at time $t$, i.e., $\Gamma(x_t; p) \geq \delta \Gamma(\alpha 1 + \hat{M}x_t; p)$, then it will also be optimal to stop the free-offering at time $s < t$, i.e., $\Gamma(x_s; p) \geq \delta \Gamma(x_{s+1}; p)$.

To formalize the above intuition, note that the Bellman equation implies the optimal stopping time $\tau^*$ as

$$\tau^* = \arg\min_{t \geq 0} \left\{ \Gamma(x_t; p) \geq \delta J(\alpha 1 + \hat{M}x_t; p) \right\}, \quad (7)$$

where $J(x; p)$ is defined in (5). Consider the one-step look-ahead policy:

$$\tau_1 = \arg\min_{t \geq 0} \left\{ \Gamma(x_t; p) \geq \delta \Gamma(\alpha 1 + \hat{M}x_t; p) \right\}.$$

That is, the firm stops the free-offering if the stop-value of the current period is greater than or equal to the stop-value in the next period. We now show that $\tau_1 = \tau^*$. First, we argue that $\tau^* \geq \tau_1$
because \( \Gamma(x; p) \leq J(x; p) \) by their respective definitions. Next, we show \( \tau^* \leq \tau_1 \), i.e., if it is optimal to stop under the one-step look-ahead policy, then it is optimal to stop under the optimal policy as well. It suffices to prove the following statement:

**Claim 1.** If \( \Gamma(x_t; p) \geq \delta \Gamma(x_{t+1}; p) \), then \( J(x_{t+1}; p) = \Gamma(x_{t+1}; p) \).

The proof of the above claim is deferred to Appendix EC.2.2, but it builds on the idea that for any \( t \), if \( \tau_1 \leq t \), then \( \Gamma(x_t; p) \geq \delta \Gamma(x_{t+1}; p) \). The above claim ensures that if \( \Gamma(x_t; p) \geq \delta \Gamma(x_{t+1}; p) \), then we also have \( \Gamma(x_t; p) \geq \delta J(x_{t+1}; p) \), and thus \( \tau^* \leq t \). Therefore, we have shown \( \tau^* \leq \tau_1 \). Combined with our earlier observation that by definition \( \tau^* \geq \tau_1 \), we have: \( \tau^* = \tau_1 \). □

### 4.2. Endogenous Price

Now, we turn to problem (3), which is a generalization of the fixed-price model studied in Section 4.1. Deriving the Bellman equation in this case follows steps similar to the previous section and leads to:

\[
J(x) = \max \left\{ \max_p \{ \Gamma(x; p) \}, \delta J(\alpha 1 + M x) \right\}. \tag{8}
\]

That is, at any point in time, the firm has to decide (i) whether or not to end the initial giveaway period and, (ii) if so, at what level the price should be set. The first term in (8) corresponds to terminating the free-offering and optimizing over price. The second term, on the other hand, corresponds to extending this offering for at least one more period.

Solving for these two decisions simultaneously, in general, presents analytical challenges. Nevertheless, we are able to decompose the problem into subproblems that can be solved separately. To do this, we first assume that the timing of price increase, \( \tau \), is exogenously set, and obtain the optimal price and revenue associated with it. We then repeat these steps for all feasible values of \( \tau \) (to be specified later), and find the optimal policy by comparing their corresponding revenues.

Suppose the duration of the initial free-offering is fixed at \( \tau \). The dynamic program (8) in this case simplifies to

\[
J(x; \tau) = \delta^\tau \max_p \{ \Gamma(x, \tau; x_0 = x; p) \}, \tag{9}
\]

whose solution is characterized as follows.

**Proposition 1.** Suppose the duration of initial free-offering, \( \tau \), is exogenously set. Then, it is optimal for the firm to increase the price to

\[
p^* (\tau) = \alpha B(\tau) \left( 1 + \frac{B(\tau)}{A} \right)
\]

at time \( \tau \), where

\[
A = \sum_{i=1}^{n} \frac{\langle v_i, 1 \rangle^2}{(1 - \delta)(1 - \delta \hat{\lambda}_i)} > 0 \quad \text{and} \quad B(\tau) = \sum_{i=1}^{n} \frac{\hat{\lambda}_i (1 - \hat{\lambda}_i) \langle v_i, 1 \rangle^2}{(1 - \delta \hat{\lambda}_i)(1 - \hat{\lambda}_i)} \geq 0.
\]

Further, \( p^* (\tau) \leq \frac{\alpha}{B} \), for any \( \tau \geq 0 \).
Notice from Proposition 1 that, in the absence of both network effects and habit persistence, i.e., when $m_{ij} = 0$ for $1 \leq i, j \leq n$ and $\mu = 0$, Proposition 1 implies that the optimal price is $\frac{\alpha}{\beta^2}$. This is the case in which periods become independent as the link between consumptions in consecutive periods disappears. Thus, $\frac{\alpha}{\beta^2}$, the usual revenue-maximizing price under the linear demand model, emerges as the optimal price in every period. According to Proposition 1, the presence of network externality effects and habit persistence allows the firm to raise his price from $\frac{\alpha}{\beta^2}$ to $\frac{\alpha}{\beta^2}(1 + \frac{B(\tau_1)}{A})$.

Given Lemma 1 and Proposition 1, we can compute the optimal price and the resulting revenue for any given $\tau$. Then, the optimal stopping time-price pair $(\tau^*, p^*)$ that jointly optimizes the firm's revenue can be found through an exhaustive search over all possible values of $\tau$. The computational burden associated with the search, however, is substantially reduced by employing the following result, which specifies an upper bound on $\tau^*$.

**Proposition 2.** Let $\pi^* = (\tau^*, p^*)$ be the firm’s optimal pricing policy. For any given network, $\tau^* \leq \bar{\tau} = \lceil -\log(4)/\log(\delta) \rceil$.

The above proposition establishes that the upper bound on $\tau^*$, denoted as $\bar{\tau}$, depends only on the discount factor $\delta$, and does not scale with the network size $n$. This allows us to limit the search for the optimal $\tau$ only to the range $0 \leq \tau \leq \bar{\tau}$.

Summarizing the steps described thus far and recalling that $x_0 = 0$, we proceed with our computation by first finding, for any $\tau \in \{0, 1, \ldots, \bar{\tau}\}$,

(i) corresponding $p^*(\tau)$ (given by Proposition 1);

(ii) consumption vector at time $\tau$, i.e., $x_{\tau}$ (as in (4));

(iii) revenue-to-go at time $\tau$, i.e., $\Gamma(x_{\tau}; p^*(\tau))$ (given by Lemma 1);

(iv) corresponding revenue, $J(0; \tau) = \delta^\tau \Gamma(x_{\tau}; p^*(\tau))$.

We then obtain the optimal policy $\pi^* = (\tau^*, p^*)$ by setting $\tau^* = \arg\max_{0 \leq \tau \leq \bar{\tau}} J(0; \tau)$ and $p^* = p^*(\tau^*)$.

We conclude this section with an illustrative example that emphasizes the role of heterogeneity in consumption trajectories for consumers that are exposed to different network effects.

**Example 1.** Consider a network with 5 consumers indexed by $1, \ldots, 5$, as depicted in the left panel of Figure 1. For all pairs $(i, j)$ that are connected, we have $m_{ij} = 0.35$. Other model parameters have the values $\delta = 0.95$, $\mu = 0.1$, and $\alpha = \beta = 0.1$. The optimal policy for this network can be easily obtained using the approach outlined above, resulting in $\tau^* = 3$, $p^* = 0.55$. The trajectory of consumption evolution for each consumer is presented in the right panel of Figure 1. We make the following observations. Consumer 2 has the highest level of interactions in the network as she is connected to the most number of neighbors (i.e., $\sum_{j=1}^{5} m_{ij}$ is the largest for $i = 2$), and consequently, her consumption level dominates that of other consumers at any time. Further, as illustrated in
Figure 1  A numerical example of consumption trajectory in the presence of network externalities.

the right panel, her consumption is least impacted by the price increase immediately after the optimal timing $\tau^* = 3$. By contrast, Consumer 4— who experiences the lowest level of interactions in the network—consumes the least amount in any period compared to other consumers, and her consumption drops significantly after the price increase. □

5. Impact of Network Structure

With the firm’s optimal pricing policy established, we now examine how the policy and its corresponding revenue is impacted by the network characteristics. In particular, we focus on the roles played by the strength (i.e., magnitude) and the heterogeneity of externality weights ($m_{ij}$’s). These two measures represent, respectively, how intensely the consumers influence one another in the network and how this influence varies across consumers. With these measures, we essentially capture the first- and second-order effects of network externality, analogous to the first- and second-moments of a random variable.

To proceed, we first introduce the following notions that help quantify the measures.

**Definition 1.** The weighted degree of consumer $i$ is $D_i = \sum_{j=1}^{n} m_{ij}$, and the average degree for a network is $D = \frac{1}{n} \sum_{i=1}^{n} D_i$. Further, a network is called balanced if $D_i = D$ for all $1 \leq i \leq n$.

Note that the weighted degree $D_i$ is the $i^{th}$ row-sum of the adjacency matrix $M$, and represents the overall influence that consumer $i$ has on her neighbors (and vice versa). The average degree $D$ is computed by taking the average of weighted degrees over all $n$ consumers. In this framework, the average degree in a network can be interpreted as a measure for the strength of network externality weights (i.e., the first order network effects). On the other hand, the heterogeneity in network externality weights (i.e., the second order network effects) can be captured by the variability in
weighted degrees across consumers. In particular, a balanced network corresponds to a homogenous network.

In what follows, we focus sequentially on the first- and second-order network effects by considering network structures that isolate these effects.

5.1. Strength of Network Externality Weights

To isolate the impact of first-order network effects, we consider a homogeneous network with a complete graph structure, in which all $n$ consumers exert identical influence on one another: $m_{ij} = \gamma$ for all $i \neq j$. This corresponds to a network in which $D_i = D = (n - 1)\gamma$ for all $i$. As such, the network is balanced and the second-order network effects are absent. For such a network, the intensity of interactions is captured by a single parameter $\gamma$, so that a larger $\gamma$ corresponds to greater network externality weights. In this case, the structure of the adjusted adjacency matrix $\hat{M}$ is simplified significantly, hence facilitating our analysis: for all eigenvectors except the first one, $\langle v_i, 1 \rangle = 0$; and the first eigenvalue is $\hat{\lambda}_1 = (n - 1)\gamma + \mu$. Hence, using the results from the last section, we can investigate the effect of $\gamma$ to establish a relationship between the strength of network externality weights and the firm’s optimal policy (and revenue). This relationship is summarized in the next proposition.

**Proposition 3.** Suppose the network is a complete graph with $m_{ij} = \gamma$ for all $i \neq j$.

(i) If price is exogenously given, firm’s optimal revenue and the duration of initial free-offering both increase with $\gamma$.

(ii) If price is endogenously determined, firm’s optimal revenue, price, and the duration of initial free-offering all increase with $\gamma$.

Proposition 3 indicates that the firm benefits from a more strongly connected network, as it establishes an unambiguous positive relationship between externality weight $\gamma$ and the optimal revenue. This is intuitive because, for any given pricing policy, a higher level of externality in the network translates into a more rapid growth in consumption vector, which in turn increases the firm’s revenue. Optimizing over the policy then enables the firm to exploit network effects more efficiently and further increase his revenue.

Less intuitive is the relationship between $\gamma$ and the optimal timing of price increase, which reflects the way externalities are exploited under the optimal policy. In fact, it is plausible to conjecture that the reverse relationship holds—namely, the duration of initial free-offering decreases in $\gamma$—because more rapid consumption growth driven by a larger $\gamma$ would incentivize the firm to monetize his service sooner. Proposition 3 suggests that this reasoning is incomplete.

To understand this result, first consider the exogenous price case (Part (i) of Proposition 3). It is important to recognize that it is not only the consumption that increases with $\gamma$ but also the
rate of consumption growth. Because the network externality creates a self-reinforcing mechanism among consumers, with the price being kept constant, consumption growth is accelerated as time passes by. This convex behavior is amplified by the strength of network externality weights, i.e., with higher \( \gamma \). Hence, a marginal increase in \( \gamma \) has a disproportionate impact on the consumers’ consumption growth in the future relative to the present. To take full advantage of this future gain, then, the firm finds it optimal to delay the timing of price increase, waiting until the growth rate in consumption reaches a critical point while tolerating a prolonged revenue loss during the initial free-offering periods.

Similar intuition extends to the case in which price is endogenously determined (Part (ii) of the proposition), except that the firm now has an additional lever to balance consumption growth against monetizing his service. In particular, as \( \gamma \) increases, the firm waits even longer to achieve a yet higher consumption level and then compensates for the forgone revenue by charging a higher price.

The proof of Proposition 3, particularly Part (ii), is fairly involved and includes multiple delicate steps. The main challenge in the proof stems from the fact that time is discrete in our framework. In the proof, we first consider a related stopping time problem in which time is continuous, and establish structural results for this surrogate problem. These auxiliary structural results then enable us to prove that, at the optimal price, the logarithm of the discounted revenue is supermodular in (continuous) time and \( \gamma \). This in turn allows us to prove the monotonicity results in the original discrete-time problem. The proof is presented in Appendix EC.3.1.

5.2. Heterogeneity of Network Externality Weights

We now turn our attention to the second-order network effects by varying the degree of heterogeneity in network externality weights while controlling for the first-order effect. To contrast the first- vs. the second-order network effects, we present the following example with two networks that differ in their second-order effects but are equivalent with respect to the first-order effects.

Example 2. Consider two networks with 5 consumers as depicted in Figure 2; these two networks have the same average degree of \( D = 0.36 \). On the other hand, their level of heterogeneity is different. In particular, the network on the right-hand side is complete and balanced, with its weighted degrees equal to \( D_i = 0.36 \) for all \( i \), while the network on the left-hand side is a star network with disparities in weighted degrees. The consumer at the center of the star network neighbors all other consumers and has weighted degree of 0.9, while the remaining consumers have weighted degrees of 0.225. □
First- vs. second-order network effects: The left graph has $m_{1i} = m_{i1} = 0.225$ for $i = 2, \ldots, 5$. The right graph has $m_{ij} = 0.09$ for all $i \neq j$.

It should be noted that balanced networks are not limited to complete graphs. In Example 2, for instance, many other balanced networks with 5 consumers can be constructed for which all consumers have the weighted degree of 0.36; the complete graph of Figure 2 (right) is only one such network. In Figure 3, we present three balanced networks with the same average degree that significantly differ in their structure. Despite the difference in structure, in the following theorem, we prove that balanced networks share two fundamental properties: (i) firm’s optimal decision and revenue are identical for all balanced networks, and (ii) these networks generate the minimum revenue.

![Figure 3](image)

Figure 3  Balanced networks with the same average degree but different structures. Left: All edges have weight 0.09. Center: The edges of the triangle have weight 0.18; the other edge has weight 0.36. Right: All edges have weight 0.18.

**Theorem 2.** Consider the class of all networks with $n$ consumers and average degree $D$.

(i) The firm’s optimal pricing policy and revenue is the same for all balanced networks in this class.

(ii) Among all networks in this class, balanced networks generate the lowest possible revenue for the firm.
Part (i) of Theorem 2 asserts that all balanced networks are identical from the firm’s perspective. This finding is particularly relevant in situations where the firm’s knowledge about the consumers’ interactions is limited. As long as consumers are known to be homogenous with respect to their overall interaction in the network, the firm does not need to acquire more information about the details of such interactions.

It is important to note that Part (i) of Theorem 2 also enables us to generalize the results of Proposition 3 to all balanced networks. That is, the proposition is not just limited to a complete network, and holds for any balanced network. More precisely, consider a balanced network with adjacency matrix $M$, and suppose we multiply all of its network externality weights, $m_{ij}$’s, by a factor $\kappa > 1$. We have: (i) If price is exogenous, firm’s optimal revenue and the duration of initial free-offering both increase with $\kappa$. (ii) If price is endogenous, firm’s optimal revenue, price, and the duration of initial free-offering all increase with $\kappa$.

Part (ii) of Theorem 2, on the other hand, highlights the downside of homogeneity in consumers’ overall interaction. The theorem states that, keeping the average degree constant, a balanced network provides the least opportunity for the firm to exploit network externalities. This points to the importance of second-order network effects.

5.2.1. Proof Sketch of Theorem 2 The proof of the first part of Theorem 2 relies on the observation that all balanced networks have the same largest eigenvalue $D$, and the same associated eigenvector $1/\sqrt{n}$. Further, even though other eigenvalues and eigenvectors vary across balanced networks, for all of them, the rest of eigenvectors are orthogonal to $1$, i.e., $\langle v_i, 1 \rangle = 0$, for $2 \leq i \leq n$. As evident from Lemma 1, firm’s discounted revenue depends on eigenvectors (and their associated eigenvalues) only through the inner products $\langle v_i, 1 \rangle$. The proof of the second part proceeds by first establishing a monotonically increasing relationship between firm’s discounted revenue and the number of (weighted) walks of any given length $k \in \mathbb{Z}_+$ on the underlying graph. It then uses a fundamental result from algebraic graph theory that relates the number of walks of length $k$ to the average degree of the graph (Blakley and Roy 1965, Täubig and Weihmann 2014). The detailed proof is presented in Appendix EC.3.2.

While the above theorem highlights the benefit of heterogeneity in weighted degrees across consumers, it does not provide any insights into how different “levels of heterogeneity” affects the firm’s pricing policy and revenue. Establishing such a result is very challenging as it requires comparing the policy and revenue of two arbitrary networks. Nevertheless, in the following proposition, we shed some light on the impact of heterogeneity level on the firm’s revenue in a fairly general setting.
Proposition 4. Let \( M \) and \( C \) be an arbitrary network and a complete network, respectively, with \( n \) consumers and average degree \( D \). Consider a new network \( M_w = wM + (1-w)C \) as a convex combination of \( M \) and \( C \). If \( \mu \geq -\lambda_n \), then for any given policy \( \pi = (\tau, p) \), the firm’s revenue from network \( M_w \) is increasing in \( 0 \leq w \leq 1 \). Consequently, the firm’s optimal revenue from network \( M_w \) also increases with \( 0 \leq w \leq 1 \).

According to Proposition 4, keeping the first-order network effects fixed (i.e., the average degree \( D \)), the firm is better off as the network moves farther away from a complete one (conditional on \( \mu \geq -\lambda_n \)). In other words, in a network that is a convex combination of an arbitrary network \( M \) and a complete network \( C \), the less weight given to the latter, the greater the firm’s revenue. Thus, keeping the average degree fixed, the firm favors more degree imbalance in the network, and his revenue decreases monotonically as the degree imbalance decreases. Effectively, in this proposition, we capture the “level of heterogeneity” by weight \( w \), while keeping the average degree fixed.

The proof of the above proposition is fairly involved (see the proof sketch presented at the end of this subsection), and requires the condition \( \mu \geq -\lambda_n \). This condition guarantees that all eigenvalues of the adjusted adjacency matrix are nonnegative. Even though our current proof technique requires us to make this additional assumption, our extensive numerical study suggests that the result holds even without it.

The benefit of heterogeneity arises from its impact on the “ripple effect,” namely, self-reinforcing feedback loops created over time due to indirect effects of network externalities beyond immediate neighbors. In a heterogeneous network, where the influences are uneven across agents, those with higher influence disproportionately benefit from such self-reinforcing loops, resulting in higher overall consumption. The underlying mechanism of this result is analogous to Jensen’s inequality for convex functions of random variables. We use the networks from Example 2 to illustrate this intuition.

Example 3. Consider the two networks presented in Example 2; let \( M \) be the star network on the left in Figure 2, and \( C \) the complete network on the right. Model parameters are set as follows: \( \delta = 0.95, \mu = 0.5, \alpha = \beta = 0.1 \). We focus on the consumption growth in the first 3 time periods assuming that the price has not increased yet during these time periods. Assign index 1 to the consumer at the center of the star network, and denote the consumption vectors of networks \( M \) and \( C \) in period \( t \) by \( x_t \) and \( y_t \), respectively. Then, we obtain the trajectory of consumption evolution as:

\[
\begin{align*}
x_1 &= \begin{bmatrix} 0.10 & 0.10 & 0.10 & 0.10 & 0.10 \end{bmatrix}^T \quad \text{with} \quad \langle x_1, 1 \rangle = 0.50, \\
x_2 &= \begin{bmatrix} 0.24 & 0.17 & 0.17 & 0.17 \end{bmatrix}^T \quad \text{with} \quad \langle x_2, 1 \rangle = 0.93,
\end{align*}
\]
\[
\begin{align*}
\mathbf{x}_3 &= [0.38 \ 0.24 \ 0.24 \ 0.24 \ 0.24]^T \quad \text{with} \quad \langle \mathbf{x}_3, 1 \rangle = 1.34, \\
\mathbf{y}_1 &= [0.10 \ 0.10 \ 0.10 \ 0.10 \ 0.10]^T \quad \text{with} \quad \langle \mathbf{y}_1, 1 \rangle = 0.50, \\
\mathbf{y}_2 &= [0.19 \ 0.19 \ 0.19 \ 0.19 \ 0.19]^T \quad \text{with} \quad \langle \mathbf{y}_2, 1 \rangle = 0.93, \\
\mathbf{y}_3 &= [0.26 \ 0.26 \ 0.26 \ 0.26 \ 0.26]^T \quad \text{with} \quad \langle \mathbf{y}_3, 1 \rangle = 1.30.
\end{align*}
\]

Consumption levels in the two networks are identical in the first period (\(\mathbf{x}_1 = \mathbf{y}_1\)). The effects of habit and externalities start to appear in the second period, creating a divergence between the two networks. In the star network, consumer 1 has a higher consumption level compared to others as she enjoys more positive externality from the consumption of her neighbors (see \(\mathbf{x}_2\)). In the complete network, on the other hand, the consumption levels grow uniformly for all consumers (see \(\mathbf{y}_2\)). Despite the divergence, the total consumption in period 2 is the same for both networks (i.e., \(\langle \mathbf{x}_2, 1 \rangle = \langle \mathbf{y}_2, 1 \rangle\)). This is due to the fact that, by the second period, the positive externality of consumption affects only up to each consumer’s immediate neighbors. Given that the average degrees of the two networks are identical, the total increase in consumption due to network externalities therefore remains the same.

In period 3, on the other hand, the consumption level of an individual consumer is not only impacted by those of her neighbors from period 2, but also by the consumption levels of the neighbors of her neighbors from period 1. With the average degree fixed, this secondary network externality (originating from the neighbors of neighbors) is stronger in the star network compared to the complete one. Such disparity in ripple effect originates from the neighbors beyond the immediate ones, and its impact is reflected in individual consumption levels at period 3 (see \(\mathbf{x}_3\) and \(\mathbf{y}_3\)) as well as in total consumption: \(\langle \mathbf{x}_3, 1 \rangle > \langle \mathbf{y}_3, 1 \rangle\). \(\square\)

Therefore, a greater degree imbalance in a network expedites consumption growth by leveraging on higher-order externalities, i.e., one’s indirect impact on the consumers in the network beyond her immediate neighbors. An important implication of this result is that having a small number of consumers whose influence reach deep into the network is preferred over having many consumers with equal level of influence distributed among them.

Studying how the optimal policy \(\pi^* = (\tau^*, p^*)\) changes in response to a change in degree imbalance is analytically intractable. Nevertheless, our extensive numerical study suggests that, similar to the case of Proposition 3, both \(\tau^*\) and \(p^*\) increase with \(w\) (where the weight \(w\) is defined in Proposition 4). The intuition for this observation is similar to that of Proposition 3. Specifically, more heterogeneity in weighted degrees positively impacts consumption as well as its growth rate, leading the firm to prolong his initial free-offering and then raising the price to a higher level. This is illustrated in our numerical study of Example 3. As Figure 4 shows, both \(\tau^*\) and \(p^*\) are non-decreasing in \(w\) under the optimal policy for \(\mathbf{M}_w = w\mathbf{M} + (1-w)\mathbf{C}\) with \(w\) ranging from 0 to 1.
5.2.2. Proof Sketch of Proposition 4: Similar to the proof of the second part of Theorem 2, we first establish a monotonically increasing relationship between firm’s discounted revenue and the number of (weighted) walks of any given length $k \in \mathbb{Z}_+$ on the underlying graph (with adjacency matrix $M_w$). For a given length $k$, we then show that “slightly” moving the adjacency matrix $M$—which corresponds to $M_w$ with $w = 1$—toward matrix $C$ results in a decrease in the number of walks of length $k$. The proof of the latter uses the following idea: the total number of walks of length $k$ can be viewed as the $k$-th moment of a discrete random variable $X \in \{\hat{\lambda}_1, \hat{\lambda}_2, \ldots, \hat{\lambda}_n\}$ with probability distribution $P(X = \hat{\lambda}_i) = \langle v_i, 1 \rangle^2/n$ for $i = 1, \ldots, n$. The detailed proof is presented in Appendix EC.3.3.

Summarizing the results in this section, we find that strength and heterogeneity of network externality weights play key roles in driving the optimal policy and its corresponding revenue for the firm. While the former exploits the direct influence of neighbors on one another to increase consumption and generate more revenue, the latter does so by leveraging on higher order (indirect) effects, e.g., neighbors of neighbors. Therefore, the firm prefers facing a network of consumers who are connected via strong externalities but differ in their overall influence in the network. In essence, the firm benefits from variability in a network by taking advantage of the indirect ripple effects it creates through network externalities. When such a condition is in place, the firm finds it optimal to prolong the duration of initial free-offering and recouping the corresponding revenue loss by raising the price to a higher level.

6. Value of Network Information

Our analyses thus far are based on the assumption that the firm has perfect knowledge of the network structure. While the assumption is reasonable for providers of internet-based services equipped to track user activities, for many firms it may be costly to mine such user-level data. This
raises the question: How valuable is network structure information in making pricing decisions? Put another way, will a firm’s revenue be affected significantly if it does not have access to such information?

Theorem 2 sheds some light on this question. In that theorem, we showed that for the class of balanced networks, the firm’s optimal pricing policy and the resulting revenue are functions of the average degree but are independent of the exact network structure. Therefore, as long as the network is known to be balanced, the firm does not need any information about the network other than its average degree.

On the other hand, network structure information may be valuable if the network is imbalanced. To quantify the value, suppose that the firm only knows the average degree $\overline{D}$, and considers the worst-case monetization opportunity by assuming that the network is complete (Theorem 2). Under this assumption, the firm follows the optimal pricing policy for the complete network, which we refer to as the heuristic policy. If the actual (unknown) network is imbalanced, this heuristic policy generates a suboptimal total discounted revenue, denoted as $J_h$, that is smaller than the optimal revenue when the network structure is known, denoted as $J$. Then the quantity $1 - J_h/J$ represents the relative revenue loss due to lack of detailed network information. If this revenue loss is outweighed by the cost of acquiring the information (e.g., developing a monitoring tool to track user activities), detailed network information is unnecessary and the firm can employ the heuristic policy. The following proposition provides a lower-bound on the ratio $J_h/J$.

**Proposition 5.** For $1 \leq i \leq n$, let $\Delta_i = \max_{1 \leq j \leq n} \{D_j\} - D_i \geq 0$, and let $\Delta = \sum_{i=1}^{n} \Delta_i/n$. We have:

$$\frac{J_h}{J} \geq \max \left\{ \frac{(1 - \delta \hat{\lambda}_1)(1 - \hat{\lambda}_1)}{(1 + \Delta/D - \delta \hat{\lambda}_1)(1 + \Delta/D - \hat{\lambda}_1)}, \frac{(1 - \delta)}{(1 + \Delta/D - \delta)(1 + \Delta/D)} \left( \frac{1}{1 + \Delta/D} \right)^{\tau+1} \right\},$$

where $\bar{\tau}$ is defined in Proposition 2.

The quantity $\Delta$ reflects the degree variation: if there is a central consumer whose degree is much higher than others, then $\Delta_i$ is large for some of the consumers and so is $\Delta$. The proposition identifies $\Delta/D$ as a key factor that determines the effectiveness of the heuristic policy. When $\Delta/D \to 0$, the lower-bound on $J_h/J$ approaches 1, and the revenue loss becomes negligible. Thus, the heuristic policy may perform well for networks whose degrees are “sufficiently balanced.” This can be viewed as a generalization of Theorem 2.

**7. Numerical Study**

In this section, we present a numerical study using a dataset that captures temporal user interactions on the website Stack Overflow, obtained from the Stanford Network Analysis Platform.
(SNAP) library (Paranjape et al. 2017). Stack Overflow is an online platform on which programmers interact with one another, exchanging information about technicalities of coding. What makes the Stack Overflow dataset unique and appealing to us is that, unlike many other datasets on social networks (that only record existence of interactions), it keeps track of the temporal information on the interactions between any pair of users over an extended period of time. We utilize this detailed recording and use the frequency of interactions between two users as a proxy for their externality weight. We describe the dataset in Section 7.1, followed by the analysis based on our previous results in Section 7.2.

7.1. Data

The dataset records all interactions between any pair of 2,601,977 users over 2,774 days. We define an “interaction” between user $i$ and user $j$ to be one of the following activities: (1) $i$ answers $j$’s question, or vice versa; (2) $i$ comments on $j$’s answer/question, or vice versa. We focus on the 250 most active users and interactions among them. For two users $i$ and $j$, we set $m_{ij} = m_{ji}$ to be the total number of interactions between them, scaled by a normalization factor $C$. Thus if users $i$ and $j$ had 100 interactions, $m_{ij}$ and $m_{ji}$ would be $100/C$. In Figure 5, we visualize the network of 50 randomly selected users among the 250 ones in the Stack Overflow dataset. Each node represents a user, and the size of each green node reflects her total number of interactions. That is, the size is proportional to the degree. Further, the edge width represents the frequency of interactions between the corresponding users, so that it is proportional to $m_{ij}$. It is evident from Figure 5 that both the number of interactions of a user and the intensity of interactions between a pair vary significantly across these 50 users; thus their network is highly imbalanced.

In Figure 6, we plot the degree distribution of all 250 users. The figure confirms the disparity in the degrees and shows that the degree distribution has a “long tail”; a few users produce a disproportionate volume of the interactions. Such a degree distribution is typical in many social networks. It further manifests the dissimilarity between the Stack Overflow network and a complete network.

Finally we study the spectral decomposition of the adjacency matrix $M$ associated with the network. As seen in Section 4, our characterizations of revenue and optimal policy are based on the eigenvalues and eigenvectors. In Figure 7, we plot the sorted eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ (after proper normalization), and their corresponding eigenvectors, i.e., $|\langle v_i, 1 \rangle|$, for $1 \leq i \leq n$. We observe that $\lambda_1 |\langle v_1, 1 \rangle|$ dominates its counterparts for $2 \leq i \leq n$. 
Figure 5  The visualization of the network between 50 randomly selected users among the top 250 most active ones in the Stack Overflow dataset.

Figure 6  The degree distribution of the top 250 most active users in the Stack Overflow dataset.
7.2. Pricing on Stack Overflow Network

Using the Stack Overflow dataset, we now study the impact of network effects on the pricing of a new service based on the analyses in the previous sections. To do this, we consider a hypothetical scenario where Stack Overflow decides to offer a new revenue-generating service to the users already on its platform. As our model prescribes, we assume that the platform initially offers the service free of charge to build up consumption, raising the price exactly once at a later point in time. The platform determines the optimal timing and magnitude of price increase to maximize its total discounted revenue. For this study, we assume the following model parameters: $\alpha = \beta = 0.1$, $\mu = 0.1$, $\delta = 0.99$, and $C = 2030$. Recall that $C$ is the normalization factor in the adjacency matrix $M$, and determines by how much $m_{ij}$ increases if consumers $i$ and $j$ have one extra interaction.

We first examine how the firm’s policy and revenue is impacted by the strength and heterogeneity of the externality weights, as discussed in Sections 5.1 and 5.2, respectively. To highlight these impacts, we construct four different networks with different levels of strength and imbalance, based on the Stack Overflow data. They are presented in the decreasing order of imbalance (or an increasing level of homogeneity):

**Network 1:** Original Stack Overflow network with distinct $m_{ij}$’s and average degree set to $D$.

**Network 2:** Unweighted version of Stack Overflow network with average degree $D$, where $m_{ij} = nD/2E$ for all $i, j$ provided that there is an edge between $i$ and $j$ in the original network. Here, $E$ is the total number of edges.

**Network 3:** A complete network with average degree $D$.

**Network 4:** No network.

Note that the average degree is preserved in Networks 1, 2, and 3, to ensure the same level of network strength. However, in Network 4 the average degree is reduced to zero. Table 1 reports...
the optimal policy \((\tau^* \text{ and } p^*)\) and revenue for these networks. From the table, we see that the optimal revenue in the presence of network effects captured in the original Stack Overflow network (Network 1) is 766%, 338%, and 290% greater than the optimal revenue under no network effects (Network 4), complete network (Network 3), and unweighted network (Network 2), respectively. These findings confirm the analytical results derived in the previous sections, and emphasizes the importance of consumer heterogeneity in monetizing local network effects.

Next, we perform comparative statics on the Stack Overflow network (Network 1) with respect to the strength of network externality weights and the habit persistence parameter \(\mu\), and report the results in Table 2. We do this by varying the normalization factor \(C\) — smaller \(C\) corresponds to larger network externality weights—as well as the habit coefficient \(\mu\). The value of other model parameters remain the same as before, with \(\alpha = \beta = 0.1\), and \(\delta = 0.99\). We confirm our earlier finding from Proposition 3 that, as one increases the network externality weights or the habit persistence \(\mu\), the optimal stopping time, price, and revenue all increase. Although Proposition 3 was derived under the assumption of a complete network, our numerical result on the Stack Overflow data (as well as additional numerical studies omitted from the paper) suggests that this finding is more general and carries over to other network structures as well.

Finally, we numerically examine the value of having information about the network structure, which was discussed in Section 6. We compute the optimal policy for four scenarios that incorporate network effects in varying degrees of detail. Particularly, we compare the following scenarios:

**Scenario 1:** The firm knows the detailed network information, i.e., the adjacency matrix \(M\).
Scenario 2: The firm has information about the existence of interactions, but not the intensity. In this case, the firm knows the average degree, and assumes $m_{ij} = n_D E$, for every pair $i$ and $j$ that interact. As mentioned before, most publicly available datasets on networks are indeed unweighted.

Scenario 3: The firm only knows the average degree $D$; this is the same setting as in Section 6.

Scenario 4: The firm ignores the network effects; equivalently assumes $M = 0$.

Table 3 reports the optimal policy $(\tau^*, p^*)$ for the above scenarios, and the revenue it achieves when implemented on the actual network. First, we observe that both elements of the optimal policy significantly differ if full network information is known. If the firm has the perfect knowledge of the underlying Stack Overflow network, it will delay the price increase and charge a higher price. Also, somewhat surprisingly, the policies derived under Scenario 2 (knowing the unweighted network structure) and Scenario 3 (knowing only the average degree) are identical. This highlights the importance of incorporating the intensity of interactions. In Scenario 2, the firm makes decisions based on significantly more detailed information compared to simply assuming a complete network. However, it does not perform any better.

Finally, note that policies corresponding to Scenarios 2-4 attain a suboptimal revenue. In particular, the revenue loss for these policies are 3.8%, 3.8%, and 4.5%, respectively. We remark that a few percent improvement in revenue can translate into a significant gain, especially for low-margin businesses. The fact that the revenue attained under Scenarios 2 and 3 is close to that of the na"ive policy of Scenario 4 (i.e., the no network information setting) once again highlights the importance of incorporating the detailed network information.

8. Conclusion

Firms offering interaction-based services—such as peer-to-peer monetary transfer, file-sharing, and online games—rely their businesses on positive network externality arising from consumers’ personal and professional connections. The significance of network externality is especially heightened for newly-introduced services, since a firm’s ability to generate and sustain revenue in the long-term depends crucially on how the initial growth of network-driven service consumption is managed. Many firms employ a pricing strategy under which they offer their services for free in the beginning in order to maximize initial consumption growth, delaying monetization of services to a later
time by raising the price only after a sufficient level of consumption is built in the network. In the presence of network externality, therefore, firms face an optimal stopping time problem for price increase. To analyze this problem, we develop a model based on the constructs of graph theory that incorporates non-stationary consumption dynamics. Since services are non-durable and used on an ongoing basis, consumers make repeated decisions on their consumption in reaction to current price and consumption of their network peers. The dynamics of these decisions in a network, coupled with a firm’s optimal pricing, present new analytical challenges that have not been addressed in the literature. We overcome these challenges by making a novel connection between the underlying high-dimensional dynamic program and several analytical techniques from algebraic graph theory.

In our analysis we isolate the impacts of first- and second-order network effects, corresponding respectively to the strength and the heterogeneity of network connections among consumers. With respect to the first-order network effect, we show that a firm facing a more strongly connected network extends the duration of initial free offering of service, and when it raises the price, it sets the price to a higher amount. Such a delay of price increase occurs despite the fact that the firm can monetize its service sooner because of rapid consumption growth fueled by a high level of consumer interactions in a strongly connected network. This seemingly contradictory conclusion arises from a convex behavior of consumption trajectory: stronger network effects result in not only a faster consumption growth overall, but also an accelerated rate of growth further into the future. To take advantage of the higher future gain, then, the firm favors delaying monetization of service.

Our analysis also reveals nontrivial impacts of second-order network effects. In particular, we show that, under mild technical conditions, introducing more heterogeneity to network connections results in a higher discounted firm revenue. In other words, the firm benefits from a network imbalance: with the average network strength fixed, having a diverse mix of consumers with strong connections and those without is preferred to having a relatively homogeneous consumer population with similar degrees of connections. This result arises from the fact that a network imbalance magnifies the “ripple effect,” specifically, self-reinforcing loops caused by indirect network effects among consumers who are not immediate neighbors in the network but are indirectly connected through their mutual neighbors.

In addition, we show that a firm’s pricing decision is insensitive to the underlying network structure as long as the network is balanced, i.e., all consumers in the network have the same level of connections. Moreover, the firm generates the least possible revenue when it faces such a network. These findings suggest that a costly investment into acquiring information about the details of network structure is undesirable if there is a reason to believe that network connections are relatively homogeneous. Finally, we validate our analytical findings with an extensive numerical
study using the user network data obtained from Stack Overflow, a knowledge-sharing website for programmers.

The primary focus in this paper is to study the impact of network externality on optimal pricing of a newly-introduced service in a monopolistic setting. By employing a novel analytical approach that incorporates the elements of spectral graph theory, we make both methodological and managerial contributions to the literature on pricing under network effects. The analytical framework developed in this paper paves ground for further studies on the impacts of the complexities that are not captured in the current model, such as forward-looking consumer behavior and competition among service providers. Furthermore, the insights derived from our analysis provide actionable guidelines to providers of network-driven services.

Acknowledgments
We would like to acknowledge Scott Rodilitz and Runyu Tang for their help with data analysis, Alireza Tahbaz-Salehi, and seminar participants at INFORMS, MSOM, the Marketplace Innovation Workshop, and the Denver Research Meeting on Networks, Matching, and Platforms for their valuable comments.

References


**Appendix**
EC.1. Preliminaries on Algebraic Graph Theory

Recall that we represent the network underlying consumers interaction by matrix \( M \) which is assumed to be symmetric and nonnegative. This matrix representation enables us to utilize techniques from algebraic graph theory—which is a part of graph theory that applies algebraic method to graph problems. In particular, in our analysis, we rely on the spectral decomposition of symmetric matrices as well as a few celebrated results for nonnegative symmetric matrices. Before proceeding to present our proofs, we review these results below.

Let \( A \in \mathbb{R}^{n \times n} \) be a symmetric \( n \times n \) matrix. We denote its eigenvalues and eigenvectors by \( \lambda_i(A) \in \mathbb{R} \) and \( v_i(A) \in \mathbb{R}^n \), \( 1 \leq i \leq n \), respectively. Eigenvalues are indexed such that \( \lambda_1(A) \geq \lambda_2(A) \geq \cdots \geq \lambda_n(A) \).

**Fact 1** The spectral decomposition of matrix \( A \) is given by \( A = \sum_{i=1}^{n} \lambda_i(A) v_i(A) v_i^T(A) \). Further, \( \langle A1, 1 \rangle = \sum_{i=1}^{n} \lambda_i(A) \langle v_i(A), 1 \rangle^2 \).

**Fact 2** For any \( k = 0, 1, \ldots \), \( \lambda_i(A^k) = \lambda_i(A)^k \), and \( v_i(A^k) = v_i(A) \).

**Fact 3** Let \( I_n \) be the identity matrix with dimension \( n \times n \). For any coefficients \( a, b \in \mathbb{R} \), we have \( \lambda_i(aI_n + bA) = a + b\lambda_i(A) \), and \( v_i(aI_n + bA) = v_i(A) \).

**Fact 4** Let \( f(A) \) be any polynomial function of matrix \( A \). Then we have \( f(A) = \sum_{i=1}^{n} f(\lambda_i(A)) v_i(A) v_i^T(A) \).

**Fact 5** The geometric sum of matrix \( A \), \( \sum_{k=0}^{\infty} A^k \), converges to \( (I_n - A)^{-1} \) iff \( \max_{1 \leq i \leq n} |\lambda_i(A)| < 1 \).

**Fact 6** Suppose that the row sum of all rows in matrix \( A \) are equal, i.e., \( \sum_{j=1}^{n} A_{ij} = \bar{A} \), for \( 1 \leq i \leq n \). Then \( \lambda_1(A) = \bar{A} \), with \( v_1(A) = (1, 1, \ldots, 1)/\sqrt{n} \). Further, \( \langle v_1, v_i \rangle = 0 \), for \( 2 \leq i \leq n \).

**Fact 7** (Perron-Frobenius Theorem) For a nonnegative and symmetric matrix \( A \), we have \( \lambda_1(A) \geq |\lambda_i(A)| \) for \( 1 \leq i \leq n \). Further \( v_1(A) \) is a nonnegative vector.

**Fact 8** (Theorem 1 in Blakley and Roy (1965)) For a nonnegative and symmetric matrix \( A \) and any nonnegative integer \( k \), we have \( \langle A^k 1, 1 \rangle \geq n^{1-k} (A1, 1)^k \).
EC.2. Proof of the Results in Section 4

EC.2.1. Proof of Lemma 1:

If the initial consumption is given by \( x_0 \) and the price is set to a constant \( p \), then by Equation (4), the consumption after \( t \) periods is given by \( x_t = (\alpha - \beta p)(I_n - \tilde{M})^{-1}(I_n - \tilde{M}')1 + \tilde{M}'x_0 \). Therefore,

\[
\Gamma(x; p) = \sum_{t=1}^{\infty} \delta^t p \langle x_t | x_0 = x, 1 \rangle \\
= \sum_{t=1}^{\infty} \delta^t p \langle (\alpha - \beta p)(I_n - \tilde{M})^{-1}(I_n - \tilde{M}')1, 1 \rangle + \sum_{t=1}^{\infty} \delta^t p (\tilde{M}'x, 1) \\
= \left\langle p(\alpha - \beta p)(I_n - \tilde{M})^{-1} \sum_{t=1}^{\infty} (\delta^t I_n - \delta^t \tilde{M}')1, 1 \right \rangle + \left\langle p \sum_{t=1}^{\infty} \delta^t \tilde{M}'x, 1 \right \rangle \\
\]

First we remark that by Assumption 1, \(|\lambda_i(\tilde{M})| = |\mu + \lambda_i| < 1\), \(1 \leq i \leq n\), where we use Fact 3 in computing \( \lambda_i(\tilde{M}) \). Therefore by Fact 5, \( \sum_{k=0}^{\infty} \tilde{M}^k \) converges to \((I_n - \tilde{M})^{-1}\). Using Facts 1 and 4, we express \( \Gamma(x; p) \) as follows:

\[
\Gamma(x; p) = p(\alpha - \beta p) \left\langle \sum_{i=1}^{n} \sum_{t=1}^{\infty} \frac{\delta^t - \delta^t \hat{\lambda}_i}{1 - \hat{\lambda}_i} \langle v_i, v_i \rangle, 1 \right \rangle + p \left\langle \sum_{i=1}^{n} \frac{\delta \hat{\lambda}_i}{1 - \delta \hat{\lambda}_i} \langle v_i, v_i \rangle, 1 \right \rangle \\
= p(\alpha - \beta p) \sum_{i=1}^{n} \frac{\delta^t - \delta^t \hat{\lambda}_i}{1 - \hat{\lambda}_i} \langle v_i, 1 \rangle^2 + p \sum_{i=1}^{n} \frac{\delta \hat{\lambda}_i}{1 - \delta \hat{\lambda}_i} \langle v_i, 1 \rangle \langle v_i, x \rangle \\
= p(\alpha - \beta p) \sum_{i=1}^{n} \left( \frac{\delta}{1 - \delta} - \frac{\delta \hat{\lambda}_i}{1 - \delta \hat{\lambda}_i} \right) \frac{1}{1 - \hat{\lambda}_i} \langle v_i, 1 \rangle^2 + \sum_{i=1}^{n} \frac{p \delta \hat{\lambda}_i \langle v_i, 1 \rangle \langle v_i, x \rangle}{1 - \delta \hat{\lambda}_i} \\
= \sum_{i=1}^{n} \frac{p(\alpha - \beta p) \delta \langle v_i, 1 \rangle^2}{(1 - \delta)(1 - \delta \hat{\lambda}_i)} + \sum_{i=1}^{n} \frac{p \delta \hat{\lambda}_i \langle v_i, 1 \rangle \langle v_i, x \rangle}{1 - \delta \hat{\lambda}_i}. \\
\]

Hence we have completed the proof. \( \square \)

EC.2.2. Proof of Theorem 1:

Proof of Part (i): we have already presented the main steps of the proof in Section 4.1.1. Here we first prove Claim 1:

Proof of Claim 1: To show the claim, we derive an alternative expression for \( \Gamma(x, p) \). By the definition, if the firm stops the free-offering, then the revenue in the current period (earned at the end and thus discounted) is \( \delta p((\alpha - \beta p)1 + \tilde{M}x, 1) \). In the next period, by the dynamics of \( x_t \), the consumption is \((\alpha - \beta p)1 + \tilde{M}x; p\); thus, the revenue from the next period onward is \( \delta \Gamma((\alpha - \beta p)1 + \tilde{M}x; p) \). Therefore,

\[
\Gamma(x; p) = \delta p((\alpha - \beta p)1 + \tilde{M}x, 1) + \delta \Gamma((\alpha - \beta p)1 + \tilde{M}x; p). \\
\]
Using the above and Lemma 1, we have:

$$\Gamma(x_i; p) \geq \delta \Gamma(\alpha 1 + \hat{M}x_i; p) \iff $$

$$\delta p \langle (\alpha - \beta p)1 + \hat{M}x_i, 1 \rangle \geq \delta \left[ \Gamma(\alpha 1 + \hat{M}x_i; p) - \Gamma((\alpha - \beta p)1 + \hat{M}x_i; p) \right] \iff$$

$$\delta p(\hat{M}x_i, 1) \geq \frac{n}{\sum_{i=1}^{\delta \hat{\lambda}_i p^2 (v_i, 1)^2} - \delta np(\alpha - \beta p) \iff}$$

$$\langle \hat{M}x_i, 1 \rangle \geq \sum_{i=1}^{\delta \hat{\lambda}_i p^2 (v_i, 1)^2} - n(\alpha - \beta p). \quad \text{(EC.1)}$$

By (4) and the fact that $x_0 = 0$, we have $\langle \hat{M}(x_{i+1} - x_i), 1 \rangle = \langle \alpha \hat{M}^{t+1}1, 1 \rangle \geq 0$ before the free-offering ends. Hence $\langle \hat{M}x_i, 1 \rangle$ is increasing in $t$.

Now we are ready to prove the claim: If $\Gamma(x_i; p) \geq \delta \Gamma(\alpha 1 + \hat{M}x_i; p)$, then (EC.1) holds and by the monotonicity of $\langle \hat{M}x_i, 1 \rangle$, we have $\Gamma(x_{i+1}; p) \geq \delta \Gamma(x_{i+1}; p)$ for all $s > t$ if the free-offering lasts until $s$. Therefore, $J(x_{i+1}; p) = \max_{s \geq t+1} \delta^{s-t} \Gamma(x_i; p) = \Gamma(x_{i+1}; p)$. This completes the proof of Claim 1. \(\Box\)

We remark that $\tau_1$ is the smallest $t$ that satisfies the condition defined in (EC.1). This is the same condition as in part (i) of Theorem 1. Thus this completes the proof of part (i).

Proof of Part (ii): The result follows from the observation that the right-hand side in (EC.1) is increasing in $p$. \(\Box\)

Proof of Corollary 1: The proof of Corollary 1 is straightforward and thus omitted. \(\Box\)

**EC.2.3. Auxiliary Lemma on Firm’s Discounted Revenue**

To show the remaining proofs, we first provide the following auxiliary lemma: Denote the total discounted revenue when the firm follows a given policy $(\tau, p)$ by $J(0; \tau, p)$, starting from $x_0 = 0$. We have:

**Lemma EC.1.** The revenue for a given policy $(\tau, p)$ is

$$J(0; \tau, p) = \sum_{i=1}^{n} \frac{p(\alpha - \beta p) \delta^{t+1}(v_i, 1)^2}{(1 - \delta)(1 - \delta \hat{\lambda}_i)} + \alpha \sum_{i=1}^{n} \frac{\delta^{t+1} \hat{\lambda}_i p(1 - \hat{\lambda}_i)(v_i, 1)^2}{(1 - \delta \lambda_i)(1 - \lambda_i)}. \quad \text{(EC.2)}$$

**Proof of Lemma EC.1:** Using Lemma 1, definition of $x_0$ given by (4), and Fact 4, we have:

$$J(0; \tau, p) = \delta^{t} \Gamma(x_{i}; p)$$

$$= \sum_{i=1}^{n} \frac{p(\alpha - \beta p) \delta^{t+1}(v_i, 1)^2}{(1 - \delta)(1 - \delta \hat{\lambda}_i)} + \sum_{i=1}^{n} \frac{\delta^{t+1} \hat{\lambda}_i p(v_i, 1)(v_i, x_i)}{1 - \delta \lambda_i}.$$
This completes the proof. □

EC.2.4. Proof of Proposition 1:

By Lemma EC.1, $J(0; \tau, p)$ is a quadratic function of $p$. Therefore, for given $\tau$, the optimal price $p^*(\tau)$ can be expressed in a closed form: $p^*(\tau) = \frac{\alpha}{2\beta} \left(1 + \frac{B(\tau)}{A}\right)$, where

\[
A = \sum_{i=1}^{n} \frac{\langle v_i, 1 \rangle^2}{(1-\delta)(1-\delta\lambda_i)} \quad \text{and} \quad B(\tau) = \sum_{i=1}^{n} \frac{\hat{\lambda}_i(1-\hat{\lambda}_i\tau)\langle v_i, 1 \rangle^2}{(1-\delta\lambda_i)(1-\lambda_i)}.
\]

By Assumption 1, $\hat{\lambda}_i < 1$ for $1 \leq i \leq n$; further $\delta < 1$, and $\langle v_i, 1 \rangle^2 > 0$ for at least one $i$, thus $A > 0$. To show $B(\tau) \geq 0$, we note that using Fact 4, we can express it as

\[
B(\tau) = \sum_{i=1}^{\tau} \langle \hat{M}^i(I_n - \delta\hat{M})^{-1}1, 1 \rangle = \sum_{i=1}^{\tau} \sum_{k=0}^{\infty} \langle \delta^k \hat{M}^t \hat{M}^{t+k}, 1 \rangle, \quad \forall \tau
\]

Note that each term in the summation is an inner product of two nonnegative vectors; therefore, $B(\tau) \geq 0$, with $B(\tau) = 0$ only if $\hat{M} = 0$.

Next we proceed to show $p^*(\tau) \leq \alpha/\beta$. First note that:

\[
p^*(\tau) \leq \alpha/\beta \iff B(\tau) \leq A \tag{EC.3}
\]

A sufficient condition for $B(\tau) \leq A$ is the following:

\[
\frac{\hat{\lambda}_i(1-\hat{\lambda}_i\tau)}{1-\lambda_i} \leq \frac{1}{1-\delta} \quad \forall i \tag{EC.4}
\]

We show that under Assumption 1, the above condition holds. Note that if $\hat{\lambda}_i \leq 0$, then (EC.4) is satisfied. Thus we focus on positive eigenvalues. We have

\[
\frac{\hat{\lambda}_i(1-\hat{\lambda}_i\tau)}{1-\lambda_i} \leq \frac{\hat{\lambda}_i}{1-\lambda_i} \leq \frac{\hat{\lambda}_i}{1-\delta} \leq \frac{1}{1-\delta}, \quad \forall \lambda_i > 0
\]

where in the second inequality we use the fact that function $\frac{\hat{\lambda}_i}{1-\lambda_i}$ is increasing for $x \in [0,1)$, and $\lambda_i \leq \lambda_1$. The last inequality holds because by Assumption 1, we have

\[
\delta \geq 2 - \frac{1}{\lambda_1} \iff \frac{\hat{\lambda}_i}{1-\lambda_i} \leq \frac{1}{1-\delta}.
\]

This completes the proof. □
EC.2.5. Proof of Proposition 2:

We first show that \( J(0; 0, \frac{\alpha}{2\beta}) \geq J(0; \tau, p^*(\tau)) \) for \( \tau > [\log(4)/\log(\delta)] \). From Lemma EC.1 and using the notations in Proposition 1, we have

\[
J(0; 0, \frac{\alpha}{2\beta}) = \frac{\alpha^2}{4\beta} \sum_{i=1}^{n} \delta \langle v_i, 1 \rangle^2 (1 - \delta)(1 - \delta \lambda_i) = \frac{\alpha^2 \delta}{4\beta} A.
\]

For \( \tau > [\log(4)/\log(\delta)] \), we plug \( p^*(\tau) = \alpha(1 + B(\tau)/A)/2\beta \) into Lemma EC.1,

\[
J(0; \tau, p^*(\tau)) = \delta^\tau + \frac{\alpha^2}{2\beta A} (A + B(\tau))^2.
\]

By (EC.3), \( B(\tau)/A \leq 1 \), so

\[
J(0; \tau, p^*(\tau)) \leq \frac{\alpha^2 \delta^{\tau+1}}{\beta} A.
\]

Therefore, when \( \tau > [\log(4)/\log(\delta)] \), we have \( J(0; \tau, p^*(\tau)) < J(0; 0, \frac{\alpha}{2\beta}) \). This implies that for \( \tau > [\log(4)/\log(\delta)] \), \( (\tau, p^*(\tau)) \) cannot be the optimal policy because it generates less revenue than policy \( (0, \alpha/2\beta) \). Therefore, \( \tau^* \leq [\log(4)/\log(\delta)] \) and we have proved the result. \( \square \)

EC.3. Proof of the Results in Section 5

EC.3.1. Proof of Proposition 3:

Using Fact 6, for a complete graph, we have: \( \hat{\lambda}_i = (n - 1)\gamma + \mu, v_i = (1/\sqrt{n}, 1/\sqrt{n}, ..., 1/\sqrt{n}) \), and \( \langle v_i, 1 \rangle = 0 \) for \( i \geq 2 \). As a result, by Lemma EC.1,

\[
J(0; \tau, p) = \frac{p(\alpha - \beta p)\delta^{\tau+1}n}{(1 - \delta)(1 - \delta \lambda_1)} + \alpha \frac{\delta^{\tau+1} \hat{\lambda}_i p(1 - \hat{\lambda}_i^\tau)n}{(1 - \delta \hat{\lambda}_i)(1 - \lambda_i^\tau)}.
\]

According to Proposition 1, we also have that for a complete graph

\[
p^*(\tau) = \frac{\alpha}{2\beta} \left( 1 + \frac{\hat{\lambda}_i(1 - \delta)(1 - \hat{\lambda}_i^\tau)}{1 - \lambda_i} \right).
\]

In addition, for a given price \( p \), we consider a relaxed version of the stopping time problem in which time is continuous. More precisely, we define:

\[
\bar{\tau}^*(p) = \min \left\{ \tau \in \mathbb{R}_+ \mid (\hat{\mathbf{M}}_x, 1) \geq \sum_{i=1}^{n} \delta \beta p \hat{\lambda}_i \langle v_i, 1 \rangle^2 / (1 - \delta \hat{\lambda}_i) - n(\alpha - \beta p) \right\}.
\]

In the following claim, we provide a closed form solution for \( \bar{\tau}^*(p) \), and relate it to \( \tau^*(p) \):

CLAIM EC.1. (i) We have:

\[
\bar{\tau}^*(p) = \begin{cases} 
0 & \text{if } f(\hat{\lambda}_1) > 1 \\
\log f(\hat{\lambda}_1) / \log(\hat{\lambda}_1) & \text{if } f(\hat{\lambda}_1) \in (0, 1]
\end{cases}
\]

where \( f(\cdot) : (0, 1) \to \mathbb{R}_+ \) is given by: \( f(\hat{\lambda}_1) = \frac{\alpha \delta \hat{\lambda}_1 - \alpha - \beta \hat{\lambda}_1 p + \beta p}{\alpha \hat{\lambda}_1 (\delta \hat{\lambda}_1 - 1)} \).
(ii) $\tau^*(p) = \lceil \bar{\tau}^*(p) \rceil$.

**Proof of Claim EC.1:**

Part (i): We start by simplifying the two sides of the condition in (EC.7) for the complete graph:

$$\langle \hat{M}x, 1 \rangle = \alpha \langle \hat{M}(I_n - \hat{M})^{-1}(I_n - \hat{M}')1, 1 \rangle = an\hat{\lambda}_1 \frac{1 - \hat{\lambda}_1'}{1 - \hat{\lambda}_1}. \quad (EC.8)$$

and

$$\sum_{i=1}^n \frac{\delta \beta p\hat{\lambda}_i(v_i, 1)^2}{1 - \delta \lambda_i} - n(\alpha - \beta p) = \frac{\delta n \beta p\hat{\lambda}_1}{1 - \delta \lambda_1} - n(\alpha - \beta p). \quad (EC.9)$$

Equating (EC.8) and (EC.9), we get:

$$an\hat{\lambda}_1 \frac{1 - \hat{\lambda}_1'}{1 - \hat{\lambda}_1} = \frac{\delta n \beta p\hat{\lambda}_1}{1 - \delta \lambda_1} - n(\alpha - \beta p) \iff \hat{\lambda}_1' = f(\hat{\lambda}_1) \equiv 1 - \frac{1 - \hat{\lambda}_1'}{\alpha \hat{\lambda}_1} \left( \frac{\delta \beta p\hat{\lambda}_1}{1 - \delta \lambda_1} - (\alpha - \beta p) \right) \iff (EC.10)$$

$$\tau^*(p) = \log f(\hat{\lambda}_1)/\log(\hat{\lambda}_1).$$

Simplifying $f(\hat{\lambda}_1)$ yields

$$f(\hat{\lambda}_1) = \frac{\alpha \delta \hat{\lambda}_1 - \alpha - \beta \hat{\lambda}_1 p + \beta p}{\alpha \hat{\lambda}_1 (\delta \hat{\lambda}_1 - 1)}. \quad (EC.11)$$

Next we determine the range of $f(\cdot)$ for $\hat{\lambda}_1 \in (0,1)$. Clearly, the denominator is negative for $\hat{\lambda}_1 \in (0,1)$. By Proposition 1, we know that the relevant price range is $p \in [0, \alpha/\beta]$, for which, the numerator satisfies:

$$\alpha \delta \hat{\lambda}_1 - \alpha - \beta \hat{\lambda}_1 p + \beta p \leq \alpha \hat{\lambda}_1 - \alpha - \beta \hat{\lambda}_1 p + \beta p = -(\alpha - \beta p)(1 - \hat{\lambda}_1) \leq 0.$$ 

Therefore, the range of $f(\cdot)$ is $(0, +\infty)$, implying that:

$$\bar{\tau}^*(p) = \begin{cases} 0 & \text{if } f(\hat{\lambda}_1) > 1 \\ \log f(\hat{\lambda}_1)/\log(\hat{\lambda}_1) & \text{if } f(\hat{\lambda}_1) \in (0,1] \end{cases}$$

Part (ii): Because $\langle \hat{M}x, 1 \rangle$ is increasing in $t$, the discrete $\tau^*(p)$ is the ceiling of $\bar{\tau}^*(p)$. \qed

Next we use Claim EC.1 to establish a few structural results in the following claim.

**Claim EC.2.**

1. For given $\tau$ and $p$, the firm’s total discounted revenue $J(0; \tau, p)$ as defined in Lemma EC.1 is increasing in $\gamma$.
2. For given $\tau$, the optimal price $p^*(\tau)$ is strictly increasing in $\gamma$. 

---

\[e-companion to Alizamir et al.: Impact of Network Structure on New Service Pricing ee7\]
(3) For given $p$, the optimal stopping time $\tau^*(p)$ is increasing in $\gamma$, and so is $\tau^*(p)$.

(4) For $\tau \in \mathbb{R}_+$, the function $J(0; \tau, p^*(\tau))$ is either unimodal or strictly decreasing.

**Proof of Claim EC.2:**

**Part (1):** Note that in (EC.5), $\lambda_1 = (n - 1)\gamma + \mu$ is increasing in $\gamma$ and $J(0; \tau, p)$ is increasing in $\hat{\lambda}_1$. Thus, $J(0; \tau, p)$ is increasing in $\gamma$.

**Part (2):** It is clear from (EC.6) that $p^*(\tau)$ is strictly increasing in $\hat{\lambda}_1$, and thus strictly increasing in $\gamma$.

**Part (3):** By the definition $\tau^*(p)$ and $f(\cdot)$, in order to show $\tau^*(p)$ is increasing in $\gamma$, it suffices to show the following two results:

(a) the domain $\{x| x \in (0,1), f(x) \in (0,1]\}$ is an interval;

(b) when $\hat{\lambda}_1 \in (0, 1)$ and $f(\hat{\lambda}_1) \in (0, 1]$, $\log f(\hat{\lambda}_1)/\log(\hat{\lambda}_1)$ is increasing in $\hat{\lambda}_1$.

To show (a), recall that the range of $f(\cdot)$ is $(0, +\infty)$ for $x \in (0, 1)$. Moreover,

$$f(x) \leq 1 \iff \alpha \delta x^2 - (\alpha + \alpha \delta + \beta p)x + \alpha - \beta p \leq 0.$$  

Since we can show that $f(1) = 1$ and $\lim_{x \to 0} f(x) = +\infty$, by the property of quadratic functions, $\{x| x \in (0,1), f(x) \in (0,1]\}$ is an interval of the form $[x', 1)$ for some $x' \in (0, 1)$.

To show (b), we take derivative of $\log f(x)/\log(x)$ with respect to $x$:

$$\frac{\partial d(\log f(x)/\log(x))}{dx} = \frac{f'(x)\log(x)/f(x) - \log(f(x))/x}{\log(x)^2}. \quad (EC.12)$$

Our goal is to show the above derivative is non-negative for $f(x) \in (0,1]$, or equivalently, $x \in [x', 1)$. When $f'(x) \leq 0$, (EC.12) is already non-negative. When $f'(x) > 0$, the non-negativity is equivalent to $x \log(x) \geq f(x) \log(f(x))/f'(x)$. Since $1 \log 1 = 0 = f(1) \log(f(1))$, it suffices to show that for $x \in [x', 1)$

$$\left(\frac{f(x) \log f(x)}{f'(x)}\right)' \geq (x \log x)'.$$  

(Ec.13)

Note that

$$\left(\frac{f(x) \log f(x)}{f'(x)}\right)' - (x \log x)' = \log f(x) - \frac{f''(x)f(x) \log f(x)}{(f'(x))^2} - \log x. \quad (EC.14)$$

If we can show that $f''(x)f(x) \geq (f'(x))^2$, then because $f(x) \leq 1$ and $x \leq 1$, the right-hand side of (EC.14) is greater than or equal to $\log(f(x)) - (f'(x))^2 \log(f(x))/(f'(x))^2 - \log(x) \geq 0$. Therefore, it suffices to show

$$\frac{f''(x)f(x) - (f'(x))^2}{(f(x))^2} \geq 0 \iff \frac{f'(x)}{f(x)} \text{ is increasing in } x.$$
By the expression (EC.11), we have

\[
\frac{f'(x)}{f(x)} = \frac{-(\alpha - \beta p)(1-\delta x)}{ax^2} + \frac{\beta p(1-\delta)}{\alpha x}.
\]

The denominator is decreasing in \( x \), the numerator is increasing in \( x \), and the ratio is positive. Therefore, \( f'(x)/f(x) \) is increasing in \( x \). This in turn completes our proof of (b). So far we have proved that \( \bar{\tau}^*(p) \) increasing in \( \gamma \); by part (ii) of Claim EC.1, \( \tau^*(p) \) is an increasing function of \( \bar{\tau}^*(p) \), and therefore itself is increasing in \( \gamma \).

**Part (4):** To show part (4), note that using definition of Proposition 1, we have:

\[
J(0; \tau, p^*(\tau)) = \delta^{r+1} \frac{\alpha^2(A + B(\tau))^2}{4A\beta}.
\]  

(EC.15)

Denote \( x = \delta^{r/2} \), \( a_1 = \frac{\lambda_1(\gamma_1, 1)^2}{(1-\delta \lambda_1)(1-\lambda_1)} \) and \( b_1 = 2\log(\hat{\lambda}_1)/\log(\delta) \). We have:

\[
\delta^{r/2}(A + B(\tau)) = x(A + a_1(1 - x^{b_1})) \triangleq H(x).
\]

Its derivative with respect to \( x \) is

\[
H'(x) = A + a_1 - a_1(b_1 + 1)x^{b_1}.
\]

Because \( a_1 > 0 \), \( a_1(b_1 + 1)x^{b_1} \) is strictly increasing as \( x \) increases from 0 to 1 (which is the range of \( \delta^{r/2} \) for \( \tau \in [0, \infty) \)), we have \( H'(0) > 0 \), and \( H'(x) \) is strictly decreasing in \( x \) and it crosses zero at most once. Therefore, \( H(x) \) is unimodal or strictly increasing in \( x \), or equivalently, \( \delta^{r/2}(A + B(\tau)) \) is unimodal or strictly decreasing in \( \tau \). This completes the proof of part (4). \( \square \)

With Claim EC.2, now we are ready to complete the proof of Proposition 3:

**Part (i):** The duration of the initial free-offering, \( \tau^*(p) \), is increasing in \( \gamma \) given by the Part (3) of Claim EC.2. The optimal revenue is increasing in \( \gamma \) because of Part (1) of Claim EC.2: since \( J(0; \tau, p) \) is increasing in \( \gamma \) for given \( \tau \) and \( p \), it is still increasing when we optimize over \( \tau \), i.e., \( J(0; \tau^*(p), p) \) is increasing in \( \gamma \).

**Part (ii):** The optimal revenue increases in \( \gamma \) by Part (1) of Claim EC.2. Next we show that the optimal solution \( (\tau^*, p^*) \) to \( J(0; \tau, p) \) are both increasing in \( \gamma \). Consider the revenue function \( J(0; \tau, p^*(\tau)) \) as a function of \( \tau \in \mathbb{Z}_+ \) when we optimize over \( p \). We will first show that \( \log (J(0; \tau, p^*(\tau))) \) is supermodular in \( \tau \) and \( \lambda_1 \) (it is clear that the domain of \( (\tau, \lambda_1) \) forms a lattice). Using (EC.15), we have

\[
\log (J(0; \tau, p^*(\tau))) = (\tau + 1) \log(\delta) + \log\left(\frac{\alpha^2}{4\beta}\right) - \log(A) + 2\log(A + B(\tau)).
\]
Note that \( \log(J(0; \tau, p^*(\tau))) \) is defined for \( \tau \in \mathbb{R}^+ \) and \( \lambda_1 \in (0, 1) \). To emphasize the continuous domain, within this proof, we use \( \bar{\tau} \) to denote the continuous variable. In the following we prove that \( \log(J(0; \bar{\tau}, p^*(\bar{\tau}))) \) is supermodular in \( \bar{\tau} \) and \( \hat{\lambda}_1 \). We compute the cross-derivative of \( \log(J(0; \bar{\tau}, p^*(\bar{\tau}))) \) w.r.t \( \bar{\tau} \) and \( \hat{\lambda}_1 \); Noting that \( A \) is not a function of \( \bar{\tau} \), we have:

\[
\frac{\partial^2 \log(J(0; \bar{\tau}, p^*(\bar{\tau})))}{\partial \hat{\lambda}_1 \partial \bar{\tau}} = \frac{\partial^2 \log(A + B(\bar{\tau}))}{\partial \hat{\lambda}_1 \partial \bar{\tau}} = \frac{\partial}{\partial \hat{\lambda}_1} \left( -\hat{\lambda}_1^{\tau+1} \log(\hat{\lambda}_1) \right)
\]

\[
= \frac{\partial}{\partial \hat{\lambda}_1} \left( -\hat{\lambda}_1^{\tau+1} \log(\hat{\lambda}_1) \left( -\hat{\lambda}_1^{\tau+1} \log(\hat{\lambda}_1) \right) \right)
\]

\[
= -\hat{\lambda}_1^{\tau+1} \left( (\bar{\tau} + 1)\hat{\lambda}_1^{\tau+1} + (\bar{\tau} + 1)(1 - \hat{\lambda}_1) \right)\left( \frac{1}{1 - \delta} + \hat{\lambda}_1^{\tau+1} \right)
\]

Therefore,

\[
\frac{\partial^2 \log(J(0; \bar{\tau}, p^*(\bar{\tau})))}{\partial \hat{\lambda}_1 \partial \bar{\tau}} \geq 0 \iff -\log(\hat{\lambda}_1) \left( \frac{\hat{\lambda}_1 \delta}{1 - \delta} + (\bar{\tau} + 1)\hat{\lambda}_1^{\tau+1} + (\bar{\tau} + 1)(1 - \hat{\lambda}_1) \right) \geq \frac{1 - \hat{\lambda}_1}{1 - \delta} + \hat{\lambda}_1(1 - \hat{\lambda}_1).
\]

Because \( 1 - \hat{\lambda}_1^\tau \leq \bar{\tau}(1 - \hat{\lambda}_1) \) for \( \hat{\lambda}_1 < 1 \), and \( -\log(\hat{\lambda}_1) > 1 - \hat{\lambda}_1 \), it suffices to show that

\[
\hat{\lambda}_1 \delta + (\bar{\tau} + 1)\hat{\lambda}_1^{\tau+1} + (\bar{\tau} + 1)(1 - \hat{\lambda}_1) \geq \frac{1 - \hat{\lambda}_1}{1 - \delta} + \hat{\lambda}_1 \bar{\tau} \iff \hat{\lambda}_1 + \hat{\lambda}_1 \delta + (\bar{\tau} + 1)(1 - \hat{\lambda}_1) \geq \frac{1}{1 - \delta}
\]

The last inequality holds because

\[
\frac{1}{1 - \delta} = \hat{\lambda}_1 + \hat{\lambda}_1 \delta + (1 - \hat{\lambda}_1) \leq \hat{\lambda}_1 + \hat{\lambda}_1 \delta + (\bar{\tau} + 1)(1 - \hat{\lambda}_1) \leq \frac{1 - \hat{\lambda}_1}{1 - \delta} + \hat{\lambda}_1 \bar{\tau}
\]

The nonnegativity of cross-derivatives implies that \( \log(J(0; \bar{\tau}, p^*(\bar{\tau}))) \) is supermodular in \( \bar{\tau} \) and \( \hat{\lambda}_1 \).

Next we argue that \( \log(J(0; \tau, p^*(\tau))) \) is supermodular in \( \tau \in \mathbb{Z}_+ \) and \( \hat{\lambda}_1 \) (note that we use \( \tau \), instead of \( \bar{\tau} \) for the discrete domain). This follows from the definition of supermodularity: let \( f(\bar{\tau}, \hat{\lambda}_1) : \mathbb{R}_+ \times (0, 1) \rightarrow \mathbb{R} \) be equal to \( \log(J(0; \bar{\tau}, p^*(\bar{\tau}))) \). We have:

\[
f(\bar{\tau} \lor \hat{\lambda}_1, \hat{\lambda}_1^\tau) + f(\bar{\tau} \land \hat{\lambda}_1, \hat{\lambda}_1^\tau) \geq f(\bar{\tau}, \hat{\lambda}_1) + f(\bar{\tau}, \hat{\lambda}_1^\tau)
\]

(Ec.16)

for \( \bar{\tau}, \hat{\lambda}_1, \hat{\lambda}_1^\tau \in (0, 1) \). Here we use the notation \( x \lor y = \max\{x, y\} \) and \( x \land y = \min\{x, y\} \). Inequality (Ec.16) also holds for any \( \bar{\tau}, \hat{\lambda}_1, \hat{\lambda}_1^\tau \in (0, 1) \) ensuring that \( \log(J(0; \tau, p^*(\tau))) \) is supermodular when we restrict the domain of \( \tau \) to be \( \mathbb{Z}_+ \). Note that the maximizer of \( \log(J(0; \tau, p^*(\tau))) \) coincides with that of \( J(0; \tau, p^*(\tau)) \), which is \( \tau^* \).
By Part (4) of Claim EC.2, the set of maximizers $\tau^* \in \mathbb{Z}_+$ is either a singleton or consists of two consecutive integers. By the lattice theory, the smallest maximizer and the largest maximizer are both increasing in $\hat{\lambda}_1$. Because $\hat{\lambda}_1$ is an increasing function of $\gamma$, we also have that $\tau^*$ is increasing in $\gamma$.

In order to prove that the optimal $p^*$ also increases with $\gamma$, note that we can write it as $p^*(\tau^*(\gamma), \gamma)$, with a slight abuse of notation. From Part (2) of Claim EC.2, fixing $\tau$, $p^*(\tau, \gamma)$ is increasing in $\gamma$. Moreover, for fixed $\gamma$, $p^*(\tau, \gamma)$ is increasing in $\tau$ from Proposition 1. We just proved that $\tau^*(\gamma)$ is increasing in $\gamma$. Therefore, $p^*(\tau^*(\gamma), \gamma)$ also increases in $\gamma$. $\square$

EC.3.2. Proof of Theorem 2:

Part (i): Suppose $M_1$ and $M_2$ are two balanced networks in the class. To show that $M_1$ and $M_2$ yield the same pricing policy and optimal revenue, we only need to show that for any policy $(\tau, p)$, the discounted revenue $J(0; \tau, p)$ given by (EC.2) are identical for the two networks: Note that for any balanced matrix with average degree $D$, all the row sums are all equal to $D$ as well. Using Fact 6, we have:

$$
\lambda_1(M_1) = \lambda_1(M_2) = D,
$$
$$
v_1(M_1) = v_1(M_2) = 1/\sqrt{n},
$$
$$
\langle v_i(M_1), 1 \rangle = \langle v_i(M_2), 1 \rangle = 0, \quad 2 \leq i \leq n.
$$

This implies that for any given policy, the discounted revenue of any two balanced networks are identical. This completes the proof of part (i).

Part (ii): Recall that we represent the eigenvalues and eigenvectors of $\hat{M}$ by $\hat{\lambda}_i$ and $v_i$. In order to prove part (ii), for any given policy $(\tau, p)$, we first decompose the discounted revenue $J(0; \tau, p)$ into infinite terms each containing a unique power of the eigenvalues. More specifically, we first prove the following claim:

**Claim EC.3.** For given $\tau$ and $p$, the discounted revenue $J(0; \tau, p)$ given by (EC.2) can be expressed as the following infinite sums:

$$
J(0; \tau, p) = \sum_{k=0}^{\infty} a_k \sum_{i=1}^{n} \hat{\lambda}_i^k \langle v_i, 1 \rangle^2 
$$

(EC.17)

$$
J(0; \tau, p) = \sum_{k=0}^{\infty} b_k \sum_{i=1}^{n} \lambda_i^k \langle v_i, 1 \rangle^2
$$

(EC.18)

where $a_k, b_k \geq 0$, for all $k \in \mathbb{Z}_+$, and they are independent of the eigenvalues and eigenvectors.
Proof of Claim EC.3: Equality (EC.17) follows from the following observation: the two terms in (EC.2), \( \frac{1}{1-\delta \hat{\lambda}_i} \) and \( \hat{\lambda}_i(1-\hat{\lambda}_i)/(1-\delta \hat{\lambda}_i)(1-\hat{\lambda}_i) \), can be expressed as an infinite-order polynomial of \( \hat{\lambda}_i \) with nonnegative coefficients:

\[
\frac{1}{1-\delta \hat{\lambda}_i} = \sum_{k=0}^{\infty} (\delta \hat{\lambda}_i)^k \\
\frac{\hat{\lambda}_i(1-\hat{\lambda}_i)/(1-\delta \hat{\lambda}_i)(1-\hat{\lambda}_i)}{(1-\delta \hat{\lambda}_i)} = (\hat{\lambda}_i + \hat{\lambda}_i^2 + \ldots + \hat{\lambda}_i^\tau) \sum_{k=0}^{\infty} (\delta \hat{\lambda}_i)^k.
\]

Equality (EC.18) follows from (EC.17) and the additional observation that \( \hat{\lambda}_i = \lambda_i + \mu \). \( \square \)

With Claim EC.3, we are now ready to prove part (ii): It suffices to show that for given \( \tau \) and \( p \), the revenue of an arbitrary network \( M \) is higher than that of a balanced network of the same average degree. Because of Part (i) of Theorem 2, it suffices to focus on comparing revenue of \( M \) with that of a complete graph. Further, utilizing (EC.18), we compare the revenues of the two networks term by term, and show that:

\[
\sum_{i=1}^{n} \lambda_i^k(v_i, 1)^2 \geq \sum_{i=1}^{n} (\lambda_i^C)^k(v_i^C, 1)^2 = nD^k, \quad k \in \mathbb{Z}_+
\]  

(EQ.19)

where \( \lambda_i^C \) and \( v_i^C \) are the eigenvalues and eigenvectors of a complete network with average degree \( D \).

In order to prove inequality (EC.19), we first use Facts 1 and 2 to express both sides of (EC.19) as follows:

\[
\sum_{i=1}^{n} \lambda_i^k(v_i, 1)^2 = \langle M^k1, 1 \rangle \\
nD^k = n1^{-k}\langle M1, 1 \rangle^k
\]

Note that \( \langle M^k1, 1 \rangle \) is the total number of walks of length \( k \) in the graph associated with the adjacency matrix \( M \). We complete the proof by applying Theorem 1 in Blakley and Roy (1965) as stated in Fact 8. \( \square \)

EC.3.3. Proof of Proposition 4:

Denote the entries of matrix \( C \) as \( \gamma \) and that of \( M \) as \( m_{ij} \). By the definition of \( C \) we have \( \gamma = D/(n-1) \). We only need to show that for given \( \tau \) and \( p \), the discounted revenue \( J(0; \tau, p) \) given by (EC.2) for a network \( M_w \) is increasing in \( w \).

Let \( \hat{\lambda}_i(w) \) and \( v_i(w) \) denote the eigenvalues and eigenvectors of \( \hat{M}_w \). Further note that, by definition, we have:

\[
\hat{m}_{ij} = \begin{cases} 
    m_{ij} & i \neq j \\
    \mu & i = j
\end{cases}, \quad \hat{m}_{w,ij} = \begin{cases} 
    (1-w)\gamma + wm_{ij} & i \neq j \\
    \mu & i = j
\end{cases}
\]
Similar to our proof strategy for part (ii) of Theorem 2, we use Claim EC.3 to show the desired monotonicity for each term $\sum_{i=1}^{n} \hat{\lambda}_i^k(w)(v_i(w),1)^2$ ($k \in \mathbb{Z}_+$) separately. For $k = 0$ and 1, we have $\sum_{i=1}^{n} v_i(w,1)^2 = n$ and $\sum_{i=1}^{n} \hat{\lambda}_i(w)(v_i(w),1)^2 = (\hat{\mathbf{M}}_w1,1) = n(D+\mu)$, both of which do not change with $w$. For $k \geq 2$, we have

$$\sum_{i=1}^{n} \hat{\lambda}_i^k(w)(v_i(w),1)^2 = (\hat{\mathbf{M}}_w^1,1) = \sum_{1 \leq i_1, \ldots, i_{k+1} \leq n} \prod_{j=1}^{k} (\hat{m}_{i_j,i_{j+1}} + (1-w)(\gamma + I_{i_j=i_{j+1}}(\mu-\gamma) - \hat{m}_{i_j,i_{j+1}})), \tag{EC.20}$$

where $I$ is the indicator function. In the above expression, we enumerate all the entries of the matrix power $\hat{\mathbf{M}}_w^k$. The next lemma allows us to relate the derivative of (EC.20) at $w = 1$ to the number of walks of different lengths on the network $\hat{\mathbf{M}}$.

**Lemma EC.2.** For $k \geq 2$, we have:

$$\frac{\partial}{\partial w} \sum_{i=1}^{n} \hat{\lambda}_i^k(w)(v_i(w),1)^2 \bigg|_{w=1} = \sum_{1 \leq i_1, \ldots, i_{k+1} \leq n} \sum_{j=1}^{k} \left( (\hat{\mathbf{M}}_w^1,1) - (\mu-\gamma)(\hat{\mathbf{M}}_w^{k-1},1) - \gamma(\hat{\mathbf{M}}_w^{k-2},1)(\hat{\mathbf{M}}_w^{k-1},1) \right).$$

Note that for an adjacency matrix $A$, the entry on the $i$th row and $j$th column of $A^k$ represents the number of walks of length $k$ from node $i$ to $j$ in the corresponding graph. Therefore, the inner product of the form $(A^k1,1)$ represents the total number of walks of length $k$ in the graph.

**Proof of Lemma EC.2:** We start by noting that $\sum_{i=1}^{n} \hat{\lambda}_i^k(w)(v_i(w),1)^2$ is a polynomial (of degree $k$) in $(1-w)$. Therefore, only the coefficient of the linear term $(1-w)$ contributes to $\frac{\partial}{\partial w} \sum_{i=1}^{n} \hat{\lambda}_i^k(w)(v_i(w),1)^2 \bigg|_{w=1}$. Therefore, we have:

$$\frac{\partial}{\partial w} \sum_{i=1}^{n} \hat{\lambda}_i^k(w)(v_i(w),1)^2 \bigg|_{w=1} = \sum_{1 \leq i_1, \ldots, i_{k+1} \leq n} \sum_{j=1}^{k} \left[ \hat{m}_{i_j,i_{j+1}} - I_{i_j=i_{j+1}}(\mu-\gamma) - \gamma \prod_{l=1 \ldots, k \neq j}^k \hat{m}_{i_l,i_{l+1}} \right].$$

We decompose the outer summation into three parts and simplify each part separately as follows:

$$\sum_{1 \leq i_1, \ldots, i_{k+1} \leq n} \sum_{j=1}^{k} \hat{m}_{i_j,i_{j+1}} = k \sum_{j=1}^{k} \prod_{l=1, l \neq j}^k \hat{m}_{i_l,i_{l+1}},$$

$$\sum_{1 \leq i_1, \ldots, i_{k+1} \leq n} \sum_{j=1}^{k} I_{i_j=i_{j+1}} \prod_{l=1, l \neq j}^k \hat{m}_{i_l,i_{l+1}} = k \prod_{j=1}^{k-1} \sum_{l=1, l \neq j}^k \hat{m}_{i_l,i_{l+1}},$$

$$\sum_{1 \leq i_1, \ldots, i_{k+1} \leq n} \sum_{j=1}^{k} \prod_{l=1, l \neq j}^k \hat{m}_{i_l,i_{l+1}} = k \sum_{j=1}^{k-1} \prod_{l=1, l \neq j}^k \hat{m}_{i_l,i_{l+1}}.$$
Plugging these three equations back into (EC.21) and some algebra, we get:

\[
\frac{\partial}{\partial w} \sum_{i=1}^{n} \lambda_i^k(w)\langle v_i(w), 1 \rangle^2 \bigg|_{w=1} = k \sum_{i=1}^{n} \sum_{j=1}^{k} \prod_{\ell=1}^{j} \hat{m}_{ij+1} - k(\mu - \gamma) \sum_{i=1}^{n} \prod_{j=1}^{k} \hat{m}_{ij+1} \\
- \gamma \sum_{i=1}^{n} \sum_{j=1}^{k} \left( \prod_{\ell=1}^{j-1} \hat{m}_{ij+1} \right) \sum_{l=1}^{k-1} \prod_{j=1}^{k} \hat{m}_{ij+1} \\
= k(\hat{M}^{k-1}, 1) - k(\mu - \gamma)(\hat{M}^{k-1}, 1) - \gamma \sum_{j=1}^{k} (\hat{M}^{j-1}, 1)(\hat{M}^{k-j}, 1) \\
= \sum_{l=1}^{k} \left( (\hat{M}^{l}, 1) - (\mu - \gamma)(\hat{M}^{k-l}, 1) - \gamma(\hat{M}^{l-1}, 1)(\hat{M}^{k-l}, 1) \right),
\]

which completes the proof of the lemma. □

Denote \( p_i = \langle v_i, 1 \rangle^2 \geq 0 \). We have: \( \sum_{i=1}^{n} p_i = n \) and \( \sum_{i=1}^{n} \lambda_i p_i = n \mu + n(n - 1) \gamma \). By Facts 1 and 2, we can express \( \sum_{i=1}^{n} \lambda_i \hat{\lambda}_i^k \) and thus using Lemma EC.2, we have:

\[
\frac{\partial}{\partial w} \sum_{i=1}^{n} \lambda_i^k(w)\langle v_i(w), 1 \rangle^2 \bigg|_{w=1} = \sum_{i=1}^{n} \left( \frac{1}{n} \sum_{j=1}^{k} p_j \lambda_i^j - \frac{\mu - \gamma}{n} \sum_{j=1}^{k} \lambda_i^j \lambda_i^{k-j} - \gamma \sum_{j=1}^{k} \lambda_i^{k-j} \lambda_i^{j-1} \right) \\
= \sum_{i=1}^{n} \sum_{j=1}^{k} p_j \left( \frac{1}{n} \lambda_i^j - \frac{\mu - \gamma}{n} \lambda_i^{k-j} - \gamma \lambda_i^{j-1} \lambda_i^{k-j} \right).
\] (EC.22)

In the following, we rewrite (EC.22) in terms of moments of a related random variable that we define next. Let \( X \in \{\hat{\lambda}_1, \hat{\lambda}_2, \ldots, \hat{\lambda}_n\} \) be a random variable with the following distribution: \( P(X = \hat{\lambda}_i) = p_i/n \) for \( i = 1, \ldots, n \). Under the proposition assumption that \( \mu \geq -\lambda_n \), we have \( X \geq 0 \). Moreover, \( E[X] = \sum_{i=1}^{n} p_i \hat{\lambda}_i/n = \mu + (n - 1) \gamma \). Let \( Y \) be an independent copy of \( X \). Using these random variables, we rewrite (EC.22) as follows:

\[
\frac{\partial}{\partial w} \sum_{i=1}^{n} \lambda_i^k(w)\langle v_i(w), 1 \rangle^2 \bigg|_{w=1} = n \sum_{i=1}^{k} E \left[ X^k - (\mu - \gamma)X^{k-1} - n \gamma X^{k-1}Y^{l-1} \right] \\
= n \sum_{l=1}^{k} \left( E[X^k] + \gamma E[X^{k-1}] - \mu E[X^{k-1}] - n \gamma E[X^{k-l}]E[X^{l-1}] \right).
\] (EC.23)

For the moments of nonnegative random variables, applying the Holder’s inequality we have that for \( 0 < r \leq s \), \( E[X^r]^{1/r} \leq E[X^s]^{1/s} \). Therefore,

\[
E[X^{k-l}] = E[X^{k-1-l}]/(k-l) \geq E[X^{k-l}]E[X^{l-1}] \\
E[X^{k-l}] \geq E[X]E[X^{k-1}] = (\mu + (n - 1) \gamma)E[X^{k-1}].
\]

Plugging both inequalities into (EC.23), we have

\[
\frac{\partial}{\partial w} \sum_{i=1}^{n} \lambda_i^k(w)\langle v_i(w), 1 \rangle^2 \bigg|_{w=1} \geq n \sum_{l=1}^{k} \left( E[X^k] - \mu E[X^{k-1}] - (n - 1) \gamma E[X^{k-l}] \right) \geq 0.
\]
This implies that the revenue is increasing in \( w \) when \( w = 1 \).

For \( w \in [0, 1) \), we define \( M'_v \triangleq vM_w + (1 - v)C = vwM + (1 - vw)C \). By the previous result, the revenue for \( M'_v \) is increasing in \( v \) at \( v = 1 \). This is equivalent to the monotonicity of the revenue at \( w \) for \( M_w \). Thus, we have completed the proof. \( \square \)

**EC.4. Proof of the Results in Section 6**

Before proceeding to prove Proposition 5, we first define a few notations, and state an auxiliary lemma. Denote the complete network that the heuristic is based on by \( C \) with \( c_{ij} = D/(n - 1), \) \( 1 \leq i, j \leq n \). Denote its eigenvalues by \( \lambda_i^C \). We first introduce the following lemma.

**Lemma EC.3.**

\[
\lambda_1 \geq \lambda_i^C \geq \frac{D}{D + \Delta} \lambda_1 > 0
\]

**Proof of Lemma EC.3:** By the Perron-Frobenius theorem (Fact 7), we have \( \lambda_1, \lambda_i^C > 0 \). By Fact 6 and a standard result in spectral graph theory, we have: \( \lambda_i^C = D = \sum_{i,j} m_{ij}/n \leq \lambda_1 \) (for example see Chung 1997). To show the second inequality, consider a diagonal matrix \( D = diag(\Delta_1, \Delta_2, \ldots, \Delta_n) \). Clearly, \( M + D \) is a balanced network and its first eigenvalue \( \lambda_1^D \) is the row sum \( (\sum_{i,j} m_{ij} + \sum_{i=1}^n \Delta_i)/n \). Therefore,

\[
\lambda_1^C = \frac{\sum_{i,j} m_{ij}}{\sum_{i,j} m_{ij} + \sum_{i=1}^n \Delta_i} \lambda_1^D.
\]

Next we argue that \( \lambda_1^D \geq \lambda_1 \). By an alternative representation of eigenvalues, we have

\[
\lambda_1 = \max_{x^T x = 1} x^T M_x \quad \text{and} \quad \lambda_1^D = \max_{x^T x = 1} x^T (M + D) x.
\]

Because \( D \) is a diagonal matrix with nonnegative entries, we have \( \lambda_1^D \geq \lambda_1 \). Therefore,

\[
\lambda_1^C \geq \frac{\sum_{i,j} m_{ij}}{\sum_{i,j} m_{ij} + \sum_{i=1}^n \Delta_i} \lambda_1^D \geq \frac{\sum_{i,j} m_{ij}}{\sum_{i,j} m_{ij} + \sum_{i=1}^n \Delta_i} \lambda_1 = \frac{D}{D + \Delta} \lambda_1
\]

and we have completed the proof. \( \square \)

**EC.4.1. Proof of Proposition 5:**

Let \( \hat{J} \) denote the revenue when the underlying network is \( C \). Let \( \lambda_i^C \) denote \( \lambda_i^C + \mu \). By Lemma EC.3, \( \lambda_1 \geq \lambda_i^C \geq D \lambda_1/(D + \Delta) \) and \( \hat{\lambda}_1 \geq \hat{\lambda}_i^C \geq D \hat{\lambda}_1/(D + \Delta) \). By those inequalities, we can derive the following inequalities, which will be used later on.

\[
\hat{\lambda}_1^C \geq \frac{D}{D + \Delta} \hat{\lambda}_1 \quad \text{(EC.24)}
\]
\[
\frac{1}{1 - \delta \lambda_i^C} \geq \frac{D + \Delta}{D + \Delta - \delta D \lambda_i} = \frac{(1 + \Delta/D)(1 - \delta \hat{\lambda}_i)}{1 + \Delta/D - \delta \lambda_i} \frac{1}{1 - \delta \lambda_i} \quad \text{ (EC.25)}
\]

\[
\frac{1}{1 - \lambda_i^C} \geq \frac{D + \Delta}{D + \Delta - \delta \lambda_i} = \frac{(1 + \Delta/D)(1 - \hat{\lambda}_i)}{1 + \Delta/D - \lambda_i} \frac{1}{1 - \lambda_i} \quad \text{ (EC.26)}
\]

\[
\frac{1 - (\lambda_i^C)^\tau}{1 - \lambda_i^C} = \frac{\sum_{k=0}^{\tau - 1} (\lambda_i^C)^k}{\sum_{k=0}^{\tau - 1} \lambda_i^C} \geq \min_{k=0, \ldots, \tau - 1} \frac{(\lambda_i^C)^k}{\lambda_i^C} \frac{1 - \hat{\lambda}_i}{1 - \lambda_i} \geq \left(\frac{\lambda_i^C}{\hat{\lambda}_i}\right)^{\tau - 1} \frac{1 - \hat{\lambda}_i}{1 - \lambda_i}. \quad \text{ (EC.27)}
\]

From the proof of Part (ii) of Theorem 2, a complete graph \( C \) generates the lowest revenue for any given policy \((\tau, p)\). Thus \( J(0; \tau, p) \geq \hat{J}(0; \tau, p) \) (With a slight abuse of notation, we use \( \hat{J} \) for the optimal revenue of \((\tau^*, \bar{p}^*)\) for the complete network \( C \), and \( \hat{J}(0; \tau, p) \) for a given policy \((\tau, p)\).) Therefore, \( J_h \geq \hat{J}(0; \tau^*, \bar{p}^*) \geq \hat{J}(0; \tau, p) \) when we plug in the optimal solution for \( C \) into the inequality.

We first show

\[
\frac{J_h}{\hat{J}} \geq \frac{(1 - \delta \hat{\lambda}_i)(1 - \hat{\lambda}_i)}{(1 + \Delta/D - \delta \lambda_i)(1 + \Delta/D - \lambda_i)}. \quad \text{(EC.28)}
\]

According to (EC.2), for any policy \( \tau \) and \( p \), we have

\[
J_h \geq \hat{J} \geq \hat{J}(0; \tau, p) = \frac{p(\alpha - \beta p)\delta^{\tau + 1}n}{(1 - \delta)(1 - \delta \lambda_i)} + \frac{\alpha \delta^{\tau + 1} \hat{\lambda}_i p(1 - \hat{\lambda}_i)^\tau n}{(1 - \delta)(1 - \hat{\lambda}_i)}
\]

Using (EC.24), (EC.25), (EC.26) and the fact that \( 1 - (\lambda_i^C)^\tau \geq 1 - \hat{\lambda}_i^\tau \), we have:

\[
J_h \geq \frac{(1 + \Delta/D)(1 - \delta \hat{\lambda}_i)}{1 + \Delta/D - \delta \lambda_i} \frac{p(\alpha - \beta p)\delta^{\tau + 1}n}{(1 - \delta)(1 - \delta \lambda_i)} + \frac{\alpha \delta^{\tau + 1} \hat{\lambda}_i p(1 - \hat{\lambda}_i)^\tau n}{(1 - \delta)(1 - \hat{\lambda}_i)}
\]

Because \( \sum_{i=1}^{n} \langle \mathbf{v}_i, \mathbf{1} \rangle^2 = n \), we have:

\[
J_h \geq \frac{(1 - \delta \hat{\lambda}_i)(1 - \hat{\lambda}_i)}{1 + \Delta/D - \delta \lambda_i} \frac{\sum_{i=1}^{n} p(\alpha - \beta p)\delta^{\tau + 1} \langle \mathbf{v}_i, \mathbf{1} \rangle^2}{(1 - \delta)(1 - \delta \lambda_i)} + \frac{\alpha \delta^{\tau + 1} \hat{\lambda}_i p(1 - \hat{\lambda}_i)^\tau n}{(1 - \delta)(1 - \hat{\lambda}_i)}
\]

Because \( \hat{\lambda}_i \geq |\hat{\lambda}_i| \) (Perron-Frobenius theorem as stated in Fact 7), \( \frac{\hat{\lambda}_i(1 - \hat{\lambda}_i)}{(1 - \delta \lambda_i)(1 - \lambda_i)} \leq \frac{\hat{\lambda}_i(1 - \hat{\lambda}_i)}{(1 - \delta \lambda_i)(1 - \lambda_i)} \) and we have:

\[
J_h \geq \frac{(1 - \delta \hat{\lambda}_i)(1 - \hat{\lambda}_i)}{1 + \Delta/D - \delta \lambda_i} \frac{\sum_{i=1}^{n} p(\alpha - \beta p)\delta^{\tau + 1} \langle \mathbf{v}_i, \mathbf{1} \rangle^2}{(1 - \delta)(1 - \delta \lambda_i)} + \frac{\alpha \delta^{\tau + 1} \hat{\lambda}_i p(1 - \hat{\lambda}_i)^\tau n}{(1 - \delta)(1 - \hat{\lambda}_i)}
\]

Because the inequality holds for any \( \tau \) and \( p \), it also holds for the optimal solution \((\tau^*, \bar{p}^*)\) for \( \mathbf{M} \). Thus we have shown (EC.28).
Next we show
\[
\frac{J_h}{J} \geq \frac{(1 - \delta)}{(1 + \Delta/D - \delta)} \left( \frac{1}{1 + \Delta/D} \right)^{\tau+1}.
\]  
(EC.29)

Similarly, by Lemma EC.1 and (EC.24) - (EC.27), for \( \tau \leq \lceil -\log(4)/\log(\delta) \rceil \), we have
\[
J_h \geq \tilde{J}(0; \tau, p) = \frac{p(\alpha - \beta_p)\delta^{\tau+1}n}{(1 - \delta)(1 - \delta \lambda_1^C)} + \frac{\alpha \delta^{\tau+1}\lambda_1^C p(1 - (\lambda_1^C)\tau)n}{(1 - \delta \lambda_1^C)(1 - \lambda_1^C)}
\]
\[
\geq \frac{(1 + \Delta/D)(1 - \delta \lambda_1)}{1 + \Delta/D - \delta \lambda_1} \frac{p(\alpha - \beta_p)\delta^{\tau+1}n}{(1 - \delta)(1 - \delta \lambda_1)} + \frac{1 - \delta \lambda_1}{1 + \Delta/D - \delta \lambda_1} \left( \frac{D}{D + \Delta} \right)^{\tau} \frac{\alpha \delta^{\tau+1}\lambda_1^C p(1 - \lambda_1^C)n}{(1 - \delta \lambda_1)(1 - \lambda_1)}
\]
\[
\geq \frac{1 - \delta \lambda_1}{1 + \Delta/D - \delta \lambda_1} \left( \frac{D}{D + \Delta} \right)^{\tau} \frac{\alpha \delta^{\tau+1}\lambda_1^C p(1 - \lambda_1^C)(\nu_i, 1)^2}{(1 - \delta \lambda_1)(1 - \lambda_1)} + \sum_{i=1}^{n} \frac{\alpha \delta^{\tau+1}\lambda_1^C p(1 - \lambda_1^C)(\nu_i, 1)^2}{(1 - \delta \lambda_1)(1 - \lambda_1)}
\]
\[
\geq \frac{1 - \delta \lambda_1}{1 + \Delta/D - \delta \lambda_1} \left( \frac{D}{D + \Delta} \right)^{\tau} \frac{\alpha \delta^{\tau+1}\lambda_1^C p(1 - \lambda_1^C)(\nu_i, 1)^2}{(1 - \delta \lambda_1)(1 - \lambda_1)} J(0; \tau, p).
\]

In Proposition 2, we proved the following upper-bound on the optimal stopping time: \( \tau^* \leq \tilde{\tau} = \lceil -\log(4)/\log(\delta) \rceil \). Therefore,
\[
J_h \geq \tilde{J}(0; \tau^*, p^*) \geq \frac{1 - \delta}{1 + \Delta/D - \delta} \left( \frac{1}{1 + \Delta/D} \right)^{\tau+1} J.
\]

Thus we have proved (EC.29). Taking the maximum of the RHS of (EC.28) and (EC.29) completes the proof. □