NET Institute*

www.NETinst.org

Working Paper #17-13

September 2017

**Blockbuster or Niche?
Competitive Strategy under Network Effects**

Yinbo Feng

School of Management, Fudan University

Ming Hu

Rotman School of Management, University of Toronto

* The Networks, Electronic Commerce, and Telecommunications (“NET”) Institute, <http://www.NETinst.org>, is a non-profit institution devoted to research on network industries, electronic commerce, telecommunications, the Internet, “virtual networks” comprised of computers that share the same technical standard or operating system, and on network issues in general.

Blockbuster or Niche?

Competitive Strategy under Network Effects*

Yinbo Feng¹, Ming Hu²

1. School of Management, Fudan University, Shanghai, China, 200433, yinbofeng13@fudan.edu.cn

2. Rotman School of Management, University of Toronto, Toronto, Ontario, Canada M5S 3E6, ming.hu@rotman.utoronto.ca

We provide a theory unifying the long tail and blockbuster phenomenon. Specifically, we analyze a three-stage game where the firms first make entry decisions, then decide on the investment in its product and lastly customers sequentially arrive to make purchase decisions based on product quality and historic sales under the network effect. We analytically show that a growing network effect always contributes to the demand concentration on a small number of products. However, product variety and quality investments, as an outcome of firms' ex-ante competitive decisions, may increase or decrease, as the network effect grows. When the network effect parameter is smaller than a threshold, the increasing network effect would shift more demand towards the products with higher qualities, preempting more products from entering the market ex ante and inducing firms to adopt the blockbuster equilibrium strategy by making larger quality investment. When the network effect is stronger than the threshold, the increasing network effect would make the market easily concentrated to a few products. Even some low quality ones may have chances to become a "hit." Interestingly, in this case, the ex-ante equilibrium product variety would be broader and firms adopt the niche equilibrium strategy by making smaller quality investment. We empirically test the theory with the movie box office data and find strong supporting evidence.

1. Introduction

The "long tail" theory was celebrated by *BusinessWeek* as the biggest idea of the year 2004, soon after the book "*The Long Tail*" by Chris Anderson was published. The book argues that our culture and economy is shifting away from a relatively small number of "hit" products towards a huge number of niche products. In other words, demand would become less concentrated and the tail of the sales rank distribution would be longer and fatter over time.

The long tail theory predicted that there would be more marketing resources to be allocated to niche offerings than in "hit" products, i.e., shifting towards the niche strategy, especially in the

*

We are grateful for the financial assistance provided by the NET Institute, www.netinst.org.

digitized world. For example, Chris Anderson pointed out in his book that a significant portion of Amazon.com’s sales come from obscure books that are not available in brick-and-mortar stores. The same seems also true for the music industry. However, some other cultural industries may tell a different story. [Elberse \(2008\)](#) argued that the importance of individual best sellers in the home-video industry is not diminishing over time. The concentration of most popular titles is growing. Evidence emerges against the long tail theory in the movie and video game industries as well. [Stewart \(2013\)](#) reported that studios have been shifting their resources towards the big-budget movies. That is what is known as a blockbuster strategy—focusing the limited production and marketing resources on a small number of likely best-selling products. This strategy is based on the notation that the “hit” product still plays an important role in today’s world.

In this paper, we aim at providing a theory unifying the long tail theory and blockbuster phenomenon. We postulate one key factor behind these phenomena as the *network effect*—the effect that the consumption of a good by one has an impact on the value of that product to others ([Economides 1996](#)). Intuitively, the blockbuster phenomenon has a lot to do with the (positive) network effect. For one thing, people who are apt to purchase the popular product due to the network effect would make the popular even more so. For another thing, producers, like studios, are willing to put more resources, e.g., advertising budgets and production resources, to produce a possibly blockbuster anticipating that the product will become a “hit” in future due to the network effect. However, no one can exactly predict a blockbuster. A high-budget, mainstream product—designed to cater to a broad range of tastes ([Sun 2012](#)) does not necessarily turn out a blockbuster. A low-budget, niche product—designed to cater to only a small group of customers ([Sun 2012](#)) may be a surprising blockbuster ex post, if it becomes popular sufficiently early on due to the market uncertainty and strong network effect. This can attract more and more niche products to enter the market while very few of them may gain significant market shares. In this sense, the (positive) network effect seemingly can also contribute to the phenomenon of long and flat tail. Thus, given the increasing network effect over time it is not surprising that in the iTunes App Store the downloads are more concentrated to a handful of top games while a rapidly increasing number of games sell very rarely or never. Roughly estimated, the number of available gaming apps in the iTunes App Store increased from 151,460 in 2013 to 631,091 in 2016. The sales of the 100th game at the yearly grossing top increased from \$2.8 million in 2013 to \$6.3 million in 2016 but the sales of the 500th game decreased from \$0.38 million in 2013 to \$0.25 million in 2016. Around the 5000th, games’ sales are very close to zero in 2013–2016. Nevertheless, how the network effect exactly leads

to the blockbuster or long tail phenomenon seems not clear. This would be the central focus of this paper.

In particular, we analyze a three-stage game where the firms first make entry decisions, then decide on the investment in its product and lastly customers sequentially arrive to make purchase decisions based on product quality and historic sales subject to the network effect. We backwards solve this Stackelberg game. In the third stage, sequentially arriving customers, upon arrival, make a purchase choice among the set of available products in the market, based on a customer discrete choice model. We assume each firm operates with a single product. The attraction value for each product depends on two attributes: the intrinsic quality of the product invested by the firm before the product release, and the current market share of the product. The term “quality” in economics and marketing is often referred to as a measure to vertically differentiate a group of products in the sense that a high-quality product generates higher surplus than a low-quality product for all customers. We abuse the term slightly and interpret “quality” as the *intrinsic appeal* for a product. In this paper, a high-quality product means it has a high probability to be purchased by a randomly chosen customer in the absence of network effect. In this sense, we define a high-quality product as a mainstream product with a high budget investment and a low-quality product as a niche product with a low budget investment. Moreover, the network benefit from consuming a product is determined by the current market share of the product and a network effect magnitude parameter (denoted by β in our model). In the base model, the network effect parameter β is assumed to be the same for all products in the industry. We adopt the special type of stochastic processes, called *nonlinear Pólya urn process*, to estimate the long-term (random) market share for each product. Before the (random) sales process starts, in the first stage, potential competitive firms decide on entering the market with an entry cost and in the second stage, all firms entering the market simultaneously decide on the quality for their product, in anticipation of future competitive market dynamics.

Using this analytical model, we show that an increasing network effect always contributes to the demand concentration on a small number of products regardless of whether the prevailing network effect is weak or strong. However, product variety in the market, as an outcome of firms’ ex-ante competitive decisions, may expand or shrink, as the network effect increases. When the network effect parameter is smaller than a threshold, the increasing network effect would shift more demand towards the products with higher qualities, preempting more products from entering the market ex ante. When the network effect is stronger than that threshold, the increasing network effect would make the market easily concentrated to a few products; even some low quality ones may have

chance to become a “hit.” Interestingly, in this case, the ex-ante equilibrium product variety would be broader if the network effect becomes stronger, as more firms bet on their products to become a “hit.” That is, increasing network effect not only intensifies the “demand concentration” ex post but also leads to the “longer tail” ex ante, observed in the curve of sales rank distribution. This result is consistent with the anecdotal evidence mentioned earlier. In our model, firms also make budget decisions for their products. With the same model, we show that an increasing network effect can intensify budget (quality) competition, i.e., shifting towards the blockbuster strategy, when the network effect parameter is weaker than the same threshold mentioned above; but can also alleviate budget (quality) competition, i.e., shifting towards the niche strategy, when the network effect parameter is greater than that threshold.

The main contribution of our analytical model lies in linking the blockbuster / long tail phenomenon and blockbuster / niche strategy in an industry to the network effect parameter in that industry. Hence, the network effect magnitude can be viewed as a key to understand a specific cultural industry. One can estimate the network effect parameter in an industry from data of that industry by statistical estimation tools. For example, in this paper, we resort to the Maximum Likelihood Estimation (MLE) method to estimate the network effect parameter in the movie industry from a comprehensive dataset collected from IMDb and Box Office Mojo. One concern, however, is that, in many industries such as the movie industry, the network effect parameter varies depending on the characteristics of a product. Directly estimating the parameter for the overall industry can be problematic. Instead, in the empirical analysis to test our theory, we categorize all movies into different groups such that movies in the same group released in the same week, and then identify each group by the maximum budget in that group. We find that the network parameter for each group is decreasing on average in the maximum budget of the group. That is, when at least one high-budget movie joins the competition, the network effect parameter would be small; and when all competitive movies are low-budgets niche ones, the network effect parameter would be large. This finding supports that the weak and strong network effects are both prevailing in the movie industry. According to our theory, on one hand, the weak network effect contributes to the “blockbuster” phenomenon and the “demand concentration” on the high-budget movies. On the other hand, the strong network effect contributes to the “demand concentration” on some movies which could be high-budget products or even low-budget ones. Moreover, the weak network effect leads to the intensified quality competition, i.e., shifting towards the blockbuster strategy, which seems adopted by major Hollywood studios nowadays aiming at producing a “hit” product. The strong network effect leads to softened quality competition, i.e., shifting towards the niche strategy,

which seems adopted by a large number of low-budgeted producers of digital films. This empirical analysis provides strong evidence to support our theory. More importantly, it provides a concrete approach to apply our framework into a specific industry and interpret and predict the evolution of the sales rank distribution curve, and to predict and prescribe in practice which type of products is more appropriate for the blockbuster or niche strategy, as the network effect is expected to be even stronger over time, due to the growing impacts of social media, review websites and recommendation systems.

The rest of this paper is organized as follows. Section 2 reviews the most related literature. Section 3 presents the analytical model. Section 4 contains the equilibrium analysis and key analytical results. The complementary numerical study is developed in Section 5. We apply our theory into the movie industry with an empirical analysis in Section 6. Concluding remarks are provided in Section 7. All the proofs and preliminaries are relegated to the appendix.

2. Literature Review

One motivation of our research is the emergence of products that exhibit network effects. Most recent works attempt to explicitly address the issue of dynamic competition between proprietary networks. Doganoglu (2003), Mitchell and Skrzypacz (2006) and Cabral (2011) derive the Markov perfect equilibrium of an infinite-period game where a customer’s utility is an increasing function of past market shares. In many aspects, their analysis goes beyond our work by analyzing compatibility decisions and dynamic price competition. However, they do not take the endogenous number of competitive firms into consideration. In practice, like in the movie industry, the studio can decide whether to produce a movie and invest how much. In our work, we focus on the firms’ entry decisions that determine the product variety in an industry and the budget decision that determines the product quality.

Analytical studies have recently begun to examine how the network effect affects firms’ policy on assortment (Wang and Wang 2016, Abeliuk et al. 2016). Different from these studies, which focus on a monopolistic setting, in this paper, we characterize the equilibrium in a competitive setting and shed light on the evolution of the product variety in an industry as the network effect changes over time. Instead of analyzing the optimal assortment and positioning policies for a specific firm, our work, from a macro perspective, contributes to interpret and predict which competitive strategy, blockbuster or niche, is/will be the trends in the market.

Our paper contributes to the long tail theory and blockbuster phenomenon. The “long tail” theory, popularized by Anderson (2006), predicts that niche products would comprise increasing

market share, while the demand for popular or “hit” products would continue to decrease. The long tail effect has generated widespread interest in academic circles. [Brynjolfsson et al. \(2010\)](#) provide a review of the research on the long tail effect. They categorize the plausible drivers of the long tail effect into demand-side and supply-side drivers. [Cachon et al. \(2008\)](#) points out by an analytical framework that lowered search cost can further encourage firms to enlarge their assortment to satisfy customers’ heterogeneous preferences. [Brynjolfsson et al. \(2011\)](#) and [Zentner et al. \(2013\)](#) empirically show that the online channel exhibits lower demand concentration because of the lower search cost for customers. [Oestreicher-Singer and Sundararajan \(2010\)](#) find that the peer-based recommendations contribute to the redistribution of the demand from blockbusters to niches. Other theories and empirical evidence are found to support the long tail effect by [Tucker and Zhang \(2011\)](#), [Kumar et al. \(2014\)](#), etc. In contrast, other industry studies highlight that despite the promise of digitization and the long tail, the greatest revenue continues to be concentrated in a small percentage of blockbusters, and not in the long tail ([Elberse and Oberholzer-Gee 2006](#), [Elberse 2008](#), [Hosanagar et al. 2013](#)). Some analytical models by [Hervás-Drane \(2015\)](#) and [Fleder and Hosanagar \(2009\)](#) show that selection-biased recommendation systems can lead to sales concentration and reduce the demand diversity. In this work, our analytical framework, coupled with the dataset in movie industry, shows that the blockbuster phenomenon and long tail effect can be due to different levels of network effect in an industry or with different market segments of an industry.

Some works empirically evaluate the effect of increasing product variety on demand concentration ([Hinz et al. 2011](#), [Zhou and Duan 2012](#), [Tan et al. 2016](#)). Different from these studies, we consider the network effect as a main driver for demand generation in an industry and, more importantly, endogenize the product variety in that industry.

Our work is also related to the active research on strategic positioning in response to technological changes in industries. [Benner and Waldfogel \(2016\)](#) take a first step in empirically studying how the technological change in the recorded music industry alters the different labels’ strategic positioning. The authors find that, due to technological changes, the independent record labels appear to adopt the cost reductions, pursuing more music releases for smaller and lower revenue audiences but the major record labels continue to shift resources towards releasing successful artists that have broad market appeal. Our finding is consistent with theirs. The difference is that we exhibit an analytical framework to study this problem and we further identify a more specific driver—the network effect that moderates the relationship between technological changes and firms’ responsive positioning.

From a methodological point of view, our framework has various similarities with [Maldonado et al. \(2015\)](#). In their framework, customer choices are captured by a discrete choice model in which

product utilities depend on the song’s quality, position bias, and a social influence signal representing past purchases. The authors define a social signal r which is equivalent to the parameter β in our model. They characterize the unique consumption equilibrium on the demand side when $r < 1$ and the asymptotic behavior of the consumption equilibrium when r closes to 1. They show that the demand would be concentrated on a smaller number of products and hence the product variety would shrink when r goes from 0 to 1, and that the market becomes unpredictable and goes most likely to a monopoly for *some* product when $r > 1$. Based on their framework, we further characterize the case of $r \geq 1$ analytically by obtaining the probability that each product goes to a monopoly when $r \geq 1$ and showing that the product variety would expand when r increases from 1 to infinity. [Maldonado et al. \(2015\)](#) focus on the position assignment that ranks the products in cultural markets with social influence. Unlike their paper, our work focuses on the product variety in an industry with network effect and on unifying the long tail theory and blockbuster phenomenon in one model.

The closest theoretical work to ours is [Bar-Isaac et al. \(2012\)](#), because both papers provide a theory to unify the long tail and blockbuster phenomena and both analysis can lead to long tail and blockbuster effects arising simultaneously. In particular, [Bar-Isaac et al. \(2012\)](#) and our paper both allow for a richer choice of strategies for firms on designing their products and consider consumers’ choices among different products. They model the search cost as key driver, based on the search literature as mentioned above. Different from their work, we develop a theory by highlighting the network effect as a key driver.

3. The Model

We consider a marketplace with a large number of potential competitive firms, who simultaneously make entry decisions. Any firm who enters the market must incur a fixed entry cost $K \geq 0$. A firm, say i , sells a single product incurs an investment cost, cq_i , in designing and production that is proportional to the quality level $q_i \geq 0$. The term “quality” in the economics, marketing and operations is often labeled as a measure to vertically differentiate a group of goods in the sense that a high-quality good results in higher utilities than a low-quality good for all customers. We slightly abuse the term “quality” in our setting and interpret it as the intrinsic appeal for a product. In this paper, a higher-quality product means it has a higher probability to be purchased by a randomly chosen customer in the absence of network effects. In this sense, we refer to a high-“quality” product as a mainstream product and a low-“quality” product as a niche product. For simplicity, the parameters $K \geq 0$ and $c \geq 0$ are assumed to be identical for all firms. They

characterize the production technology prevailing in an industry, with higher c and K corresponding to less efficient production technology. For a professional film studio that wishes to release a movie in the forthcoming summer, the entry cost K can be thought of as the costs associated with story selection and market interests research, and the cost of the quality improvement cq_i can be thought of as the costs of contents design, casting, production and visual effects. As in the film industry, fulfilling each consumption by consumers incurs a relatively negligible cost, after the moving is made and distributed. Hence we assume $K + cq_i$ is the total possible cost for firm i . If there are n firms entering the market, without loss of generality, we always label them by $i = 1, 2, \dots, n$ and the number of competing firms, n , is common knowledge to all firms. Let $\mathbf{q} = (q_1, \dots, q_n)$ denote the vector of every product's quality. We assume the price p for consumers is exogenously determined and the same for all products. This assumption allows us to focus on the competition among products from a single dimension of quality design.

We consider an infinite-period, dynamic model where customers arrive sequentially, one per time period. When a customer enters the market and faces the n products in the market, she is interested in buying one from the n number of choices. For simplicity, we do not consider the “no purchase” option, however, our results carry over the case with a constant no purchase option. The customer's purchase decision is affected by both the quality of the n products and the network effect. More specifically, upon arrival at time $t + 1$, a customer is able to observe the aggregate purchase decisions of her predecessors, denoted by $\mathbf{d}(t) = \{d_i(t) : i = 1, \dots, n\}$, with $\mathbf{d}(0) = \{1, 1, \dots, 1\}$. Let $\mathbf{x}(t) = \{x_i(t) : i = 1, \dots, n\}$ be the vector of the market share of the n products at time t , i.e.,

$$x_i(t) = \frac{d_i(t)}{d_1(t) + \dots + d_n(t)}.$$

A customer's utility from purchasing one unit of product i at time $t + 1$ is assumed to be

$$U_i(q_i, d_i(t)) = \ln q_i + \beta \ln d_i(t) - p + \xi_i,$$

where $\beta > 0$. The first and the last terms represent the customer's stand-alone valuation (intrinsic utility) of buying the product, where ξ_i is a random variable that represents customer-specific idiosyncrasies. The second term represents the customer's utility from the network externality or the utility from conforming to other people. Under the assumption that $\{\xi_i\}$'s are i.i.d., following a Gumbel distribution, the probability that the customer selects product i is given by the following discrete choice model, analogous to the MNL model:

$$\phi_i(\mathbf{x}(t)) = \frac{q_i d_i^\beta(t)}{\sum_{i=1}^n q_i d_i^\beta(t)} = \frac{q_i x_i^\beta(t)}{\sum_{i=1}^n q_i x_i^\beta(t)}. \quad (1)$$

We can see from (1) that the parameter β has a significant impact on customers' choice behavior and measures to what extent customer choice behavior is affected by the network sizes. In this paper, we use the parameter β to measure the magnitude of the network effect, with a higher value of β corresponding to a stronger network effect. Note that β is allowed to take values from 0 to ∞ . Later in the paper, we show that when β takes different values, the market dynamics may behave very differently (see Section 4.1).

If the customer selects product i at time $t+1$, then the vector of sales volumes for each product becomes

$$d_j(t+1) = \begin{cases} d_j(t) + 1 & \text{if } j = i, \\ d_j(t) & \text{otherwise,} \end{cases}$$

and the vector of market shares is correspondingly updated as

$$\mathbf{x}(t+1) - \mathbf{x}(t) = \frac{1}{t+1+n} [\boldsymbol{\phi}(\mathbf{x}(t)) - \mathbf{x}(t) + \mathbf{e}(t+1) - \mathbb{E}[\mathbf{e}(t+1)|\mathbf{x}(t)]], \quad (2)$$

where $\boldsymbol{\phi}(\mathbf{x}(t)) = \{\phi_1(\mathbf{x}(t)), \dots, \phi_n(\mathbf{x}(t))\}$ and $\mathbf{e}(t+1)$ is the random unit vector whose i th entry is 1 if product i is purchased at time $t+1$ and 0 otherwise. In this paper, the total market size is normalized to be 1 and hence firm i 's market share in the limit as $t \rightarrow \infty$ is equivalent to its demand.

To conclude this section, we summarize the sequence of events as follows (see also Figure 1):

Stage 1 (Entry Decision). Each firm decides whether to enter the market. A firm would exit the market if the expected profit is negative. The product variety is the number of firms/products entering the market.

Stage 2 (Quality Competition). Suppose there are n firms entering the market. Given n is commonly known, all firms simultaneously choose their quality levels, so as to maximize their own expected profit in the competitive selling stage.

Stage 3 (Demand Process with Network Effect). The customers sequentially arrive to select one product maximizing her utility at the time of arrival. Each customer arrival is able to observe all products' quality levels and their current market shares.

4. Equilibrium Analysis

In this section, we study the subgame-perfect equilibrium of the three-stage game, using the backward induction approach. We begin our analysis by the third stage in which the price p , all firms' quality $\mathbf{q}$, and the network effect parameter β are given and known. It is noted that $\mathbf{d}(t)$ is what is referred to as *nonlinear Pólya urn process* and $\mathbf{x}(t)$ is a corresponding discrete stochastic process. By Nevel'son and Khas' minskii (1973, Theorem 7.3, Chapter 2), $\mathbf{x}(t)$ almost surely converges to

Figure 1 Sequence of Events

the set of fixed points for $\phi(\mathbf{x})=\mathbf{x}$, as $t \rightarrow \infty$. In this paper, the vector of market shares in the limit $\lim_{t \rightarrow \infty} \mathbf{x}(t)$, which may be a point or a random variable, is defined as the demand vector for the n firms under all firms' joint quality level vector $\mathbf{q}$, and the strength of the network effect β . Then we study the equilibrium at the second stage in which all incumbent firms simultaneously choose their quality levels based on their anticipated demand functions which will be formed according to the third consumption stage. Finally, we study the market entry stage, which determines how many products would be available in the market. Our objective is to obtain qualitative insights on how the network effect β affects the sales distribution within an industry and hence affects firms' entry decisions that shape the product variety of the industry.

4.1. Demand Concentration under Network Effect

At the consumption stage, customers whose individual purchase decision depends on the aggregate sales of her predecessors but is independent of the future sales arrive sequentially at the market one per period. In this subsection, we study the limit of market shares dynamics under network effect given that there are n firms in the market and their quality levels are $\mathbf{q}=(q_1, \dots, q_n)$.

Most papers studying the *urn processes* (e.g., [Zhu 2009](#)) often discuss the limiting behavior of the stochastic process (2) by three cases: $0 < \beta < 1$, $\beta = 1$, $\beta > 1$, because the process converges to different points or random variables in different cases as t goes to infinity. [Hu et al. \(2016\)](#) is the first to study the effects of the three cases of network effect on a newsvendor problem. We also study such three cases, but endow heterogeneous attributes (i.e., quality level) to each product, which distinctively differs from [Hu et al. \(2016\)](#) and the large body of literature on urn processes.

Given the asymmetric product quality $\mathbf{q}=(q_1, \dots, q_n)$ and the symmetric initial sales volumes $\mathbf{d}(0)=(1, \dots, 1)$, the first customer makes her purchase decision only depending on her intrinsic utility for each product. But from the second customer, each customer's choice decision not only depends on her intrinsic utility for each product but also depends on the (random) social utility

from the (random) sales volumes. Thus, the asymptotic distribution of the vector of market shares must be a function of the vector of intrinsic quality $\mathbf{q}$ and the strength of network effect β . Denote by $\mathbf{x}^*(\mathbf{q}, \beta)$ the limit of consumption dynamics as $t \rightarrow \infty$. We show that the limit $\mathbf{x}^*(\mathbf{q}, \beta)$ must be unique and we provide a closed-form expression for it in the following theorem, where we assume $n \geq 2$ since the case $n = 1$ is trivial.

THEOREM 1 (LIMIT OF CONSUMPTION DYNAMICS). *Given the product quality $\mathbf{q} = (q_1, \dots, q_n) \neq \mathbf{0}$ and the initial sales volumes $\mathbf{d}(0) = (1, \dots, 1)$, the consumption dynamics $\mathbf{x}(t)$ converges to $\mathbf{x}^*(\mathbf{q}, \beta)$ almost surely and $\mathbf{x}^*(\mathbf{q}, \beta)$ can be characterized by the following three cases.*

- (i) *If $0 \leq \beta < 1$, then $\mathbf{x}^*(\mathbf{q}, \beta) = (q_1^{\frac{1}{1-\beta}}, \dots, q_n^{\frac{1}{1-\beta}}) / \sum_{i=1}^n q_i^{\frac{1}{1-\beta}}$.*
- (ii) *If $\beta = 1$, then $\mathbf{x}^*(\mathbf{q}, \beta)$ is uniformly distributed on the set $\{\mathbf{x} \in [0, 1]^n : \sum_{i \in L} x_i = 1, \sum_{i \notin L} x_i = 0\}$, where $L = \{i : q_i \text{ is the largest entry of } \mathbf{q}\}$.*
- (iii) *If $\beta > 1$, then $\mathbf{x}^*(\mathbf{q}, \beta)$ follows a multi-point distribution with support on $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ where $\mathbf{e}_i$ is the unit vector with i th entry being 1 and 0 otherwise, and*

$$\mathbb{P}\{\mathbf{x}^*(\mathbf{q}, \beta) = \mathbf{e}_i\} = \lim_{N \rightarrow \infty} \sum_{j_1=1}^N \cdots \sum_{j_n=1}^N \frac{q_i j_i^\beta}{q_1 j_1^\beta + \cdots + q_n j_n^\beta} \prod_{k=1}^n H_{j_k, N}, \quad (3)$$

$$\text{where } H_{j, N} \triangleq \prod_{\substack{l=1 \\ l \neq j}}^N \frac{l^\beta}{l^\beta - j^\beta}, \text{ for } j = 1, \dots, N.$$

We categorize network effects by the parameter β into three scenarios: $0 \leq \beta < 1$, $\beta = 1$, $\beta > 1$ (see also [Hu et al. 2016](#)). Theorem 1 shows that for a different network effect magnitude β , both the long-term and the short-term behavior of the market share dynamics exhibits dramatically different properties.

Theorem 1(i) says that if $0 \leq \beta < 1$, the stochastic process $\mathbf{x}(t)$ has a *unique* limit regardless of what the n products' quality are and what the aggregate demand distribution in any finite period is. The limit of market evolution solely depends on the quality levels of the n products. An unpopularity of a product over the short term can always win back its market share as long as it has a higher quality. In this case, the market evolution has a predictable outcome at the end. The higher the quality of a product, the more market share it will gain eventually.

Theorem 1(ii) says that when β approaches to 1 from the left, the product with the highest quality will dominate all other products, whose market share approximates to 1 in the end. If two or more products achieve the same highest quality, they divide the market randomly and the rest firms obtain zero market share. This scenario manifests the importance of the quality for a product in the sense that the market evolution eventually goes to a “winner-takes-all” pattern in which those products with the largest quality get “all” but others get “nothing”.

Theorem 1(iii) says that if $\beta > 1$, the market is eventually dominated by one *unique* product and other products all obtain zero fraction of the market. That is because, if one of the products has a sufficiently large market share at some point of time (which is referred to as the *attraction time*) then all subsequent customers purchase this product from that time onward. In the appendix, Theorem A.4 shows that such an attraction time is finite. In this scenario, i.e., $\beta > 1$, if a product fails to gain traction for a while, it becomes impossible to win back in the end. Interestingly, anyone can be the winner regardless of its quality, and there is one and only one winner, in contrast with scenarios (i) and (ii). A product even with the lowest quality may be purchased by patrons enough early on and becomes the sole winner. The strong network effect reinforces this advantage sufficiently such that all other products will never win back. A similar phenomenon is observed in other contexts with different models. For example, Fleder and Hosanagar (2009) make similar observations by studying two products' market share dynamics via a random walk framework. Different from Fleder and Hosanagar (2009), our model can be used to study the case with an arbitrary number of products.

The three scenarios have one thing in common. That is, an increasing network effect always contributes to the demand concentration on a small number of products. The difference, however, is: the increasing network effect would shift more demand towards the products with higher qualities when the prevailing network effect is generally weak, e.g., $\beta \leq 1$; but it would make the market more easily concentrated to one product that can be a low-quality one when the prevailing network effect is strong, e.g., $\beta > 1$. To further highlight this difference, we refer to the case $\beta \leq 1$ as the case with *dominated network effect* in which the demand concentration is predictable by solely observing initial qualities of all products, and the case $\beta > 1$ as the case with *dominant network effect* in which demand concentration is unpredictable due to the network effect. This theoretical finding lends some support to the practice in the movie industry. Nowadays, major movie studios tend to aim at investing in high-budget movies and younger movie making entities are more likely to pursue low-budget movies. One can see from the outcomes that a high-budget product has a higher revenue than a niche on average but some high-budget products can be flops in the box office and some low-budget movies may become surprising dark horses, though the latter has lower chances (Natividad 2013). We will empirically show in Section 6 that “ $\beta \leq 1$ ” and “ $\beta > 1$ ” co-exist in the movie industry but separately within different market segments.

Theorem 1(iii) presents the probabilities that each firm becomes the monopoly. The following lemma further shows a property of those probabilities and is used in the subsequent analysis.

LEMMA 1. For $\beta > 2$, Equation (3) can be rewritten as

$$\mathbb{P}\{\mathbf{x}^*(\mathbf{q}, \beta) = \mathbf{e}_i\} = \sum_{j_1=1}^{\infty} \cdots \sum_{j_n=1}^{\infty} \frac{q_i j_i^\beta}{q_1 j_1^\beta + \cdots + q_n j_n^\beta} \prod_{k=1}^n H_{j_k},$$

where $H_j \triangleq \prod_{\substack{l=1 \\ l \neq j}}^{\infty} \frac{l^\beta}{l^\beta - j^\beta}$, for $j \in \mathbb{N}$.

As a complement to Theorem 1 and Lemma 1, the following proposition takes a close look at the market shares in the symmetric equilibrium. The terms, “increasing” and “decreasing”, are in its weak sense in the rest of the paper.

PROPOSITION 1. For $\beta \geq 0$ and $i = 1, \dots, n$, $\mathbb{E}[x_i^*(\mathbf{q}, \beta)]$ is increasing in q_i and decreasing in q_j , $\forall j \neq i$.

Proposition 1 says that the expected market share for each firm is increasing in its own quality but decreasing in any competitor’s quality, regardless of whether the magnitude of network effect is strong or weak. This means that the market always endows the firm with a higher quality with a larger market share in the end on expectation. For $\beta \leq 1$, this insight can be clearly observed from Theorem 1 parts (i) and (ii). For $\beta > 1$, Proposition 1 demonstrates that the firm with a higher quality has a larger probability of becoming the sole dominant player in the market.

In the competitive market, each firm aims at growing its market share as much as possible. For this purpose, each firm has to improve the quality of its product to attract as many customers as possible and gain the network effect for its product at the early selling stage. Each firm enters the competitive market as long as its expected revenue exceeds its cost. A natural question would be: how does an increase in the network effect magnitude have an impact on each firm’s incentive to improve its quality, given the number of firms is fixed? From Theorem 1 parts (i) and (ii) and Proposition 1, it can be seen that if $0 \leq \beta \leq 1$, each firm would like to enhance the quality of its product as the magnitude of network effect β increases. Hence, the stronger network effect leads to a more competitive market. Then what is the answer for $\beta > 1$? We will rigorously answer this question in the next section.

4.2. Quality Competition

If the expected demand of a product, as a function of its quality, is known, a rational firm will naturally choose its quality level to make a trade-off between the corresponding expected revenue and the production cost, so to maximize its expected profit. In Section 4.1, we characterize the demand functions for the n firms by $\mathbf{x}^*(\mathbf{q}, \beta)$ under three scenarios. Based on these results, in this

subsection, we study the n firms' competition on quality levels and model it as a simultaneous-move game in which the strategy space for the n firms is $E \triangleq [0, +\infty)^n$. We denote by $\mathcal{G} = \{\{1, \dots, n\}, \{\pi_1, \dots, \pi_n\}, E\}$ the second-stage game with n firms. Since all firms are assumed to be identical in this paper, it makes sense to focus on the symmetric equilibrium of the game $\mathcal{G}$.

For $i = 1, \dots, n$, given other products' qualities $\mathbf{q}_{-i} = (q_1, \dots, q_{i-1}, q_{i+1}, \dots, q_n) \in E$ and the network effect parameter β , firm i 's problem is to choose a quality level q_i to maximize:

$$\pi_i(q_i, \mathbf{q}_{-i}, \beta) = p\mathbb{E}[x_i^*(q_i, \mathbf{q}_{-i}, \beta)] - cq_i, \quad (4)$$

where $\pi_i(\cdot)$ is firm i 's expected profit at the second stage, independent of the entry cost K . In a Nash Equilibrium (NE), the quality for any firm is the best response to others'. Let $q^*(n, \beta)$ and $\pi^*(n, \beta)$ denote each firm's quality and expected profit respectively in a symmetric equilibrium of $\mathcal{G}$. Recall that at the consumption stage the market evolution behaves differently as the network effect parameter β varies. In this subsection, we study the quality competition for any given β and characterize the equilibrium by two cases: $0 \leq \beta \leq 1$ and $\beta > 1$. The differences between equilibria under different cases reflect the impact of the magnitude of network effect on the quality competition among n firms. The following theorems summarize the equilibrium results of the game $\mathcal{G}$ by the two cases (for which again we ignore the trivial case $n = 1$).

THEOREM 2 (DOMINATED NETWORK EFFECT). *For $0 \leq \beta \leq 1$, define $m(\beta, K) = \lfloor \frac{p}{K+(p-K)\beta} \rfloor$ and the equilibrium of $\mathcal{G}$ is characterized as follows:*

- (i) *If $2 \leq n \leq m(\beta, 0)$, then there exists a unique symmetric equilibrium where $q^*(n, \beta) = \frac{(n-1)p}{n^2c(1-\beta)}$ and furthermore, for any $K \geq 0$, $\pi^*(n, \beta) \geq K$ if and only if $n \leq m(\beta, K)$.*
- (ii) *If $n > m(\beta, 0)$, then there is no symmetric pure-strategy equilibrium. In any symmetric mixed-strategy equilibrium, $\pi^*(n, \beta) = 0$.*

THEOREM 3 (DOMINANT NETWORK EFFECT). *For $\beta > 1$, define*

$$m(\beta, K) = \max \left\{ n : \lim_{N \rightarrow \infty} \sum_{j_1=1}^N \cdots \sum_{j_n=1}^N \frac{(j_1^\beta)^2}{(j_1^\beta + \cdots + j_n^\beta)^2} \prod_{k=1}^n H_{j_k, N} \geq \frac{K}{p} \right\}. \quad (5)$$

- (i) *If $2 \leq n \leq m(\beta, 0)$, then there exists a unique symmetric equilibrium where*

$$q^*(n, \beta) = \frac{p}{c} \left(\frac{1}{n} - \lim_{N \rightarrow \infty} \sum_{j_1=1}^N \cdots \sum_{j_n=1}^N \frac{(j_1^\beta)^2}{(j_1^\beta + \cdots + j_n^\beta)^2} \prod_{k=1}^n H_{j_k, N} \right), \quad (6)$$

and furthermore, for any $K \geq 0$, $\pi^(n, \beta) \geq K$ if and only if $n \leq m(\beta, K)$.*

- (ii) *If $n > m(\beta, 0)$, then there is no symmetric pure-strategy equilibrium. In any symmetric mixed-strategy equilibrium, $\pi^*(n, \beta) = 0$.*

In Theorems 2 and 3, the term $m(\beta, K)$, a function of the network effect parameter β and the entry cost K , can be considered as the *market capacity* for product variety, namely, the largest number of competitive products such that each product would earn a profit exceeding K in equilibrium. A market with a larger market capacity accommodates a larger variety of products. Once the number of competitive products entering the market is beyond its capacity, the competition among them is too intense such that no firm profits from further entering the market. Here, we introduce the notion of $m(\beta, K)$ to facilitate our elaboration on the equilibrium of the quality competition. More properties about $m(\beta, K)$ will be discussed in Section 4.3.

As a complement to Theorem 3, the following lemma shows alternative formulations for the market capacity and the equilibrium quality, and will be used in the subsequent analysis.

LEMMA 2. *For $\beta > 2$, (5) and (6) can be rewritten as*

$$m(\beta, K) = \max \left\{ n : \sum_{j_1=1}^{\infty} \cdots \sum_{j_n=1}^{\infty} \frac{(j_1^\beta)^2}{(j_1^\beta + \cdots + j_n^\beta)^2} \prod_{k=1}^n H_{j_k} \geq \frac{K}{p} \right\},$$

$$q^*(n, \beta) = \frac{p}{c} \left(\frac{1}{n} - \sum_{j_1=1}^{\infty} \cdots \sum_{j_n=1}^{\infty} \frac{(j_1^\beta)^2}{(j_1^\beta + \cdots + j_n^\beta)^2} \prod_{k=1}^n H_{j_k} \right).$$

To further highlight the impacts of the network effect and the number of products on the quality competition, we present the following proposition.

PROPOSITION 2. *The equilibrium profit and quality satisfy the following properties:*

- (i) *For $\beta \geq 0$ and $n < m(\beta, 0)$, $\pi^*(n, \beta)$ and $q^*(n, \beta)$ are both decreasing in n , that is, $\pi^*(n, \beta) \geq \pi^*(n+1, \beta)$ and $q^*(n, \beta) \geq q^*(n+1, \beta)$.*
- (ii) *(Dominated network effect) For $0 \leq \beta < 1$ and $n \leq m(\beta, 0)$, $q^*(n, \beta)$ is increasing in β .*
- (iii) *(Dominant network effect) There exists a $\tilde{\beta} \geq 1$ such that, for $\beta > \tilde{\beta}$ and $n \leq m(\beta, 0)$, $q^*(n, \beta)$ is decreasing in β .*

Proposition 2(i) shows that the expected equilibrium profit for each firm is decreasing in the number of competitive products no matter whether the network effect is strong or weak. The more products are in the market, the more intense the competition is and hence the lower expected profit each firm can obtain, as we assume the total market size is fixed. This result is fairly intuitive in traditional economic theory. We verify that this result holds even with the network effect. Proposition 2(i) also shows that the equilibrium quality level is decreasing in the number of products in the market. That is because each firm has to reduce its cost by lowering its quality as more firms enter the competition.

Proposition 2 parts (ii) and (iii) illustrate the impact of the network effect on the quality competition. More specifically, Proposition 2(ii) shows that the equilibrium quality increases, which means the quality competition is intensified, as the network effect parameter increases when the prevailing network effect is *dominated*. It is interesting to contrast this part to Proposition 2(iii) which presents the case when the prevailing network effect is *dominant*. There the equilibrium quality decreases, which means the quality competition is alleviated, as the network effect parameter increases. Note that the insight in Proposition 2(iii) holds if the prevailing network is greater than a threshold $\tilde{\beta}$. We will numerically show in Section 5 that this threshold is in fact 1.

From Proposition 2 parts (ii) and (iii), we can explore the impact of the increasing network effect on the quality investment competition. In this paper, investment strategies can be generally categorized to two types: blockbuster and niche strategy. A blockbuster strategy has the feature of releasing high-quality products targeted at a large group of consumers. In contrast, a niche strategy aims at commercializing low investment products expected to appeal to a smaller group of customers. On one hand, we find from Proposition 2(ii) that when the prevailing network effect is dominated each firm has a high incentive to improve its product's quality since the higher quality would attract more demand to it. This induces the equilibrium investment strategies to shift towards blockbuster strategies as the network effect increases. On the other hand, we find from Proposition 2(iii) that when the prevailing network effect is dominant each firm turns to be confident with the network benefit, even if its product is targeted to a small group of consumers, since a small advantage in sales could be significantly amplified as long as such an advantage occurs in the early of the horizon. In this case, it is reasonable to believe even if more than one product is allowed to be produced, each firm would allocate its limited resources to a broader range of low-quality products, i.e., shifting towards the niche strategy, as the network effect parameter increases.

So far, we have revealed the impact of the increasing network effect on the equilibrium quality investment strategy. As mentioned above, Proposition 2 has an implication for quality investment strategies in an industry with increasing network effect, e.g., music, movie, books online industry, etc, due to the growing influence of social media, review websites and self-enforcing recommendation systems. In some other industries, however, the network effect may not be simply increasing but could evolve non-monotonically over time. Our result may also shed some light for those industries depending on where their current network effect magnitude lies. In fact, we can see that the equilibrium strategy shifts towards the blockbuster strategy when the network effect parameter moves close to 1 from either side of 1, but the equilibrium strategy shifts towards the niche strategy when the network effect parameter moves away from 1 to either side of 1.

4.3. Product Variety

In the first stage of market entry competition, any firm who wishes to enter the market must incur an entry cost K . Firms simultaneously decide whether to enter the market by trading off the entry cost and their expected profit from the second and third stages. If n firms decide to enter the market, each one's expected profit in the first stage is given by

$$\Pi^*(n, \beta) = \pi^*(n, \beta) - K.$$

Hence n firms will join the competition if and only if $\Pi^*(n, \beta) > 0$. The product variety in the market is defined by the number of products entering the market, as an outcome of firms' ex-ante entry decisions. We consider only the pure strategy equilibrium.

In Section 4.2, we have denoted by $m(\beta, K)$ the market capacity, namely the largest number of products entering the market such that all products entering the market can earn a profit of at least K . We have shown that in Theorems 2 and 3 that the n firms are willing to enter the market or indifferent to whether he enters or not if $n \leq m(\beta, K)$. In this section, we view $m(\beta, K)$ as a measure of the product variety in the market, given the magnitude of network effect β and the entry cost K . The following theorem presents some properties of $m(\beta, K)$.

THEOREM 4 (PRODUCT VARIETY). *For $K > 0$, the market capacity $m(\beta, K)$ satisfies the following properties:*

- (i) $m(\beta, K) \geq 1$ if and only if $p \geq K$; for $\beta \geq 0$, $m(\beta, K)$ is decreasing in K .
- (ii) (Dominated Network Effect) $m(\beta, K)$ is decreasing in β on $[0, 1)$.
- (iii) (Dominant Network Effect) There exists a $\tilde{\beta} \geq 1$ such that $m(\beta, K)$ is increasing in β on $(\tilde{\beta}, +\infty)$, where $\tilde{\beta}$ is the same as the threshold in Proposition 2.

Part (i) of this theorem is intuitive. Note that p is not only the price but also the revenue for the whole industry as we normalize the total market size to be 1. The higher the entry cost, the less the product variety in equilibrium in the industry. When the entry cost is higher than p , no firm would enter the market. A low entry cost is frequently quoted as a reason to support the long tail phenomenon. Our framework shows that the entry cost indeed plays a significant role in an industry with network effect.

Theorem 4 parts (ii) and (iii) characterize the impact of the increasing network effect on the product variety. The product variety in an industry may expand or shrink as the network effect increases. With the dominated network effect, an increasing network effect would prevent more products from entering the market ex ante, since the firms would expect a more intense competition

as the network effect becomes stronger (see Proposition 2(ii)). In contrast, with the dominant network effect, the ex-ante equilibrium product variety would be broader if the network effect becomes stronger, since the firms would expect a less competitive market (see Proposition 2(iii)).

In Section 4.1, we have shown that an increasing network effect always contributes to the demand concentration on a small number of products, e.g., ex-post blockbusters. As a complement, Theorem 4 parts (ii) and (iii) help to exhibit how the tail in the curve of sales rank distribution varies as the network effect parameter increases. More specifically, when the network effect is dominated, an increasing network effect contributes to a shrinking tail in the curve of sales rank distribution; interestingly, when the network effect is dominant, an increasing network effect contributes to an expanding tail in which most of products earn sales close to 0 ex post. In Section 5, we will demonstrate these results graphically through a numerical experiment.

In Section 4.2, we have studied the quality competition with network effect among multiple firms and shown that an increasing network effect can intensify or alleviate the competition. The following proposition continues to study the quality competition with network effect, but, differently, it incorporates the endogenized product variety.

PROPOSITION 3. *For $K > 0$ and $\beta \geq 0$, suppose there are $m(\beta, K)$ firms entering the market. Let $q^*(\beta)$ be the quality that they offer in the symmetric equilibrium at the second stage. Then $q^*(\beta)$ is increasing in β on $[0, 1)$ and decreasing on $(\tilde{\beta}, +\infty)$, where $\tilde{\beta}$ is the same as the threshold in Proposition 2.*

We find from Proposition 3 that the main insights in Section 4.2 still hold when the firms' entry decisions are endogenized. The equilibrium quality increases, i.e., the quality investment strategy shifts towards blockbuster strategy, as the network parameter increases if the prevailing network effect is dominated, but it turns out to decrease, i.e., the equilibrium quality strategy shifts towards niche strategy, as the network parameter increases if the prevailing network effect is dominant.

5. Numerical Study

We conduct extensive numerical experiments for our model, since some questions are difficult to be answered analytically. In particular, our experimental design centers around two main questions: (a) What is equilibrium outcome for the case when $1 < \beta < \tilde{\beta}$ and (b) what will be the equilibrium if some part of the assumptions are changed, e.g., when firms are heterogeneous with respect to the marginal cost of improving quality, and when the cost structure becomes more general than being linear in the quality level?

5.1. $\tilde{\beta}$ is exactly 1

In Proposition 2, we defined a parameter $\tilde{\beta}$ and analytically showed its existence but without providing its precise value. The term $\tilde{\beta}$ stands for a threshold on the magnitude of the network parameter above which the network effect is *dominant* and below which the network effect is *dominated*. It is critical for part of our results, e.g., Propositions 2, 3 and Theorem 4. In this subsection, we will numerically show the parameter $\tilde{\beta}$ is exactly 1.

The term $\tilde{\beta}$ is first adopted in Proposition 2(iii) which characterizes the equilibrium quality, i.e., $q^*(n, \beta)$, at the second stage game $\mathcal{G}$ when the network effect parameter $\beta > 1$. To precisely estimate $\tilde{\beta}$, it suffices to show how the equilibrium quality $q^*(n, \beta)$ varies as the network parameter β changes, given the number of firms n is fixed. Recall that we have shown the market capacity $m(\beta, 0)$ depends on β . That is, for a given n , β must be sufficiently large to guarantee $n \leq m(\beta, 0)$. Thus we define

$$\beta_{min}(n) = \inf_{\beta > 1} \{\beta : n \leq m(\beta, 0)\}.$$

In Table 1, we compute values of $\beta_{min}(n)$ for certain number of firms. In line with Theorem 4(iii), Table 1 to some extent confirms the insight that the market capacity is increasing in β in the case with dominant network effect.

Table 1 $m(\beta, 0)$ and β when $\beta > 1$

n	2	3	10	50	100	150
$\beta_{min}(n)$	1.91	2.87	5.02	6.14	6.34	6.42

It is noteworthy that $q^*(n, \beta)$, defined by Equation (6), is difficult to compute directly by Equation (6). We will approximate such a series in Equation (6) by an integral using a technique which can be found in the proof Theorem A.5. The numerical results are summarized in Figure 2.

As Figure 2 displays, $q^*(n, \beta)$ is decreasing in β and n respectively. Moreover, this monotonicity holds as long as $\beta \geq \beta_{min}(n)$, i.e., $n \leq m(\beta, 0)$. Based on these numerical results, we tend to believe $\tilde{\beta}$ in our analytical framework should 1, though we are not able to analytically demonstrate this.

5.2. Heterogeneous Firms and More General Cost Function

Different from the base model which focus on the symmetric equilibrium for analytical tractability, this subsection develops the asymmetric equilibrium for the second stage game $\mathcal{G}$ in the setting in which the marginal cost of improving the quality among firms is heterogeneous. In practice, different business entities may have different technology levels in the development and production, and as a result, their quality levels are different in a realistic quality competition. This difference would

Figure 2 The Equilibrium Quality at the Second Stage Game

Note. The parameter values are $p = 1$, $c = 0.5$.

play a fundamental role in determining the shape of ex-post sales rank curve. In this subsection, we will numerically show how the network effect parameter affects the shape of the ex-post sales rank curve, given a number of firms with different technology levels.

Figure 3 Qualities in Asymmetric Equilibria

Note. The numbers located at the x axis are the labels for products. For example, “2” stands for product 2. The label with no bar on it means the corresponding product does not enter the market.

Another focus is on the cost structure. The base model assumes a linear cost structure in improving the quality, i.e., cq . However, a general form of convex cost structure, i.e., cq^γ , $\gamma > 1$, may be more common in most of the industries. Hence, we adopt this kind of convex cost structure in the numerical experiments. In the base model, we characterize the quality competition by two cases: $\beta \leq 1$ and $\beta > 1$. Thus $\beta = 1$ is a cutoff between the two cases. On different sides of this cutoff, the

Figure 4 Ranked ex post Sales Volumes

Note. The sales volumes for the bottom 10 products when $\beta = 6$ are: 1749, 1373, 1363, 1348, 1312, 1244, 1235, 1234, 1203, 1182.

equilibrium results in the quality competition demonstrate different properties (see Section 4.2 for more details). Through the numerical study, we seek to identify if the cutoff varies when the cost structure changes.

The parameters $\mathbf{c}$ and γ , capturing the cost structure in an industry, critically influence the quality levels of products offered in the market and the distribution of their sales. A large body of literature on the blockbuster vs. long tail phenomena has studied this aspect, see, e.g., Bar-Isaac et al. (2012) and Brynjolfsson et al. (2006). Therefore, we incorporate both the heterogeneity of $\mathbf{c}$ and a convex cost function in the numerical experiments. To obtain the equilibrium of the quality competition, we resort to solving for the fixed point of the First Order Conditions (FOCs) derived from each firm's profit functions (see the proofs of Theorems 2 and 3). Then, given the products offered and their qualities in equilibrium, we compute all firms' ex post sales volumes by simulating the demand process through a discrete choice model defined by (1). In the following numerical studies, we consider a market where there are 20 potential firms (labeled from 1 to 20) with $\mathbf{c}$ evenly located from $c_1 = 0.04$ to $c_{20} = 0.116$. That is, c_1 is the lowest entry. Each firm entering the market has an initial sales volume of 1000. There are 100,000 customers sequentially arrive one at a time. Other parameters are $p = 1$, $\gamma = 2$, $K = 10^{-5}$. The numerical results are summarized in Figures 3 and 4.

The first observation from Figure 3 is that, as β increases from 0.1 to 0.4, product variety shrinks from 7 to 2; but, as β increases from 2 to 6, product variety expands from 2 to 11. That is, the equilibrium product variety has a U-shape relationship with the the network effect magnitude β . This observation is consistent with our analytical results. In this numerical example, the market

capacity would be reduced to 1 if $0.49 < \beta < 1.98$. Hence, $\beta = 1$ can still be viewed of as a cutoff on the network effect parameter above which the network effect is dominant and below which the network effect is dominated. Secondly, we see from Figure 3 that the quality competition is intensified as β increases from 0.1 to 0.4, which drives the product variety to shrink; but the quality competition becomes alleviated as β increases from 2 to 6, which contributes to the expansion of product variety. Naturally, higher-technology firms, namely those firms with lower marginal costs for improving the quality, would offer higher qualities in the (asymmetric) equilibrium, regardless of the network parameter. The lower-technology firms are gradually forced out of the market as β increases but is still below 1. Interestingly, when the network effect is sufficiently large, the lower-technology firms have a chance again to enter the market.

Next, we consider the ex-post sales rank distribution. Figure 4 plots the distribution of ex-post sales for 4 different network effect parameters. Obviously, more demand is concentrated on a smaller number of products as β increases from 0.1 to 6. Comparing sales at different values of β , the demand shifts towards the high-quality products from low-quality products as β increases but is less than 1; but this trend no longer exists when $\beta > 1$. As Figure 3(b) suggests, the firm with the highest quality has only a few sales volume, and many lower-quality products attract more demand than higher-quality products, especially in the tail of the ranked sales curve. This insight is exactly implied by Theorem 1.

6. An Application to Movie Industry

In this section, we consider the application of our general framework to the study of films release. This application aims at fitting with the practical data to estimate the specific network effect in the movie industry. The network effects of a film mainly come from two sources. The first is the social utility gained by individuals from conforming to others who also watched the same film. The second is the availability of the film. A film that is successful in the opening week will often be available at more theaters in the subsequent weeks and often have a long lifespan in theater. We use the parameter β in our model to measure the network effect in the movie industry. The value of β can be estimated with the approach of Maximum Likelihood Estimation (MLE) by fitting our analytical model to weekly sales data in the movie industry. The values of β may be different when different data segments are adopted. We then use the theory developed in the previous sections to study: (a) the blockbuster or long tail phenomena, (b) the quality (budget) competition and (c) the product variety in the movie industry.

Our data is a comprehensive dataset that consists of the weekly sales and other characteristics of US feature movie releases from 1980 to 2017, collected from multiple online sources, including

IMDb, Box Office Mojo, etc. The dataset includes 9,294 US movies in total. We focus on 6,715 US movies since each of them has an estimated budget, a box office revenue in US and at least two values of sales for the first two weeks, and others are excluded because of missing values. In this section, all sales means the sales volumes in US. We divide these movies into a set of groups according the release date. In each group, there are at least two movies and all movies must be released at the same week which is defined by boxofficemojo.com as beginning from a Friday to next Thursday. We do this grouping to guarantee that movies in one group have almost the same initial conditions on sales. Each group can be viewed as a “local market” within which movies compete with each other. We obtain 1,117 groups in total from the data. We assume the probability distribution of a customer’s choice among movies in a group is independent of the probability distribution of this customer’s choices among other groups. To apply our theoretical model, we define by b_i the budget of movie i and view b_i^α as the intrinsic quality of movie i . The parameter α will be estimated later through MLE as well.

A customer’s choice behavior among a group of n movies is modeled as follows. For $i = 1, \dots, n$, the probability that a customer selects movie i in the first two weeks¹ since its release is $b_i^\alpha / \sum_{i=1}^n b_i^\alpha$, and the probability that a customer selects movie i after the first two weeks is

$$\frac{b_i^\alpha \cdot (\text{the first two weeks' sales of movie } i)^\beta}{\sum_{i=1}^n b_i^\alpha \cdot (\text{the first two weeks' sales of movie } i)^\beta}.$$

This choice model is built based on our base model (see Equation (1)). In the base model, we consider an infinite-period dynamics in which one customer arrives in each period. In contrast, we simplify the base model into a two-period model in which the first two weeks can be seen as the first period and other screen-weeks can be seen as the second period. We make such a simplification because (a) movies played for an average of 13 weeks, and a median of 7 weeks, which are estimated from the data in 2016 in US and 96 percent of screen capacity was used by the top 200 movies in that year; hence most movies have a short screen-time that it is impossible to apply an infinite-period model; (b) people as well as media often pay most attention to the initial sales of a movie in a few days at its beginning and its ultimate box office revenue, but may not care much about the sales in the duration of the screen time; (c) the data of the sales in a middle week are missing for a part of the movies. For tractability, let $Initial_i$ be the sales of movie i in the first two weeks

¹ Here, we use the first two weeks’ sales, instead of the first week’s sales, as the initial condition. Note that the value of the first week’s sales we observed from boxofficemojo.com may be not the value of a full week’s sales, namely the first seven days’ sales, because a week is defined by boxofficemojo.com as beginning from Friday to Thursday but not all movies in a group released on Friday. Hence, a large amount of errors may be introduced from using the first week’s sales. The possible error is reduced if we use the first two weeks’ sales.

since its release and $Total_i$ be the total gross sales of movie i in US. Using this choice model, the logarithm of the likelihood function for the observed data of a group of movies can be written as

$$l(\alpha, \beta, G) = \sum_{g \in G} \sum_{i \in g} \left\{ Initial_i \cdot \ln \left(\frac{b_i^\alpha}{b_1^\alpha + \dots + b_n^\alpha} \right) + (Total_i - Initial_i) \cdot \ln \left(\frac{b_i^\alpha Initial_i^\beta}{b_1^\alpha Initial_1^\beta + \dots + b_n^\alpha Initial_n^\beta} \right) \right\},$$

where g is a group defined above and G is a set of groups that satisfy specific requirements that we will outline below.

We first separately estimate α and β through MLE for each group. The results are summarized in Table 2. As the table displays, the estimated parameters α and β vary among different groups. About 82 percent groups generate a $\alpha > 0$, which means in most cases a higher-budget movies has a higher intrinsic appeal to the market. About 24 percent groups generate a $\beta > 1$. That is, the dominated network effect and dominant network effect coexist in the movie industry and both contribute to the market structure of the industry.

According to our theory, as the network effect parameter increases, under a dominated network effect, the demand tends to be concentrated on a smaller number of high-budget movies, whereas under a dominant network effect, the demand can be easily concentrated on a smaller number of movies that may have a low budget. These results are consistent with the increase of the blockbusters in the movie industry. Our theory also suggests that the budget competition among studios would be heating up as the network parameter increases when the prevailing network effect is dominated, but would be waning when the prevailing network effect is dominant. Moreover, as our theory suggests, the product variety in the movie industry may shrink or may expand over time, depending on the intensity of the budget competition. Our theory explains why there is a substantial increase in movie production, i.e., shifting towards blockbuster strategy, over time, but, meanwhile, there are more producers targeting niche movies, especially as digitization unfolds, which helps to increase the network effect. Since US theaters exhibit only about 500 movies per year, it is hard to see how the variety of movies on theaters varies over time. However, it seems likely that most niches are distributed digitally through online platforms or channels, such as Amazon, Netflix and iTunes. This phenomenon is consistent with the evidence of [Goldmanis et al. \(2010\)](#) on bookstores, new-auto dealers and travel agencies.

Another question may be what kind of group generates a large or small value of α and β . This can be an intricate question. The following analysis provides a sketch map. We identify each group by the maximum budget movie—the most likely blockbuster in that group. Then we study how

Table 2 Descriptive statistics for the α and β estimated by the data of each group

	Mean	St. Dev.	Probability of <0	Probability of >1
α	1.4157	7.1835	0.1772	0.4494
β	0.4518	4.0503	0.3536	0.2435

the maximum budget affects the estimated parameters α and β . We do this by defining different grouping G then estimating the parameters by maximizing $l(\alpha, \beta, G)$. For example, we can define G as the set of groups including a high-budget product with budget > 100 million to estimate the parameters α and β within the grouping by MLE. The results are summarized in Table 3.

Table 3 A Rationale For Blockbuster and Niche

Max Budget (million USD)	≥ 200	≥ 100 & < 200	≥ 50 & < 100	≥ 10 & < 50	≥ 4 & < 10	< 4
α	1.0465	0.7795	0.5724	0.6054	0.4403	0.3174
β	0.0297	0.1871	0.2740	0.3256	0.7022	2.5060

From Table 3, we observe that α for a group is increasing and β for a group is decreasing on average in the maximum budget in that group. When at least one high-budget movie enters the fray, the network effect parameter would be small and a customer's choice depends heavily on the *intrinsic appeal* of each movie. This provides the rationale for the blockbuster strategy adopted by major studios in Hollywood which aim at a "hit" product. On the contrary, a customer's choice depends heavily on the *sales* as the high-budget product is out of the picture. Hence the network effect is more influential in this case. This is consistent in the expansion of the movie variety on Amazon, Netflix and iTunes. The insight from Table 3 is similar to Zhu and Zhang (2010) that studies the impact of online reviews on sales of video games and indicates that online reviews are more influential for less popular games.

7. Conclusion

We formulate a three-stage game to study how the growing network effect affects the evolution of market structure in culture-related industries such as the movie and music industry. We show that the increasing network effect always contributes to the demand concentration on a small number of products no matter whether the current network effect is weak or strong. However, the equilibrium behavior of product variety and quality level critically depends on the strength of the network effect. We show that the product variety in equilibrium has a U-shape relationship with the network effect magnitude and the product quality level in equilibrium has an inverted U-shape relationship with the network effect magnitude. This result can be used to unify the long tail theory and blockbuster phenomenon. In the theoretical proof, we cannot exactly identify the

changing point shared by both the U-shape and the inverted U-shape relationships as 1, though we numerically verify this is the case. The threshold we identify is shown close to 1 and we hope to close the gap in future research. Lastly, future research can also focus on testing our theory with full-blown empirical analysis in a specific industry.

References

- Abeliuk, A., G. Berbeglia, M. Cebrian, P. Van Hentenryck. 2016. Assortment optimization under a multinomial logit model with position bias and social influence. *4OR* **14**(1) 57–75.
- Anderson, C. 2006. *The long tail: Why the future of business is selling less of more*. Hachette Books.
- Arthur, W. B., Y. M. Ermoliev, Y. M. Kaniovski. 1986. Strong laws for a class of path-dependent stochastic processes with applications. *Stochastic Optimization*. Springer Berlin Heidelberg, 287–300.
- Athreya, K. B., S. Karlin. 1968. Embedding of urn schemes into continuous time Markov branching processes and related limit theorems. *Ann. Math. Statist.* **39**(6) 1801–1817.
- Athreya, K.B., P.E. Ney. 1972. *Branching Processes*. Springer, Berlin.
- Bar-Isaac, H., G. Caruana, V. Cuñat. 2012. Search, design, and market structure. *The American Economic Review* **102**(2) 1140–1160.
- Benaïm, M. 1999. Dynamics of stochastic approximation algorithms. J. Azéma, M. Émery, M. Ledoux, M. Yor, eds. *Séminaire de Probabilités XXXIII*. Lecture Notes in Mathematics, Vol. 1709 (Springer, Berlin) 1–68.
- Benner, M. J., J. Waldfogel. 2016. The song remains the same? technological change and positioning in the recorded music industry. *Strategy Science* **1**(3) 129–147.
- Brynjolfsson, E., Y. Hu, D. Simester. 2011. Goodbye pareto principle, hello long tail: The effect of search costs on the concentration of product sales. *Management Science* **57**(8) 1373–1386.
- Brynjolfsson, E., Y. Hu, M. D. Smith. 2010. Research commentary?long tails vs. superstars: The effect of information technology on product variety and sales concentration patterns. *Information Systems Research* **21**(4) 736–747.
- Brynjolfsson, Erik, Yu Jeffrey Hu, Michael D Smith. 2006. From niches to riches: Anatomy of the long tail. *sloan management Review* **47**(4) 67–71.

-
- Cabral, L. 2011. Dynamic price competition with network effects. *The Review of Economic Studies* **78**(1) 83–111.
- Cachon, G. P., C. Terwiesch, Y. Xu. 2008. On the effects of consumer search and firm entry in a multiproduct competitive market. *Marketing Science* **27**(3) 461–473.
- Chung, F., S. Handjani, D. Jungreis. 2003. Generalizations of polya’s urn problem. *Annals of combinatorics* **7**(2) 141–153.
- Doganoglu, T. 2003. Dynamic price competition with consumption externalities. *Netnomics* **5**(1) 43–69.
- Economides, N. 1996. The economics of networks. *Int. J. Ind. Organ.* **14**(6) 673–699.
- Elberse, A. 2008. Should you invest in the long tail? *Harvard business review* **86**(7/8) 88.
- Elberse, A., F. Oberholzer-Gee. 2006. *Superstars and underdogs: An examination of the long tail phenomenon in video sales*. Division of Research, Harvard Business School.
- Fleder, D., K. Hosanagar. 2009. Blockbuster culture’s next rise or fall: The impact of recommender systems on sales diversity. *Management Sci.* **55**(5) 697–712.
- Goldmanis, M., A. Hortaçsu, C. Syverson, Ö. Emre. 2010. E-commerce and the market structure of retail industries. *The Economic Journal* **120**(545) 651–682.
- Hervás-Drane, Andres. 2015. Recommended for you: The effect of word of mouth on sales concentration. *International Journal of Research in Marketing* **32**(2) 207–218.
- Hinz, O., J. Eckert, B. Skiera. 2011. Drivers of the long tail phenomenon: an empirical analysis. *Journal of management information systems* **27**(4) 43–70.
- Hosanagar, K., D. Fleder, D. Lee, A. Buja. 2013. Will the global village fracture into tribes? recommender systems and their effects on consumer fragmentation. *Management Science* **60**(4) 805–823.
- Hu, M., J. Milner, J. Wu. 2016. Liking and following and the newsvendor: Operations and marketing policies under social influence. *Management Sci.* **62**(3) 867–879.
- Kumar, A., M. D. Smith, R. Telang. 2014. Information discovery and the long tail of motion picture content. *Management Information Systems Quarterly* **38**(4) 1057–1078.
- Maldonado, F., P. Van Hentenryck, G. Berbeglia, F. Berbeglia. 2015. Popularity signals in trial-offer markets Working Paper.

- Mitchell, M. F., A. Skrzypacz. 2006. Network externalities and long-run market shares. *Economic Theory* **29**(3) 621–648.
- Natividad, G. 2013. Financial slack, strategy, and competition in movie distribution. *Organization Science* **24**(3) 846–864.
- Nevel'son, M. B., R. Z. Khas' minskii. 1973. *Stochastic approximation and recursive estimation*. American Mathematical Society, Providence.
- Oestreicher-Singer, G., A. Sundararajan. 2010. Recommendation networks and the long tail of electronic commerce Working Paper.
- Pemantle, R. 1990. Nonconvergence to unstable points in urn models and stochastic approximations. *The Annals of Probability* **18**(2) 698–712.
- Shanthikumar, J. G., D. D. Yao. 1986. The preservation of likelihood ratio ordering under convolution. *Stochastic Processes and their Applications* **23**(2) 259–267.
- Smaili, K., T. Kadri, S. Kadry. 2013. Hypoexponential distribution with different parameters. *Applied mathematics* **4**(04) 624.
- Stewart, J. B. 2013. The flop looms as studios lean on blockbusters. *New York Times* <http://www.nytimes.com/2013/05/18/business/as-hollywood-leans-on-blockbusters-the-flop-looms.html>.
- Sun, M. 2012. How does the variance of product ratings matter? *Management Science* **58**(4) 696–707.
- Tan, T., S. Netessine, L. M. Hitt. 2016. An empirical study of the impact of product variety on demand concentration .
- Tucker, C., J. Zhang. 2011. How does popularity information affect choices? a field experiment. *Management Science* **57**(5) 828–842.
- Wang, R., Z. Wang. 2016. Consumer choice models with endogenous network effects. *Management Science* .
- Zentner, A., M. Smith, C. Kaya. 2013. How video rental patterns change as consumers move online. *Management Science* **59**(11) 2622–2634.
- Zhou, W., W. Duan. 2012. Online user reviews, product variety, and the long tail: An empirical investigation on online software downloads. *Electronic Commerce Research and Applications* **11**(3) 275–289.
- Zhu, F., X. Zhang. 2010. Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics. *Journal of marketing* **74**(2) 133–148.

Zhu, T. 2009. Nonlinear Pólya urn models and self-organizing processes. Dissertation, University of Pennsylvania, PA.

Appendix

A. Mathematical Preliminaries for Proofs

A.1. A Stochastic Approximation Process

This subsection reviews some basic definitions and concepts that are useful in the proofs for the theorems in this paper. Consider the stochastic approximation process defined by recurrence relation (2). This process is defined on a n -dimensional unit simplex, denoted by $S = [0, 1]^n$. Let $\Phi(\mathbf{x}) = \phi(\mathbf{x}) - \mathbf{x}$ and $\zeta(t+1) = \mathbf{e}(t+1) - \mathbb{E}[\mathbf{e}(t+1)|\mathbf{x}(t)]$. Then, for $t=0, 1, 2, \dots$, this stochastic approximation process can be rewritten as

$$\mathbf{x}(t+1) - \mathbf{x}(t) = \frac{1}{t+1+n} [\Phi(\mathbf{x}(t)) + \zeta(t+1)], \quad (7)$$

where $\mathbb{E}[\zeta(t+1)|\mathbf{x}(t)] = 0$.

Let $B = \{\mathbf{x} \in S : \Phi(\mathbf{x}) = \mathbf{0}\}$ be the set of all critical points of $\Phi(\mathbf{x})$ (or fixed points of $\phi(\mathbf{x})$). A point $\mathbf{x}$ in B must satisfy

$$\frac{q_i x_i^\beta}{q_1 x_1^\beta + \dots + q_n x_n^\beta} = x_i, \quad (8)$$

for $i = 1, \dots, n$. If $x_i > 0$ for all i and $\beta \neq 1$, from (8), we have $\mathbf{x} = (q_1^{\frac{1}{1-\beta}}, \dots, q_n^{\frac{1}{1-\beta}}) / \sum_{i=1}^n q_i^{\frac{1}{1-\beta}}$, which is the unique critical point in the interior of S (see Theorem 5.1 in [Maldonado et al. 2015](#), for details). Let B_0 be the intersection of B and the interior of S . We have that $B_0 = \{(q_1^{\frac{1}{1-\beta}}, \dots, q_n^{\frac{1}{1-\beta}}) / \sum_{i=1}^n q_i^{\frac{1}{1-\beta}}\}$ if $\beta \neq 1$ and is empty otherwise. If $x_i = 0$ for some i , then $\mathbf{x}$ satisfying (8) must be on the boundary of S . We categorize the boundary critical points into three sets:

Set 1 The set of vertices, denoted by $B_1 = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ where $\mathbf{e}_i$ is the unit vector with i th entry being 1 and 0 otherwise.

Set 2 The set of points in the *relative interior* of any facet of the simplex, denoted by $B_2 = \{\mathbf{x} : x_i = q_i^{\frac{1}{1-\beta}} / \sum_{i \in Q} q_i^{\frac{1}{1-\beta}} \text{ for } i \in Q, x_i = 0 \text{ for } i \notin Q, Q \text{ is any subset of } \{1, \dots, n\}, \text{ with } 1 < |Q| < n\}$. Note that B_2 is nonempty if and only if $\beta \neq 1$.

Set 3 The union of all convex hulls of two or more vertices $\mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_k}$ whenever $q_{i_1} = \dots = q_{i_k}$ and $\beta = 1$.

We denote this set by B_3 . B_3 is a union of intervals or a polyhedrons whose vertices are vertices of S .

Note that B_3 is nonempty if and only if $\beta = 1$ and two or more firms' qualities are equal.

Consequently, we have $B = \bigcup_{i=0}^3 B_i$. Thus B is a union of finitely many connected components. According to Theorem 7.3 in [Nevel'son and Khas' minskii \(1973\)](#), Chapter 2, or Theorem 3.1 in [Arthur et al. \(1986\)](#),

the process $\mathbf{x}(t)$ converges with probability 1 either to an isolated point in B or to one of its connected components for any initial condition.

We now consider the stability of a critical point $\mathbf{p}$ in B , depending on whether it has the property: the process $\mathbf{x}(t)$ with any initial point that lies in a sufficiently small neighborhood of $\mathbf{p}$ converges to $\mathbf{p}$ almost surely.

DEFINITION A.1. (Arthur et al. 1986). A point $\mathbf{p} \in B$ is called a *stable point* if for some neighborhood U of $\mathbf{p}$, there is a symmetric positive definite matrix D such that

$$\langle D \cdot \Phi(\mathbf{v}), \mathbf{v} - \mathbf{p} \rangle < 0, \quad \forall \mathbf{v} \neq \mathbf{p} \text{ and } \mathbf{v} \in U \cap S,$$

where $\langle \cdot, \cdot \rangle$ stands for the inner product. Similarly, a point $\mathbf{p} \in B$ is called an *unstable point* if for some neighborhood U of $\mathbf{p}$, there is a symmetric positive definite matrix D such that

$$\langle D \cdot \Phi(\mathbf{v}), \mathbf{v} - \mathbf{p} \rangle > 0, \quad \forall \mathbf{v} \neq \mathbf{p} \text{ and } \mathbf{v} \in U \cap S.$$

DEFINITION A.2. (Pemantle 1990). A point $\mathbf{p} \in B$, is *linearly stable* if all the eigenvalues of $\partial\Phi(\mathbf{p})/\partial\mathbf{x}$ (the *Jacobian* of Φ at $\mathbf{p}$) have negative real parts. If some eigenvalue of $\partial\Phi(\mathbf{p})/\partial\mathbf{x}$ has a positive real part, then $\mathbf{p}$ is called a *linearly unstable point*.

LEMMA A.1. A point $\mathbf{p} \in B$ is linearly stable then it is stable.

Proof. For the proof, see Page 107 in Nevel'son and Khas' minskii (1973). □

THEOREM A.1. If $\mathbf{p} \in B$ is a stable point, then $\mathbb{P}\{\mathbf{x}(t) \rightarrow \mathbf{p}\} > 0$.

Proof. See Theorem 5.1 in Arthur et al. (1986). □

THEOREM A.2. If $\mathbf{p} \in B$ is linearly unstable, then $\mathbb{P}\{\mathbf{x}(t) \rightarrow \mathbf{p}\} = 0$.

Proof. See Theorem 1 in Pemantle (1990). □

A.2. Branching Process

One of the tools to study the distribution of the stochastic sequence (7) is the *exponential embedding* method, introduced by Athreya and Karlin (1968). Consider an n -dimensional continuous Markov branching process $\mathbf{Y}(s) = \{Y_1(s), \dots, Y_n(s)\}$ with initial size $\mathbf{Y}(0) = \{1, \dots, 1\}$. $\{Y_i(s)\}_{i=1}^n$ are all *pure-birth* processes and mutually independent. We will build the relationship between the branching process $\mathbf{Y}$ and the demand process $\mathbf{d}$ in the base model. For each group i , the k th split time $\gamma_i(k) \triangleq \inf\{s' : Y_i(s + s') = k + 1 | Y_i(s) = k\}$ is exponentially distributed with mean $q_i k^\beta$. Let τ_t , $t = 1, 2, \dots$, denote the successive times at which splits occur

in the whole collection of n groups. Given τ_t and $\mathbf{Y}(\tau_t)$, we know τ_{t+1} follows an exponential distribution with mean $\sum_{i=1}^n q_i Y_i(\tau_t)^\beta$ and $\mathbb{P}\{Y_i(\tau_{t+1}) = Y_i(\tau_t) + 1 | Y_i(\tau_t)\} = q_i Y_i(\tau_t)^\beta / \sum_{i=1}^n q_i Y_i(\tau_t)^\beta$ for $i = 1, \dots, n$.

THEOREM A.3. *The stochastic processes $\{\mathbf{Y}(\tau_t), t = 0, 1, \dots\}$ and $\{\mathbf{d}(t), t = 0, 1, \dots\}$ are equivalent.*

Proof. See Theorem 1 in [Athreya and Karlin \(1968\)](#) for two-dimensional case; see Theorem 2 in [Athreya and Ney \(1972\)](#), Section 9.1, for the case with more than 2 dimensions. $\square$

Let $\Gamma_i(N) = \gamma_i(1) + \dots + \gamma_i(N)$ be the sum of the first N split times in group i . Thus $Y_i(s) = N + 1$ for $\Gamma_i(N) \leq s < \Gamma_i(N + 1)$. It is known that $\Gamma_i(N)$ follows a *hyperexponential* distribution and its P.D.F. $f_{i,N}(s)$ is

$$f_{i,N}(s) = \sum_{j=1}^N q_i j^\beta e^{-(q_i j^\beta)s} H_{j,N}, \quad s \geq 0,$$

where $H_{j,N} \triangleq \prod_{\substack{l=1 \\ l \neq j}}^N \frac{l^\beta}{l^\beta - j^\beta}$, for $j = 1, \dots, N$. By Lemma 3.2.2 in [Zhu \(2009\)](#), if $\beta > 1$, $\Gamma_i(\infty) \triangleq \lim_{N \rightarrow \infty} \Gamma_i(N)$ exists almost surely and has the probability density function:

$$f_i(s) = \begin{cases} \sum_{j=1}^{\infty} q_i j^\beta e^{-(q_i j^\beta)s} H_j, & s > 0, \\ 0, & s \leq 0, \end{cases}$$

where $H_j \triangleq \prod_{\substack{l=1 \\ l \neq j}}^{\infty} \frac{l^\beta}{l^\beta - j^\beta}$, $\forall j \in \mathbb{N}$. Note that, if $\beta > 1$, by Lemma A.2(ii), $f_{i,N}(s)$ uniformly converges to $f_i(s)$, as $N \rightarrow +\infty$.

LEMMA A.2. *If $\beta > 1$ and $q_i > 0$ for all i , then*

$$(i) \quad \sum_{j=1}^N H_{j,N} = 1, \quad \forall N \geq 2;$$

$$(ii) \quad \text{As } j \rightarrow +\infty, |H_j| = O(je^{-j\pi \cot(\frac{\pi}{\beta})}); \text{ hence, } H_j \rightarrow 0 \text{ for } \beta > 2.$$

Proof. For part (i), see Corollary 3.2.3 in [Zhu \(2009\)](#).

We next prove part (ii). Consider $\ln |H_j|$ and we have

$$\begin{aligned} \ln |H_j| &= \sum_{l=1}^{j-1} \ln \frac{l^\beta}{j^\beta - l^\beta} + \sum_{l=j+1}^{\infty} \ln \frac{l^\beta}{l^\beta - j^\beta} \\ &= \int_0^j \ln \frac{x^\beta}{j^\beta - x^\beta} dx + \int_j^\infty \ln \frac{x^\beta}{x^\beta - j^\beta} dx + O(\ln j) \\ &= j \int_0^1 \ln \frac{x^\beta}{1 - x^\beta} dx + j \int_1^\infty \ln \frac{x^\beta}{x^\beta - 1} dx + O(\ln j) \\ &= j \int_0^1 \ln \frac{x^\beta}{1 - x^\beta} + \frac{1}{x^2} \ln \frac{1}{1 - x^\beta} dx + O(\ln j) \\ &= -j\pi \cot\left(\frac{\pi}{\beta}\right) + O(\ln j). \end{aligned}$$

Hence we have $|H_j| = O(je^{-j\pi \cot(\frac{\pi}{\beta})})$ and $H_j \rightarrow 0$ for $\beta > 2$. $\square$

LEMMA A.3. For $\beta > 1$ and $i = 1, \dots, n$,

(i) $f_i(s)$ is at least first-order continuous and differentiable, and $f_i(0) = f'_i(0) = \lim_{s \rightarrow +\infty} f_i(s) = \lim_{s \rightarrow +\infty} f'_i(s) = 0$;

(ii) $f_i(s)$ is logconcave on $[0, +\infty)$, which means $f_i(s)$ is first increasing then decreasing and its C.D.F, denoted by $F_i(s)$, has an increasing failure rate on $[0, +\infty)$;

Proof. We first prove part (i). We know $\Gamma_i(N)$ is the sum of N independent exponentials. Let $f_{i,N}(s)$ be the P.D.F. of $\Gamma_i(N)$. By Proposition 2 in [Smaili et al. \(2013\)](#) we have $f_{i,N}(0) = f'_{i,N}(0) = 0$, $\forall N \geq 3$. By the uniform convergence of $f'_{i,N}(s)$ and $f_{i,N}(s)$ as $N \rightarrow +\infty$, we have $f_i(0) = f'_i(0) = 0$. Integrability of f_i and f'_i implies $\lim_{s \rightarrow +\infty} f_i(s) = \lim_{s \rightarrow +\infty} f'_i(s) = 0$.

We next prove part (ii). Since the convolution of logconcave functions is also logconcave, we know $f_{i,N}(s)$ is logconcave for all i, N . Then by the convergence of $f_{i,N}(s)$ and the continuity of $f_i(s)$ we have $f_i(s)$ is logconcave. $\square$

When $\beta > 1$, the discrete process $\mathbf{d}(t)$ eventually converges to a *monopoly*, that is, for some large enough time t beyond which all subsequent customers select the same product. Define the *attraction time* by $\Gamma_a = \min\{t' : \mathbf{e}(t) = \mathbf{e}(t'), \forall t > t'\}$.

THEOREM A.4. If $\beta > 1$, then a monopoly occurs almost surely, i.e., $\mathbb{P}\{\Gamma_a < \infty\} = 1$.

Proof. See Theorem 3.3.1 in [Zhu \(2009\)](#). $\square$

Next, we consider which product will become monopoly if $\beta > 1$. By Theorem A.3 and Theorem A.4, the probability that product i eventually goes to monopoly must be equal to $\mathbb{P}\{\Gamma_i(\infty) < \Gamma_j(\infty), \text{ for all } j \neq i\}$, where “ $<$ ” is equivalent to “ $\leq$ ” since all the variables $\Gamma_i(\infty)$ have continuous distribution, which implies that they are distinct with probability one. These probabilities are formally established in the following theorem.

THEOREM A.5. If $\beta > 1$, for $i = 1, \dots, n$,

$$\mathbb{P}\{\Gamma_i(\infty) < \Gamma_j(\infty), \text{ for all } j \neq i\} = \lim_{N \rightarrow \infty} \sum_{j_1=1}^N \cdots \sum_{j_n=1}^N \frac{q_i j_i^\beta}{q_1 j_1^\beta + \cdots + q_n j_n^\beta} \prod_{k=1}^n H_{j_k, N},$$

where $H_{j,N} \triangleq \prod_{\substack{l=1 \\ l \neq j}}^N \frac{l^\beta}{l^\beta - j^\beta}$, for $j = 1, \dots, N$. Furthermore if $\beta > 2$, for $i = 1, \dots, n$,

$$\mathbb{P}\{\Gamma_i(\infty) < \Gamma_j(\infty), \text{ for all } j \neq i\} = \sum_{j_1=1}^{\infty} \cdots \sum_{j_n=1}^{\infty} \frac{q_i j_i^\beta}{q_1 j_1^\beta + \cdots + q_n j_n^\beta} \prod_{k=1}^n H_{j_k},$$

where $H_j \triangleq \prod_{\substack{l=1 \\ l \neq j}}^{\infty} \frac{l^\beta}{l^\beta - j^\beta}$, $\forall j \in \mathbb{N}$.

Proof. For $\beta > 1$ and $i = 1, \dots, n$, we have

$$\begin{aligned}
& \mathbb{P}\{\Gamma_i(N) < \Gamma_j(N), \text{ for all } j \neq i\} \\
&= \int_{s_i < s_j \text{ for all } j \neq i} f_{1,N}(s_1) \cdots f_{n,N}(s_n) ds_1 ds_2 \cdots ds_n \\
&= \int_0^\infty f_{i,N}(s_i) ds_i \int_{s_i}^\infty f_{1,N}(s_1) ds_1 \cdots \int_{s_i}^\infty f_{n,N}(s_n) ds_n \\
&= \int_0^\infty f_{i,N}(s_i) ds_i \int_{s_i}^\infty f_{1,N}(s_1) ds_1 \cdots \int_{s_i}^\infty \sum_{j_n=1}^N q_n j_n^\beta e^{-(q_n j_n^\beta) s_n} H_{j_n,N} ds_n \\
&= \sum_{j_n=1}^N \left(\int_0^\infty f_{i,N}(s_i) e^{-(q_n j_n^\beta) s_i} ds_i \int_{s_i}^\infty f_{1,N}(s_1) ds_1 \cdots \int_{s_i}^\infty f_{n-1,N}(s_{n-1}) ds_{n-1} \right) H_{j_n,N} \\
&= \cdots \\
&= \sum_{j_1, \dots, j_{i-1}, j_{i+1}, \dots, j_n=1}^N \left(\int_0^\infty f_{i,N}(s_i) e^{-(q_1 j_1^\beta + \cdots + q_{i-1} j_{i-1}^\beta + q_{i+1} j_{i+1}^\beta + \cdots + q_n j_n^\beta) s_i} ds_i \right) \prod_{\substack{k=1 \\ k \neq i}}^n H_{j_k,N} \\
&= \sum_{j_1, \dots, j_n=1}^N \frac{q_i j_i^\beta}{q_1 j_1^\beta + \cdots + q_n j_n^\beta} \prod_{k=1}^n H_{j_k,N}.
\end{aligned}$$

Thus we have

$$\mathbb{P}\{\Gamma_i(\infty) < \Gamma_j(\infty), \text{ for all } j \neq i\} = \lim_{N \rightarrow \infty} \sum_{1 \leq j_1, \dots, j_n \leq N} \frac{q_i j_i^\beta}{q_1 j_1^\beta + \cdots + q_n j_n^\beta} \prod_{k=1}^n H_{j_k,N}.$$

Finally, by Lemma A.2(ii), the rest of this theorem is proved. $\square$

B. Proofs

In the following proofs, some notation is defined in Appendix A and we will use them without any specification.

Proof of Theorem 1. As mentioned in Appendix A, the stochastic approximation process $\mathbf{x}(t)$ converges with probability 1 either to a point of B or to the boundary of one of its connected components. That is, $\lim_{t \rightarrow \infty} \mathbf{x}(t) = \mathbf{x}^*(\mathbf{q}, \beta)$ exists almost surely, that may be a single point or a random variable. Next, we will give the explicit form of $\mathbf{x}^*(\mathbf{q}, \beta)$ under three cases.

For any β , $0 \leq \beta < 1$, the fixed points set $B = \bigcup_{i=0}^2 B_i$ (see Section A.1). To prove this part, we resort to a continuous dynamic process:

$$\frac{d\hat{\mathbf{x}}(t)}{dt} = \Phi(\hat{\mathbf{x}}(t)). \tag{9}$$

By Proposition 4.1 in Benaïm (1999), the interpolation of the discrete process $\mathbf{x}(t)$ is almost surely an asymptotic pseudo trajectory of the flow induced by the associated ODE (9). Thus, if the continuous flow $\hat{\mathbf{x}}(t)$ defined by (9) converges to B_0 for any initial point $\mathbf{x}(0)$ in the interior of S , then the discrete process

$\mathbf{x}(t)$ defined by (2) converges to B_0 almost surely, because the discrete process $\mathbf{x}(t)$ with initial point $\mathbf{x}(0) = \{\frac{1}{n}, \dots, \frac{1}{n}\}$ cannot reach the boundary of S in an arbitrarily finite time. The convergence of the flow $\hat{\mathbf{x}}(t)$ has been proved according to the proof of Theorem 6.2 in [Maldonado et al. \(2015\)](#) and hence the proof of part (i) is complete.

For $\beta = 1$, the fixed points set $B = B_1 \cup B_3$, where each connected component of B_1 is an isolated point and each connected component of B_3 is a convex hull of some vertices of S . Define $L = \{i : q_i \text{ is the largest entry of } \mathbf{q}\}$ and $B_L = \{\mathbf{x} \in S : \sum_{i \in L} x_i = 1, \sum_{i \notin L} x_i = 0\}$. Note that $B_L \subset B_3$. We will prove $\mathbf{x}(t)$ converges to B_L almost surely.

We first prove that for any point $\mathbf{e}_i \in B_1$ and $i \notin L$, $\mathbb{P}\{\mathbf{x}(t) \rightarrow \mathbf{e}_i\} = 0$. Without loss of generality, we assume $1 \notin L$ and we will prove $\mathbb{P}\{\mathbf{x}(t) \rightarrow \mathbf{e}_1\} = 0$. At the point $\mathbf{p} = \mathbf{e}_1$, we have

$$T_{\mathbf{p}} = \begin{bmatrix} 0 & -\frac{q_2}{q_1} & -\frac{q_3}{q_1} & \dots & -\frac{q_n}{q_1} \\ 0 & \frac{q_2}{q_1} & 0 & \dots & 0 \\ 0 & 0 & \frac{q_3}{q_1} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \frac{q_n}{q_1} \end{bmatrix} - I_n,$$

where I_n is an identity matrix of order n . From the diagonal of $T_{\mathbf{p}}$, it can be seen that all eigenvalues of $T_{\mathbf{p}}$ are $\{-1, \frac{q_2}{q_1} - 1, \dots, \frac{q_n}{q_1} - 1\}$ in which at least one element is positive. Thus $\mathbf{e}_1$ is linearly unstable and, by Theorem A.2 which derived from Theorem 1 in [Pemantle \(1990\)](#), $\mathbb{P}\{\mathbf{x}(t) \rightarrow \mathbf{e}_1\} = 0$. Since B_1 contains finite points, we have $\mathbb{P}\{\lim_{t \rightarrow \infty} \mathbf{x}(t) \in B_1 \setminus B_L\} = 0$.

Then we will prove that the discrete process $\mathbf{x}(t)$ converges to $B_3 \setminus B_L$ with probability 0 if $\beta = 1$. Note that, in the light of Theorem A.2, we can prove $\mathbb{P}\{\mathbf{x}(t) \rightarrow \mathbf{p}\} = 0, \forall \mathbf{p} \in B_3 \setminus B_L$, but we want to prove that $\mathbf{x}(t)$ with probability 0 converges to $B \setminus B_L$ which contains uncountable number of points. To this end, we arbitrarily choose a connected component of $B_3 \setminus B_L$, say the convex hull of $\{\mathbf{e}_1, \mathbf{e}_2\}$ without loss of generality, denoted by $\mathbf{conv}\{\mathbf{e}_1, \mathbf{e}_2\}$, and we will show that $\mathbb{P}\{\lim_{t \rightarrow \infty} \mathbf{x}(t) \in \mathbf{conv}\{\mathbf{e}_1, \mathbf{e}_2\}\} = 0$. To this end, we consider a n -order *auxiliary* matrix

$$A = \begin{bmatrix} 1 & 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & I_{n-2} \end{bmatrix} \in \mathbb{R}^{n-1 \times n}$$

which maps $(x_1, x_2, \dots, x_n)^T \in \mathbb{R}^n$ into $(x_1 + x_2, x_3, \dots, x_n)^T \in \mathbb{R}^{n-1}$ and introduce an *auxiliary* process based on (7)

$$A\mathbf{x}(t+1) - A\mathbf{x}(t) = \frac{1}{t+1+n} [A\Phi(\mathbf{x}(t)) + A\zeta(t+1)]. \quad (10)$$

Thus, to prove $\mathbb{P}\{\lim_{t \rightarrow \infty} \mathbf{x}(t) \in \mathbf{conv}\{\mathbf{e}_1, \mathbf{e}_2\}\} = 0$, it is sufficient to prove $\mathbb{P}\{\lim_{t \rightarrow \infty} A\mathbf{x}(t) = (1, 0, \dots, 0)^T\} = 0$.

Consider a new demand process in which firm 1 has amalgamated with firm 2 by adding up their market

shares. Denote by $\tilde{\mathbf{x}}(t) = A\mathbf{x}(t)$ the new vector of market shares of the $n-1$ products. As with the definition of $\Phi(\mathbf{x}(t))$, $\tilde{\Phi}(\tilde{\mathbf{x}}(t))$ can be defined in this $n-1$ -dimensional problem. Also $\tilde{\mathbf{e}}(t)$ and $\tilde{\zeta}(t)$ can be correspondingly defined. Since each product is selected with probability linear in the market share of this product in the case with $\beta = 1$ and $q_1 = q_2$, it is not hard to see that the process (10) is equivalent to

$$\tilde{\mathbf{x}}(t+1) - \tilde{\mathbf{x}}(t) = \frac{1}{t+1+n} [\tilde{\Phi}(\tilde{\mathbf{x}}(t)) + \tilde{\zeta}(t+1)]. \quad (11)$$

That is, to see the convergence of the process $A\mathbf{x}(t)$, we only need to focus on the convergence of the process $\tilde{\mathbf{x}}(t)$, defined by (11). As with the previous example in which we proved $\mathbb{P}\{\lim_{t \rightarrow \infty} \mathbf{x}(t) = \mathbf{e}_1\} = 0$, we can prove $\tilde{\mathbf{x}}(t)$, defined by (11), converges to $(1, 0, \dots, 0)^T$ in probability 0 by Theorem A.2. Hence, we have $\mathbb{P}\{\lim_{t \rightarrow \infty} \mathbf{x}(t) \in \text{conv}\{\mathbf{e}_1, \mathbf{e}_2\}\} = 0$.

It is also not hard to show that the cumulated perturbations $\sum_{\tau=0}^t \frac{1}{\tau+1+n} \zeta(\tau+1)$, where $\zeta(\cdot)$ is the noise in (7), form a martingale and converge. It follows that at a sufficiently large time, there is not enough perturbation for the process to make infinitely many “trips” to-and-fro between two or more connected components. Since $B_1 \cup B_3 \setminus B_L$ has a finite number of connected components, $\mathbf{x}(t)$ converges to $B_1 \cup B_3 \setminus B_L$ with probability 0. Therefore, $\mathbf{x}(t)$ converges with probability 1 to B_L . Given that the demand will be concentrated on the products in the set L , by Theorem 2.1 in Chung et al. (2003), the limit of $\mathbf{x}(t)$ is uniformly distributed on B_L .

We next prove part (iii). For any $\beta > 1$, the fixed points set $B = \bigcup_{i=0}^2 B_i$. Clearly B is a union of a finite number of isolated points. Following the same procedure in part (ii), we have that all points in $B_0 \cup B_2$ are linearly stable and all points in B_1 are linearly stable. By Theorem A.1 and A.2 we have $\mathbb{P}\{\lim_{t \rightarrow \infty} \mathbf{x}(t) \in B_1\} = 1$. Alternatively, we can prove this result by considering the demand process as a branching process (see Section A.2 for details) and using Theorem A.3 and A.4. Furthermore, Theorem A.5 shows the probability that the process $\mathbf{x}(t)$ converges to each point in B_1 . $\square$

Proof of Lemma 1. See Theorem A.5 for the proof. $\square$

The proof of Proposition 1 makes use of the following lemma.

$$\text{LEMMA B.1. } \lim_{N \rightarrow \infty} \sum_{j_1=1}^N \cdots \sum_{j_n=1}^N \frac{q_1 j_1^\beta \cdots q_n j_n^\beta}{(q_1 j_1^\beta + \cdots + q_n j_n^\beta)^2} \prod_{k=1}^n H_{j_k, N} > 0, \forall i, l = 1, \dots, n, i \neq l, \forall N \geq 1.$$

We first prove

$$\sum_{j_1=1}^N \cdots \sum_{j_n=1}^N \frac{q_1 j_1^\beta \cdots q_n j_n^\beta}{(q_1 j_1^\beta + \cdots + q_n j_n^\beta)^2} \prod_{k=1}^n H_{j_k, N} > 0.$$

In fact, by using the technique in the proof of Theorem A.5, we have

$$\begin{aligned}
0 &< \int_0^\infty s_1 f_{1,N}(s_1) f_{2,N}(s_1) ds_1 \int_{s_1}^\infty f_{3,N}(s_3) ds_3 \cdots \int_{s_1}^\infty f_{n,N}(s) ds \\
&= \sum_{j_1, \dots, j_n=1}^N \int_0^\infty q_1 j_1^\beta q_2 j_2^\beta s_1 e^{-(q_1 j_1^\beta + \dots + q_n j_n^\beta) s_1} ds_1 \cdot \prod_{\substack{k=1 \\ k \neq i}}^n H_{j_k} \\
&= \sum_{j_1, \dots, j_n=1}^N \frac{q_1 j_1^\beta \cdot q_2 j_2^\beta}{(q_1 j_1^\beta + \dots + q_n j_n^\beta)^2} \prod_{k=1}^n H_{j_k \cdot N}.
\end{aligned}$$

Then this lemma holds by the uniform convergence of $f_{i,N}(s)$ for $i = 1, \dots, n$. $\square$

Proof of Proposition 1. Suppose $n \geq 2$. From Theorem 1(i)(ii), we see that for $0 \leq \beta \leq 1$, $\mathbb{E}x_i^*(n, \beta)$ is increasing in q_i and decreasing in q_j for any $j \neq i$. Next we will consider the case with $\beta > 1$. We focus on firm 1 without loss of generality and have that

$$\begin{aligned}
\frac{\partial}{\partial q_1} \mathbb{E}x_1^*(n, \beta) &= \frac{\partial}{\partial q_1} \mathbb{P}\{\mathbf{x}^*(\mathbf{q}, \beta) = \mathbf{e}_1\} \\
&= \lim_{N \rightarrow \infty} \sum_{j_1, \dots, j_n=1}^N \frac{\partial}{\partial q_1} \frac{q_1 j_1^\beta}{q_1 j_1^\beta + \dots + q_n j_n^\beta} \prod_{k=1}^n H_{j_k, N} \\
&= \frac{1}{q_1} \lim_{N \rightarrow \infty} \sum_{j_1, \dots, j_n=1}^N \frac{q_1 j_1^\beta (q_2 j_2^\beta + \dots + q_n j_n^\beta)}{(q_1 j_1^\beta + \dots + q_n j_n^\beta)^2} \prod_{k=1}^n H_{j_k, N},
\end{aligned}$$

where the convergence of the limit in the second/third equality comes from the uniform convergence of $f_{i,N}(s)$ for $i = 1, \dots, n$, which uses the technique in the proof of Theorem A.5. Then by Lemma B.1, the expected market share of firm 1 is increasing in its quality level. Similarly, we have that the expected market share of firm 1 is decreasing in its competitors' quality levels. Therefore this proposition is proved. $\square$

Proof of Theorem 2. For tractability, we define $x_i^*(\mathbf{0}, \beta) = 1/n$ in the proof of this theorem. For the simplicity of notation, we write x_i^* for $x_i^*(\mathbf{q}, \beta)$, π_i for $\pi_i(q_i, \mathbf{q}_{-i}, \beta)$.

For $0 \leq \beta < 1$, $x_i^* = q_i^{\frac{1}{1-\beta}} / \sum_{i=1}^n q_i^{\frac{1}{1-\beta}}$, $\pi_i = p x_i^* - c q_i$. Define $x_i^* = 1/n$ for all i at $\mathbf{q} = \mathbf{0}$. Note that $\mathbf{x}^*$ and π^* are discontinuous at $\mathbf{q} = \mathbf{0}$. For $i = 1, \dots, n$, in the interior of E , the first-order derivative (FOD) of firm i 's problem is

$$\frac{\partial \pi_i}{\partial q_i} = \frac{p}{q_i(1-\beta)} x_i^*(1-x_i^*) - c,$$

and the second-order derivative (SOD) is

$$\frac{\partial^2 \pi_i}{\partial q_i^2} = \frac{p}{q_i^2(1-\beta)} x_i^*(1-x_i^*) \left[\frac{1}{1-\beta} (1-2x_i^*) - 1 \right],$$

which is negative if and only if $q_i^{\frac{1}{1-\beta}} > \frac{\beta}{2-\beta} \sum_{j \neq i} q_j^{\frac{1}{1-\beta}}$. Hence π_i is first convex then concave for $q_i \in (0, +\infty)$.

For $0 \leq \beta < 1$, it can be seen from the FOCs for all firms' problems that a symmetric equilibrium of $\mathcal{G}$, if it exists, must be $\mathbf{q} = \frac{(n-1)p}{n^2c(1-\beta)} \cdot (1, \dots, 1)$ where $n \geq 2$. Thus $\mathbf{q}$ is a symmetric equilibrium of $\mathcal{G}$ if and only if it satisfies the following two conditions: (a) Each firm's profit at $\mathbf{q}$ is nonnegative; (b) The SOC's hold for each firm's problem. More precisely these two conditions can be written as

$$\begin{aligned}\pi_i &= p \frac{1}{n} - c \frac{(n-1)p}{n^2c(1-\beta)} \geq 0, \quad i = 1, \dots, n, \\ \frac{\partial^2 \pi_i}{\partial q_i^2} &= \frac{c}{q_i} \left[\frac{1}{1-\beta} \left(1 - \frac{2}{n} \right) - 1 \right] \leq 0, \quad i = 1, \dots, n.\end{aligned}$$

We can find that $\pi_i \geq 0$ implies $\frac{\partial^2 \pi_i}{\partial q_i^2} \leq 0$, and $\pi_i \geq K$ if and only if $n \leq \frac{p}{K+(p-K)\beta}$ for $K \geq 0$. Define $m(\beta, K) = \lfloor \frac{p}{K+(p-K)\beta} \rfloor$. Thus we have that if $2 \leq n \leq m(\beta, 0)$, $\mathbf{q}$ must be an equilibrium of $\mathcal{G}$, and $\pi^* \geq K$ if and only if $n \leq m(\beta, K)$; if $n > m(\beta, 0)$, then there is no symmetric pure-strategy equilibrium since at least one of conditions (a) and (b) does not hold at the point $\mathbf{q}$.

For $\beta = 1$, only the firms whose qualities are largest among the n firms can have positive profits. Thus, any firm has incentive to increase his quality in order to win the whole market (as in Bertrand competition) unless his current quality is p/c , at which the production cost is sufficiently high such that the firm always gets a nonpositive profit even if he wins the whole market. When all firms offer the quality p/c , we can see that all wish to deviate and choose the quality of 0. Thus, no pure-strategy exists.

We next prove in any symmetric mixed-strategy equilibrium, each of the n firms must get an expected profit of 0 if $n > m(\beta, 0)$. Suppose the n firms randomize their qualities according an independent and identical distribution, $G(q_i)$, $i = 1, \dots, n$. Let $\underline{q}$ and $\bar{q}$ be the infimum and supremum of the set of strategies, denoted by R , that each firm possibly choose. We know that in equilibrium, given other firms' strategies, each firm (say firm 1) must be indifferent between any $q_1 \in R$. If $\underline{q} = 0$, then i 's expected profit is 0 and the proof is complete. If $\underline{q} > 0$, since i 's expected profit is continuous for $q_1 \in (0, +\infty)$, we can infer that R can be decomposed as some intervals, so we assume $R = [\underline{q}, \bar{q}]$ without loss of generality.

For $0 \leq \beta < 1$ we have that

$$p \int_{\underline{q}}^{\bar{q}} \dots \int_{\underline{q}}^{\bar{q}} \frac{q_1^{\frac{1}{1-\beta}}}{q_1^{\frac{1}{1-\beta}} + \dots + q_n^{\frac{1}{1-\beta}}} dG(q_2) \dots dG(q_n) = cq_1 + a_1, \quad \forall q_1 \in R, \quad (12)$$

where a_1 is a constant (the expected profit of firm 1). Define $h_1(\mathbf{q}_{-1}) = \left(\frac{\beta}{2-\beta} \sum_{j \neq 1} q_j^{\frac{1}{1-\beta}} \right)^{1-\beta}$. We have shown that $\frac{\partial^2 \pi_1^*}{\partial q_1^2} \leq 0$ if and only if $q_1 \geq h_1(\mathbf{q}_{-1})$. We must have that $\bar{q} < h_1(\bar{q}, \dots, \bar{q})$ because if not, $\bar{q} \geq h_1(\mathbf{q}_{-1})$, $\forall \mathbf{q}_{-1} \in R^{n-1}$ since $h_1(\cdot)$ is (strictly) increasing in q_j , $\forall j \neq 1$; then the left side of Equation (12) is strictly

concave at $q_1 = \bar{q}$ but the right hand side is linear, which is impossible. Note that $\bar{q} < h_1(\bar{q}, \dots, \bar{q})$ is equivalent to $\underline{q} < h_1(\underline{q}, \dots, \underline{q})$ if $\underline{q} > 0$, which implies $\underline{q} < h_1(\mathbf{q}_{-1})$, $\forall \mathbf{q}_{-1} \in R^{n-1}$. Thus, the left side of Equation (12) is strictly convex at $q_1 = \underline{q}$ but the right hand side is linear, which is a contradiction. Therefore, $\underline{q} = 0$.

For $\beta = 1$ we have that

$$p[G(q_1)]^{n-1} = cq_1 + a_2, \quad (13)$$

where a_2 is a constant. Substituting $\bar{q}$ and $\underline{q}$ into Equation (13), we have $p = c\bar{q} + a_2$ and $0 = c\underline{q} + a_2$. It follows that $\bar{q} - \underline{q} = p/c$, which is impossible since $\bar{q} \leq p/c$. Therefore, we must have $\underline{q} = 0$. $\square$

The proof of Theorem 3 makes use of the following lemma. Since all firms are identical, we take firm 1 as the focal firm in the lemma. The results of this lemma can be carried over to any other firm. Before this lemma, we first present some notations as follows:

$$\begin{aligned} y(n) &= \lim_{N \rightarrow \infty} \sum_{j_1=1}^N \cdots \sum_{j_n=1}^N \frac{(j_1^\beta)^2}{(j_1^\beta + \cdots + j_n^\beta)^2} \prod_{k=1}^n H_{j_k, N}, \\ z(n) &= \lim_{N \rightarrow \infty} \sum_{j_1=1}^N \cdots \sum_{j_n=1}^N \frac{(j_1^\beta)^2 \sum_{k \neq 1} j_k^\beta}{(j_1^\beta + \cdots + j_n^\beta)^3} \prod_{k=1}^n H_{j_k, N}, \\ \pi(n) &= py(n), \\ y(n, \mathbf{q}) &= \lim_{N \rightarrow \infty} \sum_{j_1=1}^N \cdots \sum_{j_n=1}^N \frac{(q_1 j_1^\beta)^2}{(q_1 j_1^\beta + \cdots + q_n j_n^\beta)^2} \prod_{k=1}^n H_{j_k, N}, \\ z(n, \mathbf{q}) &= \lim_{N \rightarrow \infty} \sum_{j_1=1}^N \cdots \sum_{j_n=1}^N \frac{(q_1 j_1^\beta)^2 \sum_{k \neq 1} q_k j_k^\beta}{(q_1 j_1^\beta + \cdots + q_n j_n^\beta)^3} \prod_{k=1}^n H_{j_k, N}. \end{aligned}$$

In fact, $\pi(n)$ is firm 1's profit at the symmetric equilibrium $\mathbf{q}^*$ characterized by (6), and $\frac{-2p}{q_1^2} z(n, \mathbf{q})$ is the SOD of firm 1's problem in the interior of E . These notations will be used in the following lemmas, propositions and theorems.

LEMMA B.2. *If $\beta > 1$, $n \geq 2$ and $N \geq 1$ then*

- (i) $y(n, \mathbf{q}) \geq 0$ implies $z(n, \mathbf{q}) > 0$, for any $\mathbf{q}$ in the interior of E ; especially, $y(n) \geq 0$ implies $z(n) > 0$;
- (ii) $y(n, \mathbf{q}) \geq 0$ implies $\frac{\partial y(n, \mathbf{q})}{\partial q_i} < 0$, for any $\mathbf{q}$ in the interior of E and for $i = 2, \dots, n$;
- (iii) $y(n) \geq 0$ implies $y(n) \geq y(n+k)$, $\forall n \geq 2, k \geq 1$; $y(n) \geq 0$ implies $y(n) \leq y(n-k)$, $\forall n \geq 3, 1 \leq k \leq n-2$;
- (iv) there exists an integer $\tilde{n}$ such that $\pi(n) \geq 0$ if $n \leq \tilde{n}$, $\pi(n) < 0$ otherwise; $\pi(n)$ is decreasing in n for $n \in \{k : \pi(k) \geq 0\}$.

Proof. Recall that $\Gamma_i(\infty)$ is the sum of infinite exponentials and $f_i(\cdot)$ is its density, for $i = 1, \dots, n$ (See Appendix A). Let $F_i(s)$ be the C.D.F. of $\Gamma_i(\infty)$, $\bar{F}_i(s) = 1 - F_i(s)$, $i = 1, \dots, n$. Denote by $f(\cdot)$ and $F(\cdot)$ as

their P.D.F. and C.D.F. respectively, if all q_i s are mutually equal. Using the same techniques in the proof of Theorem A.5, we have the following results:

$$\begin{aligned}
y(n, \mathbf{q}) &= - \int_0^\infty s_1 f'_1(s_1) ds_1 \int_{s_1}^\infty f_2(s_2) ds_2 \cdots \int_{s_1}^\infty f_n(s) ds \\
&= - \int_0^\infty s_1 f'_1(s) \bar{F}_2(s_1) \cdots \bar{F}_n(s_1) ds_1; \\
\lim_{N \rightarrow \infty} \sum_{j_1=1}^N \cdots \sum_{j_n=1}^N \frac{(q_1 j_1^\beta)^2 q_2 j_2^\beta}{(q_1 j_1^\beta + \cdots + q_n j_n^\beta)^3} \prod_{k=1}^n H_{j_k, N} &= - \frac{1}{2} \int_0^\infty s_1^2 f'_1(s_1) f_2(s_1) ds_1 \int_{s_1}^\infty f_3(s_3) ds_3 \cdots \int_{s_1}^\infty f_n(s_n) ds \\
&= - \frac{1}{2} \int_0^\infty s_1^2 f'_1(s_1) f_2(s_1) \bar{F}_3(s_1) \cdots \bar{F}_n(s_1) ds_1; \\
z(n, \mathbf{q}) &= \lim_{N \rightarrow \infty} \sum_{i=2}^n \sum_{j_1=1}^N \cdots \sum_{j_n=1}^N \frac{(q_1 j_1^\beta)^2 q_i j_i^\beta}{(q_1 j_1^\beta + \cdots + q_n j_n^\beta)^3} \prod_{k=1}^n H_{j_k, N}.
\end{aligned}$$

By the second result above, we have

$$\lim_{N \rightarrow \infty} \sum_{j_1=1}^N \cdots \sum_{j_n=1}^N \frac{(q_1 j_1^\beta)^2 q_2 j_2^\beta}{(q_1 j_1^\beta + \cdots + q_n j_n^\beta)^3} \prod_{k=1}^n H_{j_k, N} = - \frac{1}{2} \int_0^\infty s_1 f'_1(s_1) \bar{F}_2(s_1) \frac{s_1 f_2(s_1)}{\bar{F}_2(s_1)} \bar{F}_3(s_1) \cdots \bar{F}_n(s_1) ds_1.$$

We now prove part (i). Since $f_i(\cdot)$ is logconcave, $f_i(s)$ is first increasing then decreasing on $[0, +\infty)$ and $\frac{s f_i(s)}{\bar{F}_i(s)}$ is increasing. Suppose $f'_1(s)$ is positive on $(0, \epsilon)$ and negative on $(\epsilon, +\infty)$. Thus $y(n, \mathbf{q}) \geq 0$ implies

$$\begin{aligned}
- \int_0^\infty s_1 f'_1(s_1) \frac{s_1 f_2(s_1)}{\bar{F}_2(s_1)} \bar{F}_2(s_1) \cdots \bar{F}_n(s_1) ds_1 &= - \frac{\theta_1 f_2(\theta_1)}{\bar{F}_2(\theta_1)} \int_0^\epsilon s_1 f'_1(s_1) \bar{F}_2(s_1) \cdots \bar{F}_n(s_1) ds_1 \\
&\quad - \frac{\theta_2 f_2(\theta_2)}{\bar{F}_2(\theta_2)} \int_\epsilon^{+\infty} s_1 f'_1(s_1) \bar{F}_2(s_1) \cdots \bar{F}_n(s_1) ds_1 > 0,
\end{aligned}$$

because $\theta_1 < \theta_2$. By this way, we can prove that $y(n, \mathbf{q}) \geq 0$ implies $z(n, \mathbf{q}) > 0$.

We next prove part (ii). Using the result of part (i), we have

$$\frac{\partial y(n, \mathbf{q})}{\partial q_2} = \lim_{N \rightarrow \infty} \frac{-2}{q_2} \sum_{j_1=1}^N \cdots \sum_{j_n=1}^N \frac{(q_1 j_1^\beta)^2 q_2 j_2^\beta}{(q_1 j_1^\beta + \cdots + q_n j_n^\beta)^3} \prod_{k=1}^n H_{j_k, N} < 0.$$

We next prove part (iii). Set $q_1 = \cdots = q_n > 0$ then for $k_1, k_2 \geq 1$ and $n - k_2 \geq 2$ we have

$$\begin{aligned}
y(n) &= - \int_0^\infty s f'(s) \bar{F}(s)^{n-1} ds, \\
y(n + k_1) &= - \int_0^\infty s f'(s) \bar{F}(s)^{n+k_1-1} ds \\
y(n - k_2) &= - \int_0^\infty s f'(s) \bar{F}(s)^{n-k_2-1} ds.
\end{aligned}$$

Suppose $f'(s)$ is positive on $(0, \epsilon)$ and negative on $(\epsilon, +\infty)$. Thus $y(n) \geq 0$ implies

$$\begin{aligned}
y(n + k_1) - y(n) &= F(\theta_1)^{k_1} \int_0^\epsilon s f'(s) \bar{F}(s)^{n-1} ds + F(\theta_2)^{k_1} \int_\epsilon^\infty s f'(s) \bar{F}(s)^{n-1} ds < 0; \\
y(n) - y(n - k_2) &= \left(\frac{1}{\bar{F}(\theta_1)^{k_2}} - 1 \right) \int_0^\epsilon s f'(s) \bar{F}(s)^{n-1} ds + \left(\frac{1}{\bar{F}(\theta_2)^{k_2}} - 1 \right) \int_\epsilon^\infty s f'(s) \bar{F}(s)^{n-1} ds < 0;
\end{aligned}$$

Hence this part is proved.

Part (iv) can be inferred from part (iii). Therefore the proof of this lemma is complete. $\square$

Proof of Theorem 3. In this proof, shortened forms of notation have the same meaning in the proof of Theorem 2.

For $\beta > 1$, $\pi_i = p \cdot \mathbb{P}\{\mathbf{x}^* = \mathbf{e}_i\} - cq_i$. In the interior of E , the FOD of firm i 's problem is

$$\frac{\partial \pi_i}{\partial q_i} = \lim_{N \rightarrow \infty} \frac{p}{q_i} \sum_{j_1=1}^N \cdots \sum_{j_n=1}^N \frac{q_i j_i^\beta \sum_{k \neq i} q_k j_k^\beta}{(q_1 j_1^\beta + \cdots + q_n j_n^\beta)^2} \prod_{k=1}^n H_{j_k, N} - c,$$

and the SOD is

$$\frac{\partial^2 \pi_i}{\partial q_i^2} = \lim_{N \rightarrow \infty} \frac{-2p}{q_i^2} \sum_{j_1=1}^N \cdots \sum_{j_n=1}^N \frac{(q_i j_i^\beta)^2 \sum_{k \neq i} q_k j_k^\beta}{(q_1 j_1^\beta + \cdots + q_n j_n^\beta)^3} \prod_{k=1}^n H_{j_k, N}.$$

Note that $\frac{q_i j_i^\beta \sum_{k \neq i} q_k j_k^\beta}{(q_1 j_1^\beta + \cdots + q_n j_n^\beta)^2} = \frac{q_i j_i^\beta}{q_1 j_1^\beta + \cdots + q_n j_n^\beta} - \frac{(q_i j_i^\beta)^2}{(q_1 j_1^\beta + \cdots + q_n j_n^\beta)^2}$. Hence we can see from the FODs that if there exists a symmetric equilibrium then it must be

$$\mathbf{q} = (1, \dots, 1) \cdot \frac{p}{c} \left(\frac{1}{n} - \lim_{N \rightarrow \infty} \sum_{j_1=1}^N \cdots \sum_{j_n=1}^N \frac{(j_1^\beta)^2}{(j_1^\beta + \cdots + j_n^\beta)^2} \prod_{k=1}^n H_{j_k, N} \right).$$

Thus $\mathbf{q}$ is a symmetric equilibrium of $\mathcal{G}$ if and only if it satisfies $\pi_i|_{\mathbf{q}} \geq 0$ and $\frac{\partial^2 \pi_i}{\partial q_i^2}|_{\mathbf{q}} \leq 0$, which are equivalent to $\pi(n) \geq 0$ and $z(n) \geq 0$ (where $\pi(n)$, $z(n)$ are defined in Lemma B.2). By Lemma B.2(i), we have that $\pi(n) \geq 0$ implies $z(n) \geq 0$. Define $m(\beta, K) = \max\{k : \pi(k) \geq K, z(k) \geq 0\}$. Since $\pi(k) = py(k)$ in any symmetric equilibrium, we have that $m(\beta, K) = \max\{k : \pi(k) \geq K\} = \max\{k : y(k) \geq K/p\}$. Thus we have that, if $2 \leq n \leq m(\beta, 0)$, $\mathbf{q}$ must be an equilibrium of $\mathcal{G}$, and $\pi^*(n, \beta) \geq K$ if and only if $n \leq m(\beta, K)$; if $n > m(\beta, 0)$, then there is no symmetric pure-strategy equilibrium.

We next prove in any symmetric mixed-strategy equilibrium, each of the n firms must get an expected profit of 0 if $n > m(\beta, 0)$. Suppose the n firms randomize their qualities according an independent and identical distribution, $G(q_i)$, $i = 1, \dots, n$ with support R . As the same argument in the proof of Theorem 2, we assume $R = [\underline{q}, \bar{q}]$ and we will show that $\underline{q}$ must be zero. By the definition of a mixed equilibrium, we have

$$p \int_{\underline{q}}^{\bar{q}} \cdots \int_{\underline{q}}^{\bar{q}} \lim_{N \rightarrow \infty} \sum_{j_1, \dots, j_n} \frac{q_1 j_1^\beta}{q_1 j_1^\beta + \cdots + q_n j_n^\beta} \prod_{k=1}^n H_{j_k, N} dG(q_2) \cdots dG(q_n) = cq_1 + a_1, \quad \forall q_1 \in R, \quad (14)$$

where $a_1 > 0$, a constant, is the expected profit of firm 1.

By Proposition 1, we have that if $y(n, \mathbf{q}) \geq 0$ at the point $(\bar{q}, \dots, \bar{q})$ then $y(n, \mathbf{q}) \geq 0$ for $q_1 = \bar{q}$ and $\forall \mathbf{q}_{-1} \in R^{n-1}$. By Lemma B.2(ii) we have that if $y(n, \mathbf{q}) \geq 0$ at the point $(\bar{q}, \dots, \bar{q})$ then the left side of Equation (14) is strictly concave at $q_1 = \bar{q}$, which is impossible because the right side of Equation (14) is linear in q_1 . Hence $y(n, \mathbf{q}) < 0$ at the point $(\bar{q}, \dots, \bar{q})$, which implies that $y(n, \mathbf{q}) < 0$ at the point $(\underline{q}, \dots, \underline{q})$

if $\underline{q} > 0$. Also by Lemma B.2(ii) we have that $y(n, \mathbf{q}) < 0$ for $q_1 = \underline{q}$ and $\mathbf{q}_{-1} \in R^{n-1}$. Taking the first-order derivative on the both sides of Equation (14) at $q_1 = \underline{q}$, we have

$$\begin{aligned} c &= \frac{p}{\underline{q}} \int_{\underline{q}}^{\bar{q}} \cdots \int_{\underline{q}}^{\bar{q}} \lim_{N \rightarrow \infty} \sum_{j_1, \dots, j_n} \left[\frac{\underline{q} j_1^\beta}{\underline{q} j_1^\beta + \cdots + q_n j_n^\beta} - \frac{(\underline{q} j_1^\beta)^2}{(\underline{q} j_1^\beta + \cdots + q_n j_n^\beta)^2} \right] \prod_{k=1}^n H_{j_k, N} dG(q_2) \cdots dG(q_n) \\ &> \frac{p}{\underline{q}} \int_{\underline{q}}^{\bar{q}} \cdots \int_{\underline{q}}^{\bar{q}} \lim_{N \rightarrow \infty} \sum_{j_1, \dots, j_n} \frac{\underline{q} j_1^\beta}{\underline{q} j_1^\beta + \cdots + q_n j_n^\beta} \prod_{k=1}^n H_{j_k, N} dG(q_2) \cdots dG(q_n) \\ &= c + \frac{a_1}{\underline{q}}, \end{aligned}$$

where the last equality hold by Equation (14). That forms a contradiction. Therefore we must have $\underline{q} = 0$ and the proof is complete. $\square$

Proof of Lemma 2. See Lemma A.2(ii). $\square$

The proof of Proposition 2 makes use of the following lemma.

LEMMA B.3. *For any $C > 0$, there exists $\hat{\beta}(C)$ such that $|H_k| > C|H_{k+1}|$, for $k = 1, 2, \dots$, if $\beta > \hat{\beta}(C)$.*

Proof. We can see from Lemma A.2(ii) that, for $\beta > 2$, $\{|H_k|\}_{k=1}^\infty$ is bounded and hence $\{\frac{|H_{k+1}|}{|H_k|}\}_{k=1}^\infty$ is bounded. For $k = 1, 2, \dots$, $\frac{H_{k+1}}{H_k} = - \prod_{\substack{l \neq k \\ l \neq k+1}} \frac{l^\beta - k^\beta}{l^\beta - (k+1)^\beta} \frac{k^\beta}{(k+1)^\beta}$. Since $\forall l \neq k, l \neq k+1$, $\frac{l^\beta - k^\beta}{l^\beta - (k+1)^\beta}$ is decreasing in β and $\frac{k^\beta}{(k+1)^\beta}$ is decreasing in β , we have that $\frac{|H_{k+1}|}{|H_k|}$ converges to 0 as $\beta \rightarrow \infty$. Also by Lemma A.2(ii), we have that for $k = 1, 2, \dots$, $\frac{|H_{k+1}|}{|H_k|}$ uniformly converges to 0 as $\beta \rightarrow \infty$. Thus for any $C > 0$, there exists $\hat{\beta}(C)$ such that $|H_k| > C|H_{k+1}|$, for $k = 1, 2, \dots$, if $\beta > \hat{\beta}(C)$. $\square$

LEMMA B.4. *For $n = 1, 2, \dots$, there exists a $\tilde{\beta}$ (where $\tilde{\beta}$ is independent of n) such that $\lim_{N \rightarrow \infty} \sum_{j_1=1}^N \sum_{j_2=1}^N \frac{j_1^\beta j_2^\beta}{(j_1^\beta + j_2^\beta + n)^2} H_{j_1, N} H_{j_2, N}$ is decreasing in β for $\beta \in (\tilde{\beta}, +\infty)$.*

Proof. For $\beta > 2$, we can see from Lemma A.2(ii) that

$$\lim_{N \rightarrow \infty} \sum_{j_1=1}^N \sum_{j_2=1}^N \frac{j_1^\beta j_2^\beta}{(j_1^\beta + j_2^\beta + n)^2} H_{j_1, N} H_{j_2, N} = \sum_{j_1=1}^\infty \sum_{j_2=1}^\infty \frac{j_1^\beta j_2^\beta}{(j_1^\beta + j_2^\beta + n)^2} H_{j_1} H_{j_2}.$$

We can first get that

$$\begin{aligned} \frac{\partial}{\partial \beta} \frac{j_1^\beta j_2^\beta}{(j_1^\beta + j_2^\beta + n)^2} &= - \frac{j_1^\beta j_2^\beta}{(j_1^\beta + j_2^\beta + n)^2} \frac{(\ln j_1 - \ln j_2)(j_1^\beta - j_2^\beta) - n(\ln j_1 + \ln j_2)}{j_1^\beta + j_2^\beta + n}, \\ \frac{\partial}{\partial \beta} H_{j_1} &= -H_{j_1} \sum_{l \neq j_1} \frac{\ln l - \ln j_1}{l^\beta - j_1^\beta} j_1^\beta, \\ \frac{\partial}{\partial \beta} H_{j_2} &= -H_{j_2} \sum_{l \neq j_2} \frac{\ln l - \ln j_2}{l^\beta - j_2^\beta} j_2^\beta. \end{aligned}$$

Define

$$\psi(j_1, j_2, n) = - \frac{j_1^\beta j_2^\beta}{(j_1^\beta + j_2^\beta + n)^2} \left[\frac{(\ln j_1 - \ln j_2)(j_1^\beta - j_2^\beta) - n(\ln j_1 + \ln j_2)}{j_1^\beta + j_2^\beta + n} + \sum_{l \neq j_1} \frac{\ln l - \ln j_1}{l^\beta - j_1^\beta} j_1^\beta + \sum_{l \neq j_2} \frac{\ln l - \ln j_2}{l^\beta - j_2^\beta} j_2^\beta \right].$$

Obviously $\psi(j_1, j_2, n)$ is $j_1 - j_2$ symmetric. We have that

$$\frac{\partial}{\partial \beta} \sum_{j_1=1}^{\infty} \sum_{j_2=1}^{\infty} \frac{j_1^{\beta} j_2^{\beta}}{(j_1^{\beta} + j_2^{\beta} + n)^2} H_{j_1} H_{j_2} = \sum_{j_1=1}^{\infty} \sum_{j_2=1}^{\infty} \psi(j_1, j_2, n) H_{j_1} H_{j_2}.$$

Since $\sum_{l \neq j_1} \frac{\ln l - \ln j_1}{l^{\beta} - j_1^{\beta}} j_1^{\beta} > \ln j_1$, $\forall j_1 \geq 1$, we know from the definition of $\psi(j_1, j_2, n)$ that $\psi(j_1, j_2, n) < 0$, $\forall j_1 \geq 1, j_2 \geq 1, n \geq 0, \beta \geq 2$. We can also see that there exists a constant C such that $|\frac{\psi(j_1+1, j_2, n)}{\psi(j_1, j_2, n)}| < C$, $\forall j_1 \geq 1, j_2 \geq 1, n \geq 0, \beta \geq 2$.

Notice that $\{H_{j_1}\}$ is a infinite sequence with alternating signs. We have that

$$\begin{aligned} \sum_{j_1=1}^{\infty} \sum_{j_2=1}^{\infty} \psi(j_1, j_2, n) H_{j_1} H_{j_2} &= \sum_{k=1}^{\infty} \left(\psi(k, k, n) H_k^2 + 2 \sum_{j_1=k+1}^{\infty} \psi(j_1, k, n) H_{j_1} H_k \right) \\ &= \sum_{k=1}^{\infty} \left(\psi(k, k, n) H_k + 2\psi(k+1, k, n) H_{k+1} + 2 \sum_{m=1}^{\infty} \psi(k+2m, k, n) H_{k+2m} + \psi(k+2m+1, k, n) H_{k+2m+1} \right) H_k. \end{aligned}$$

By Lemma B.3, there exists a $\tilde{\beta}$ such that if $\beta > \tilde{\beta}$ then

$$\psi(k, k, n) H_k^2 + 2\psi(k+1, k, n) H_{k+1} H_k < 0,$$

$$\psi(k+2m, k, n) H_{k+2m} H_k + \psi(k+2m+1, k, n) H_{k+2m+1} H_k < 0, \quad \forall m \in \mathbb{N}.$$

Therefore $\frac{\partial}{\partial \beta} \sum_{j_1=1}^{\infty} \sum_{j_2=1}^{\infty} \psi(j_1, j_2, n) H_{j_1} H_{j_2} < 0$ if $\beta > \max\{\tilde{\beta}, 2\}$. This completes the proof. $\square$

Proof of Proposition 2. We first prove part (i). We have known that $\pi^*(n, \beta) = p_n^{\frac{1}{n}} - cq^*(n, \beta)$ is the expected second-stage profit for each firm in the symmetric equilibrium (if it exists). From Equation (6) we can see that $\pi^*(n, \beta) \geq 0$ if and only if $\lim_{N \rightarrow \infty} \sum_{1 \leq j_1, \dots, j_N \leq N} \frac{(j_1^{\beta})^2}{(j_1^{\beta} + \dots + j_N^{\beta})^2} \prod_{k=1}^n H_{j_k, N} \geq 0$, which is equivalent to $n \leq m(\beta, 0)$. Then from Lemma B.2(iv) we see that $\pi^*(n, \beta)$ is decreasing in n if $n \leq m(\beta, 0)$.

We next prove $q^*(n+1, \beta) \leq q^*(n, \beta)$ if $n < m(\beta, 0)$. We consider firm 1 as a focal firm. For $\beta \leq 1$, by Theorem 1(i), we know that $\frac{\partial^2 \pi_1}{\partial q_1 \partial q_2} < 0$ if $q_1 < q_2$. We now prove $\frac{\partial^2 \pi_1}{\partial q_1 \partial q_2} < 0$ if $q_1 < q_2$ for $\beta > 1$. Using the same technique in the proof of Theorem A.5, if $q_1 < q_2$, we have

$$\begin{aligned} \frac{\partial^2 \pi_1}{\partial q_1 \partial q_2} &= \frac{1}{q_1 q_2} \lim_{N \rightarrow \infty} \sum_{j_1, \dots, j_N=1}^N \frac{q_1 j_1^{\beta} q_2 j_2^{\beta} (q_1 j_1^{\beta} - q_2 j_2^{\beta})}{(q_1 j_1^{\beta} + \dots + q_n j_n^{\beta})^3} - \frac{q_1 j_1^{\beta} q_2 j_2^{\beta} (q_3 j_3^{\beta} + \dots + q_n j_n^{\beta})}{(q_1 j_1^{\beta} + \dots + q_n j_n^{\beta})^3} \prod_{k=1}^n H_{j_k, N} \\ &< \frac{1}{q_1 q_2} \lim_{N \rightarrow \infty} \sum_{j_1, \dots, j_N=1}^N \frac{q_1 j_1^{\beta} q_2 j_2^{\beta} (q_1 j_1^{\beta} - q_2 j_2^{\beta})}{(q_1 j_1^{\beta} + \dots + q_n j_n^{\beta})^3} \prod_{k=1}^n H_{j_k, N} \\ &= \frac{1}{2q_1 q_2} \int_0^{\infty} s^2 [f_1(s) f_2'(s) - f_1'(s) f_2(s)] \bar{F}_3(s) \dots \bar{F}_n(s) ds \\ &= \frac{1}{2q_1 q_2} \int_0^{\infty} s^2 f_1(s)^2 \left[\frac{f_2(s)}{f_1(s)} \right]' \bar{F}_3(s) \dots \bar{F}_n(s) ds \\ &< 0 \end{aligned}$$

where the first inequality uses the fact analogous to Lemma B.1 and the last inequality holds due to the decreasing likelihood ration $\frac{f_2(s)}{f_1(s)}$. Thus the marginal revenue of improving the quality for firm 1 is decreasing in product 2's quality if product 2's quality is higher than product 1's quality. In the symmetric equilibrium with $n + 1$ firms, the equilibrium quality is $q^*(n + 1, \beta)$. At the point $\mathbf{q} = \mathbf{1} \cdot q^*(n + 1, \beta)$, if the firm $n + 1$ exits, namely q_{n+1} decreases from $q^*(n + 1, \beta)$ to zero, then any other firm has an incentive to improve its quality due to the analysis above. Therefore, in the equilibrium with n firms, the equilibrium quality must be greater than or equal to that in the equilibrium with $n + 1$ firms, i.e., $q^*(n, \beta) \geq q^*(n + 1, \beta)$.

Part (ii) is clear by following Theorem 2(i). We next prove part (iii). From Equation (6), we find that it suffices to prove that there exists a $\tilde{\beta} > 1$ such that $y(n) = \lim_{N \rightarrow \infty} \sum_{j_1=1}^N \cdots \sum_{j_n=1}^N \frac{(j_1^\beta)^2}{(j_1^\beta + \cdots + j_n^\beta)^2} \prod_{k=1}^n H_{j_k, N}$ is increasing in β for $\beta > \tilde{\beta}$. Since

$$1 = \sum_{j_1=1}^N \cdots \sum_{j_n=1}^N \frac{(j_1^\beta + \cdots + j_n^\beta)^2}{(j_1^\beta + \cdots + j_n^\beta)^2} \prod_{k=1}^n H_{j_k, N}$$

by Lemma A.2(i), we have

$$n \sum_{j_1=1}^N \cdots \sum_{j_n=1}^N \frac{(j_1^\beta)^2}{(j_1^\beta + \cdots + j_n^\beta)^2} \prod_{k=1}^n H_{j_k, N} + n(n-1) \sum_{j_1=1}^N \cdots \sum_{j_n=1}^N \frac{j_1^\beta j_2^\beta}{(j_1^\beta + \cdots + j_n^\beta)^2} \prod_{k=1}^n H_{j_k, N} = 1. \quad (15)$$

Thus it suffices to prove that there exists a $\tilde{\beta} > 1$ such that $\lim_{N \rightarrow \infty} \sum_{j_1=1}^N \cdots \sum_{j_n=1}^N \frac{j_1^\beta j_2^\beta}{(j_1^\beta + \cdots + j_n^\beta)^2} \prod_{k=1}^n H_{j_k, N}$ is decreasing in β for $\beta > \tilde{\beta}$. Using the techniques in the proof of Theorem A.5, we have

$$\lim_{N \rightarrow \infty} \sum_{j_1=1}^N \cdots \sum_{j_n=1}^N \frac{j_1^\beta j_2^\beta}{(j_1^\beta + \cdots + j_n^\beta)^2} \prod_{k=1}^n H_{j_k, N} = \int_0^\infty s f(s)^2 \bar{F}(s)^{n-2} ds.$$

Before going on the proof we first presents some properties of $f(\cdot)$ and $\bar{F}(\cdot)$. As β increases from $\beta_1 > 1$ to $\beta_2 > 1$, the corresponding P.D.F. varies from $f(\beta_1, s)$ to $f(\beta_2, s)$. By Lemma 1.1 in Shanthikumar and Yao (1986) we know $\frac{f(\beta_2, s)}{f(\beta_1, s)}$ is decreasing in s . By the logconcavity of $f(s)$ we have that for $\beta > 1$, $\frac{\partial}{\partial \beta} f(s)$ is first positive then negative for $s \in [0, +\infty)$. We assume $\frac{\partial}{\partial \beta} f(s)$ is positive for $s \in (0, \epsilon)$ and negative for $s \in (\epsilon, +\infty)$. Recall that $\bar{F}(s)$ is the tail distribution of the convolution of exponentials with parameters $1, 2^\beta, 3^\beta, \dots$. As $\beta \rightarrow \infty$, $f(s) \rightarrow e^{-s}$, $\bar{F}(s) \rightarrow e^{-s}$. Since $f(\beta_1, s)$ dominates $f(\beta_2, s)$ in the likelihood ratio ordering for $\beta_1 < \beta_2$, we know $\frac{f(s)}{\bar{F}(s)}$ is increasing in β and $\bar{F}(s)$ decreasing in β , $\forall s \geq 0$. Thus we have $\frac{f(s)}{\bar{F}(s)} \leq \frac{e^{-s}}{e^{-s}} = 1$, $\forall s \geq 0$. Since $\frac{\partial}{\partial s} \ln \frac{\bar{F}(s)}{e^{-s}} = 1 - \frac{f(s)}{\bar{F}(s)} \geq 0$, we know $\frac{\bar{F}(s)}{e^{-s}}$ is increasing in s .

Consider that

$$\frac{\partial}{\partial \beta} \int_0^\infty s f(s)^2 \bar{F}(s)^{n-2} ds = \frac{\partial}{\partial \beta} \int_0^\infty s f(s)^2 e^{-(n-2)s} \left(\frac{\bar{F}(s)}{e^{-s}} \right)^{n-2} ds$$

$$\begin{aligned}
&= \int_0^\infty \left(\frac{\partial}{\partial \beta} s f(s)^2 e^{-(n-2)s} \right) \left(\frac{\bar{F}(s)}{e^{-s}} \right)^{n-2} ds + \int_0^\infty s f(s)^2 e^{-(n-2)s} \frac{\partial}{\partial \beta} \left(\frac{\bar{F}(s)}{e^{-s}} \right)^{n-2} ds \\
&= \left(\frac{\bar{F}(\theta_1)}{e^{-\theta_1}} \right)^{n-2} \int_0^\epsilon \frac{\partial}{\partial \beta} s f(s)^2 e^{-(n-2)s} ds + \left(\frac{\bar{F}(\theta_2)}{e^{-\theta_2}} \right)^{n-2} \int_\epsilon^\infty \frac{\partial}{\partial \beta} s f(s)^2 e^{-(n-2)s} ds \\
&\quad + \int_0^\infty s f(s)^2 e^{-(n-2)s} \frac{\partial}{\partial \beta} \left(\frac{\bar{F}(s)}{e^{-s}} \right)^{n-2} ds.
\end{aligned}$$

According to the properties of $f(\cdot)$ and $\bar{F}(\cdot)$, we can see that it suffices to prove $\frac{\partial}{\partial \beta} \int_0^\infty s f(s)^2 e^{-(n-2)s} ds < 0$, which is exactly the observation in Lemma B.4. This completes the proof. $\square$

Proof of Theorem 4. We first prove part (i). We have known that $m(\beta, K) = \lfloor \frac{p}{K+(p-K)\beta} \rfloor$ if $0 \leq \beta \leq 1$; $m(\beta, K) = \max\{n : y(n) \geq K/p\}$ if $\beta > 1$. From Lemma B.1 and Equation (15), we know $y(n) \leq 1$. From Lemma B.2(iii), we also know $y(n)$ is decreasing in n if $y(n) \geq 0$. Thus we conclude that $m(\beta, K) \geq 1$ if and only if $p \geq K$, and that $m(\beta, K)$ is decreasing in K .

Part (ii) comes from the definition of $m(\beta, K)$ in Theorem 3. Next we will prove part(iii). In the proof of Proposition 2(iii), we have shown that $y(n)$ is increasing in β , $\forall n \geq 1$, if $\beta > \tilde{\beta}$. It follows that $m(\beta, K)$ is increasing in β for $\beta > \tilde{\beta}$. $\square$

Proof of Proposition 3. If there are $m(\beta, K)$ firms entering the market, then in the symmetric equilibrium, $q^*(\beta) = q^*(m(\beta, K), \beta)$. We first prove part (i). If $0 \leq \beta < 1$, we can see from Theorem 2 that $q^*(\beta) = \frac{\lfloor \frac{p}{K+(p-K)\beta} \rfloor - 1}{\lfloor \frac{p}{K+(p-K)\beta} \rfloor^2} \frac{p}{c(1-\beta)}$. It is clear that $q^*(\beta)$ is increasing in β in this case.

Then we will prove part (ii). We have shown in Proposition 2 that $q^*(n, \beta) \geq q^*(n+1, \beta)$ if $n+1 \leq m(\beta, 0)$ for any $\beta \geq 0$, and that $q^*(n, \beta)$ is decreasing in β for $\beta > \tilde{\beta}$, $\forall n \leq m(\beta, 0)$. We also shown in Theorem 4(iii) that $m(\beta, K)$ is increasing in β for $\beta > \tilde{\beta}$. Based on these results, part (ii) holds naturally. $\square$