Retail Clusters in Developing Economies

Xuying Zhao

University of Notre Dame, xzhao1@nd.edu

Arthur Lim

University of Notre Dame, arthurlim@nd.edu

Hong Guo

University of Notre Dame, hguo@nd.edu

Chao Ding

The University of Hong Kong, chao.ding@hku.hk

Jing-Sheng Song

Duke University, jingsheng.song@duke.edu,

We develop a game-theoretic model to explore why retail clusters are so popular in developing economies and when governments should facilitate the formation of retail clusters to improve social welfare. First, we find two determinants of retailer clusters: valuation-cost ratio (consumers' maximum valuation over retailers' production cost) and retailer density (the number of retailers over unit transportation cost). These two determinants indicate retailers' profit potential and competition intensity, respectively. Second, the equilibrium cluster size increases in the valuation-cost ratio. This finding explains the phenomenon that clusters are usually larger in developing economies (where numerous retailers sell unrecognized brands with low profit potential) than in developed economies. Third, when the retailer density of a product market exceeds a certain threshold, the market coverages of clusters overlap with each other (i.e., the overlapping case). Furthermore, when compared to the non-overlapping case, the equilibrium cluster size in the overlapping case is larger for low-profit-potential products but smaller for high-profit-potential products. Together, valuation-cost ratio and retailer density define four types of clusters: overlapping massive clusters, non-overlapping large clusters, non-overlapping small clusters, and overlapping mini clusters. Finally, the socially optimal cluster size is larger than the equilibrium cluster size, and the gap between these two cluster sizes decreases in the valuation-cost ratio.

Key words: retail clusters; developing economies; valuation uncertainty; valuation-cost ratio; retailer density; transportation cost

History:

1. Introduction

A retail cluster refers to a collection of retailers of a particular business sector that are located in close proximity. Retail clusters commonly exist in developing economies such as Africa and Asia (Dawson et al. 2005). For example, retailer clusters exist in Jaipur, India (jewelry); Deira, Dubai (apparel); Urumuqi, China (handicrafts); and so forth. In fact, Shenzhen, China is home to the world's largest electronics retail cluster, featuring hundreds of retailers that supply a full spectrum of electronic products, parts, peripherals, and accessories. In this commercial district, the majority of retailers in the cluster are small-sized sellers who carry products made by unknown small manufacturers.

In this paper, we develop a game-theoretic model framework to explore why retail clusters are so popular in developing economies and how the governments can facilitate the formation of retail clusters to improve social welfare. Our model framework consists of a set of competing retailers offering products in the same category to a heterogeneous consumer population. Consumers differ in their physical locations on the Salop circle (Salop 1979) as well as in their valuations for the product. Retailers strategically decide their store locations on the Salop circle and their prices to attract consumers. We can then use the resulting equilibrium store locations of retailers to understand the conditions under which voluntary clusters form. Comparing this profit maximum equilibrium with the social welfare maximum equilibrium, we provide directions for government intervention.

To characterize the players and their utilities in our game setting, we recognize several intertwining features of the retailing landscape in developing economies compared to those of developed economies. First, in developing economies, there is a lack of a credible information sharing platform for consumers to learn about the quality of a product and whether a particular product would fit their needs, especially with respect to unrecognized brands. Consumers need to visit stores to find out how much they value the product. For this reason, we explicitly model consumers' valuation uncertainty before visiting a physical store.

Second, consumers' maximum valuation differs in developing and developed economies across different product types. In developed economies, consumers are generally wealthier and retailers usually carry products with well-known brands, whereas in developing economies, the consumer market is largely shared by small local retailers who sell large numbers of unrecognized brands, especially in suburban and rural regions. Consumers' maximum valuation for unrecognized brands can be much lower than well-known brands because of the unknown product quality and customer service levels.

Third, unit transportation cost also differs in developing and developed economies across different product types. Consumers in developing economies rely largely on underdeveloped public

transportation systems, which may require them to combine multiple transportation modes for a single trip rather than use private vehicles or streamlined public transportation systems. Not only do consumers consider tangible costs like fuel price and fare costs, but they also consider travel convenience, as reflected in travel time and travel frequency. We collectively define these concerns as transportation cost, and include it as an important component in consumers' utility function.

Fourth, the number of retailers differs in developing and developed economies across different product types. The number may be much higher in developing economies for products with low entry costs (e.g., small commodities). However, the statement may be reversed for products with high entry costs (e.g., imported cars and luxury brands).

Fifth, retailers' production cost may also be different in developing and developed economies across different product types. For labor-intensive products, the production cost in developing economies is lower due to lower labor cost. For technology-intensive products, the product cost in developing economies is higher due to lagging technological development.

To capture the above key characteristics identified for developing economies, we extend the model in Konishi (2005) to consider a general consumers' maximum valuation, transportation cost, number of retailers, and retailers' production cost in the model. We assume that consumers are heterogeneous in their physical locations and realized valuations after visiting stores. Before they visit a retailer, consumers are uncertain about product valuation. Consumers estimate their valuation, transportation cost, and price to decide which location to visit. Based on this model setting, we analyze the equilibrium of the game to gain insights and we summarize our key results below.

First, we find two important determinants of retail clusters: valuation-cost ratio and retailer density. Valuation-cost ratio refers to consumers' maximum valuation over retailers' production cost. This ratio indicates the profit potential for a retailer. The higher the valuation-cost ratio, the higher the profit potential. Retailer density, meanwhile, refers to the ratio of the number of retailers over unit transportation cost. Retailer density indicates the degree of competition intensity within a product market. The higher the retailer density, the more intensive the competition.

Second, we find that the equilibrium cluster size decreases in the valuation-cost ratio (i.e., decreases in profit potential). When the size of a cluster increases, more retailers compete with each other within the cluster. Consequently, the retail price goes down, driven by the price-cutting effect within the cluster. When there are more retailers clustering together, there is also a market-size effect, which attracts more customers to visit the cluster. When the valuation-cost ratio is low

(i.e., profit potential is low), the market-size effect dominates the price-cutting effect.[1] Thus, the equilibrium cluster size decreases in the valuation-cost ratio.

This finding explains the phenomenon that clusters in developing economies are usually larger compared to those in developed economies. For instance, there are many apparel clusters in both developing and developed economies. Whereas the cluster size in developing economies is usually in the thousands or even tens of thousands, the cluster size in developed economies is in the tens or hundreds. One big difference among these apparel clusters is the profit potential: the majority of apparel retailers in clusters in developing economies sell clothes from unrecognized brands at cheap prices, while most clothes sold in clusters in developed economies are usually from well-known brands. Thus, the profit potential is much higher in developed economies than in developing economies. Since the equilibrium cluster size decreases in profit potential, the cluster sizes in developing economies are larger than those in developed economies.

Third, we find that when retailer density exceeds a certain threshold, the market coverages of clusters overlap with each other, which is referred to as the overlapping case or the overlapping clustering equilibrium. In other words, when retailer density is high, competition is so intense that the consumer market is fully covered, and clusters compete with each other for market size. When compared to developed economies, retailer density in developing economies is higher for low-entry-cost products, but lower for high-entry-cost products. Consequently, the overlapping clustering equilibrium is more commonly seen in developing economies for low-entry-cost products (e.g., small commodities), while the non-overlapping clustering equilibrium is more commonly seen in developing economies for high-entry-cost products (e.g., automobiles).

Fourth, we find that when compared to the non-overlapping case, the equilibrium cluster size in the overlapping case is larger for low-profit-potential products, but smaller for high-profit-potential products. The implications of this finding are the opposite for developing and developed economies. For developing economies, the majority of products have low profit potential when compared to those in developed economies. Thus, the equilibrium cluster size is larger for high-retailer-density products than low-retailer-density products. For developed economies, this finding is reversed, i.e., the equilibrium cluster size is smaller for high-retailer-density products than for lower-retailer-density products.

Together, the valuation-cost ratio and retailer density define four types of clusters: overlapping massive clusters, non-overlapping large clusters, non-overlapping small clusters, and overlapping

---

[1] In this paper, the market-size effect and the price-cutting effect are the main drivers of our equilibrium results. We show that the two identified ratios (valuation-cost ratio $\frac{v}{c}$ and retailer density $\frac{n}{t}$) have important moderating effects on the two main drivers.

mini clusters. The first two are equilibrium clusters when the valuation-cost ratio is low (i.e., the profit potential is low). The last two occur when the valuation-cost ratio is high (i.e., the profit potential is high).

Large or massive clusters are more likely to be observed in developing economies because of low profit potential in developing economies. Furthermore, overlapping massive clusters are the equilibrium results for products with a low entry cost or low transportation cost (e.g., small commodities) such that retailer density is high. For example, the small commodities cluster in Yiwu, China and the electronics cluster in Shenzhen, China are both overlapping massive clusters with thousands of retailers. Non-overlapping large clusters are the equilibrium results for products with a high entry cost or high transportation cost (e.g., furniture) such that retailer density is low. For instance, in almost every major city in China, several furniture clusters exist that serve local customers only. Usually, these furniture clusters are non-overlapping large clusters with hundreds of retailers.

In developed economies, the retail cluster size is larger for low-retailer-density products than for high-retailer-density products. Overlapping mini clusters are the equilibrium results for products with a high retailer density, while non-overlapping small clusters are for products with a low retailer density. For example, stores selling high-retailer-density products, such as pharmaceutical retailers CVS and Walgreens, usually choose to locate close to customers rather than cluster together at a central location. Pharmaceutical clusters are very small with either a single retailer or two retailers at the same location. On the other hand, stores selling low-retailer-density products, such as automobiles, usually form a small cluster.

Finally, we explore how governments can foster retail clusters and leverage them to improve social welfare. We find that the socially optimal cluster size is larger than the equilibrium cluster size, and that the gap between these two cluster sizes decreases in the valuation-cost ratio. This finding suggests that governments should encourage retailers to form clusters. Since the valuation-cost ratio is usually lower in developing economies, the benefit from government intervention in developing economies is more than that in developed economies. In other words, governments in developing economies should be more aggressive in encouraging clustering – providing various incentives or subsidies, such as reduced rents, participation grants, tax benefits, and so forth – than those in developed economies.

The remainder of this paper is organized as follows. In the next section, we review relevant prior studies and discuss the contributions of this paper to the existing literature. We introduce the model setup in Section 3 and present our equilibrium analyses and results in Sections 4 and 5. In Section 6, we then examine retail clusters from the social planner's perspective and discuss policy implications. In Section 7, we extend the model to endogenize the entry of retailers with an entry cost. Finally, in Section 8, we conclude with a discussion of our findings and possible future research directions.

2. Literature Review

The concept of retail clusters is rooted in more general cluster theory. Porter (1998, 2000) defines clusters as "geographic concentrations of interconnected companies and institutions in a particular field." Based on the structure of the interconnection and business domains, we further differentiate between supply clusters, logistics clusters, and retail clusters. A supply cluster focuses on producers, manufacturers, and supply networks. Several important issues have been investigated in this domain, such as cooperation in the supply chain (Schmitz 1999, Deng et al. 2016, Akçay and Tan 2008), cost advantages of supply clusters (Wu et al. 2006), pricing and quantity policies in coalitions (Granot and Yin 2008, Nagarajan and Sošić 2009, Yin 2010, Huang et al. 2015), and policy instruments to promote socially preferred choices (An et al. 2015, Bagwell 2008). Logistics clusters are "geographically concentrated sets of logistics-related business activities" (Sheffi 2012); these clusters play a very important role in enhancing international trades (Amonkar 2016), gaining competitive advantages in industry (Sheffi 2010), speeding up global supply chains (Elbert and Schönberger 2009), and planning strategic economics development (Lee and Yang 2003). Retail clusters are different from supply clusters and logistics clusters. First, retailers in retail clusters do not cooperate and support each other like suppliers in supplier clusters do. Second, retail clusters directly serve the end consumer market, for which product valuation uncertainty is an important factor.

Retail clusters, also referred to as retail agglomeration, have been well studied in the economics literature. In general, agglomeration refers to the geographic concentration of economic activities (Krugman 1991, Fujita and Thisse 2013). The general agglomeration literature includes two main research streams: one stream studies the geographic concentration of diverse firms in urbanization economies, while the other stream studies the geographic concentration of similar firms in specific industries. Studies on retail clusters belong to the latter. As the focus of our research is retail clusters through the lens of operations research and game theory, we center our literature review on existing analytical studies on retail clusters. For a broader view of agglomeration, we refer readers to excellent review articles elsewhere (e.g., McCann and Folta 2008, Marianov and Eiselt 2014).

Prior studies on retail clusters attempt to explain why retailers form clusters and explore the equilibrium cluster size, considering different characteristics of retailers and consumers. We first identify several key characteristics considered in the literature of retail clusters and then review existing papers in this area based on these characteristics. Table 1 provides a summary of these key characteristics and the comparison of existing papers.

Consumers' valuation uncertainty is a key driving force for retailers to form a cluster. The underlying argument is that consumers have imperfect information about the exact product fit, causing consumers to conduct searches to sample available products, during which they incur

Table 1 **Comparison of Papers on Retail Clusters Based on Key Characteristics**

Categories	Modeling Features	Eaton and Lipsey 1979	Stahl 1982	Wolinsky 1983	Dudey 1990	Fischer and Harrington 1996	Gehrig 1998	Konishi 2005	Takahashi 2013	Our paper
Market Structure	Number of retailers	n with entry	n without entry	n with entry	n with entry	n with entry	n with entry	n without entry	n with entry	n with and without entry
	Number of clusters	k	k	1	k	1 + periphery	2	k	2	k
	Price competition	--	--	--	✓	✓	✓	✓	✓	✓
Uncertainty	Valuation uncertainty	✓	✓	✓	--	✓	✓	✓	✓	✓
	Price uncertainty	--	--	--	✓	✓	--	✓	✓	✓
Heterogeneity	Valuation heterogeneity	--	✓	✓	--	✓	✓	✓	✓	✓
	Location heterogeneity	✓	✓	--	--	--	✓	✓	✓	✓
Revenue Potential	Maximum valuation	--	✓	--	--	✓	--	✓	✓	✓
Cost	Retailers' production cost	✓	✓	✓	✓		--	--	✓	✓
	Consumers' transportation cost	1	t	t	--	--	t	1	0 or $T^{\dagger}$	t

† T is fixed transportation cost. Other parameters n, k, and t are defined the same way as in this paper.

transportation costs. Retailers located in the same cluster help expand product varieties, which in turn increases consumers' probability of finding desired products.

Another characteristic is price uncertainty. Some early studies assume that consumers have prior knowledge about product prices before they sample (Eaton and Lipsey 1979, Stahl 1982, Wolinsky 1983). However, this simplification is only practical under certain scenarios. Since small retailers in developing economies sell unrecognized brands, price information is largely unknown to consumers before they visit retailers. Therefore, price uncertainty also influences consumers' shopping decisions, especially in developing economies.

A third characteristic concerns the spatial aspect of consumers. That is, consumers are from different locations and incur transportation costs proportional to their travel distance. The unit transportation cost in this case is either considered unity (Eaton and Lipsey 1979, Konishi 2005) or as a parameter (Stahl 1982, Gehrig 1998). Other studies have taken a different approach and model the transportation cost as a fixed cost per trip (Takahashi 2013). The cost of transportation in developing economies is critical to consumers, not only in terms of tangible costs like bus fare, but also with respect to travel convenience, especially with larger products like appliances and furniture.

In the literature of retail clusters, retailers' production cost has been shown to play a central role in determining the equilibrium clustering structure. Furthermore, we note that retailers selling similar products in developing and developed economies may have different costs. However, some papers (Fischer and Harrington 1996, Gehrig 1998, Konishi 2005) do not consider this factor.

Some papers in the literature of retail clusters consider number of retailers n as an exogenously given parameter without entry (Stahl 1982, Konishi 2005), while others allow entry of retailers and consider n as an endogenously determined variable (Eaton and Lipsey 1979, Wolinsky 1983, Dudey 1990, Fischer and Harrington 1996, Gehrig 1998, Takahashi 2013). Retailers either simultaneously or sequentially decide if they want to locate in clusters or not. Scholars have investigated a single cluster plus a periphery model (Fischer and Harrington 1996), the two-market model (Gehrig 1998), the two-region model (Takahashi 2013), and the multiple-cluster model (Dudey 1990, Konishi 2005). Wolinsky (1983) finds a single-cluster equilibrium and claims the feasibility of a multiple-cluster equilibrium. Further, it is shown that competition in a cluster drives down prices. Konishi (2005) refers to this effect as the price-cutting effect, which is not analytically present in some early studies (see for example, Stahl 1982, Wolinsky 1983).

Our paper is also related to the literature of concentration in retail industries. Studies in this area investigate the driving forces of market concentration (typically measured by market concentration ratios) for various industries, such as supermarkets (Ellickson 2006, 2007), online book sellers (Latcovich and Smith 2001), newspapers and restaurants (Berry and Waldfogel 2010), and so forth. One key factor considered across these studies is endogenous product quality through firms' investment in increasing product varieties. Consequently, increased product varieties and product quality attract more consumers. This effect is similar to the market-size effect of retail clusters. However, the literature of concentration in retail industries and the literature of retail clusters have fundamental differences. Specifically, "higher concentration" in the former literature refers to a larger market share and more product varieties for one specific firm, whereas "higher concentration" in the latter literature refers to a larger cluster with more independent firms. Since the product variety decision is made by one firm, there is no internal competition in studies that examine concentration in retail industries. In contrast, in studies on retail clusters, firms located in the same cluster compete with each other (i.e., the price-cutting effect), which is one of the key driving forces of clustering results.

Our model extends Konishi (2005) to incorporate key characteristics of developing economies. In Konishi (2005), a few simplifying assumptions are made: consumers' maximum valuation is a constant 1, retailers do not incur any production cost to acquire products, and unit transportation cost is a constant 1. These assumptions are fair and suffice to demonstrate the main tradeoff between the price-cutting effect and the market-size effect. However, we believe that relaxing these assumptions is critical to capture the key characteristics of retail clusters in developing economies. By incorporating more general consumers' maximum valuation, unit transportation cost, and retailers' production cost, we identify two determinants of retailer clusters relevant to developing economies: valuation-cost ratio (consumers' maximum valuation over retailers' production cost) and retailer

density (the number of retailers over unit transportation cost). We find that the equilibrium cluster size decreases in the valuation-cost ratio, and also find that the occurrence of overlapping versus non-overlapping clustering equilibrium is determined by retailer density. Without the above three general parameters, the Konishi (2005) model is incapable of distinguishing developing from developed economies and, thus, cannot make specific recommendations regarding retail clusters in developing economies. Our study is among the first in the literature to investigate retail clusters from the standpoint of developing economies' unique characteristics.

3. Model

In this section, we consider a retail market with n retailers. Each retailer $i \in \{1, ..., n\}$ chooses her store location. Multiple retailers may choose the same store location to form a cluster. We denote the number of clusters as k. For each cluster C_j, $j \in \{1, ..., k\}$, the size of the cluster is denoted by s_j and the location of the cluster is denoted by L_j. A list of notation can be found in Table 2.

Table 2   List of Notation

Notation	Definition
n	Number of retailers
x	Consumers' location, uniformly distributed on a circle of perimeter 1
t	Unit transportation cost
c	Retailers' marginal cost
v	Consumers' maximum valuation for the product
$B_i v$	A consumer's valuation for retailer $i = 1, \cdots, n$, where B_i's are independent and identically distributed (i.i.d.) such that each B_i is uniformly distributed on [0,1]
$\beta_i v$	A consumer's ex post valuation for retailer i
k	Number of clusters
C_j	Set of retailers located at j^{th} cluster, where $j = 1, \cdots, k$
L_j	Location of cluster C_j
s_j	Size of cluster C_j, i.e., number of retailers located at L_j
r_j	Arc distance from cluster C_j to the consumer who is indifferent between visiting cluster C_j or not. All consumers locating within the arc distance r_j to cluster C_j choose to visit C_j.
$U_j(x)$	Consumer x's expected net utility (including transportation cost) from shopping at cluster C_j
$u_i(x)$	Consumer x's net utility (excluding the sunk transportation cost) from buying from retailer i
p_i	Retailer i's price
Π_i	Retailer i's profit
SW	Social welfare
c_e	Retailers' entry cost in the extended model

We consider a unit mass of heterogeneous consumers in terms of their locations, and these consumers are uniformly distributed on a circle of perimeter 1. As shown in Figure 1, without loss of generality, the rightmost point on the circle is denoted as 0 and other points are labeled counterclockwise. This way, this circle of perimeter 1 can be represented by $[0, 1)$. Given a consumer located at x and cluster C_j located at L_j, consumer x incurs a transportation cost of $t \|L_j - x\|$
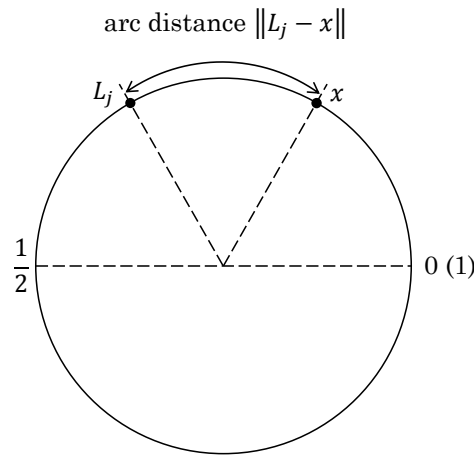
Figure 1   Locations of Consumers and Clusters

when he travels to store location L_j, where parameter t is unit transportation cost and $\|L_j - x\|$ is the arc distance between x and L_j.

Each retailer i charges a price p_i for her product. For each unit of product sold, each retailer incurs a marginal cost c, which may include operational cost, production cost, price paid to the wholesaler, etc. Following Konishi (2005), we assume that consumers are uncertain about their valuations for retailers before they visit physical stores. Consumers share the same ex ante valuations for the retailers, but their ex post valuations can vary. To capture this valuation uncertainty, consumers' valuation of retailer $i = 1, ..., n$ is modeled as $B_i v$, where v is consumers' maximum valuation, and B_i's are independent and identically distributed (i.i.d.) such that each B_i is uniformly distributed on $[0, 1]$. Figure 2 shows the timing of the game.
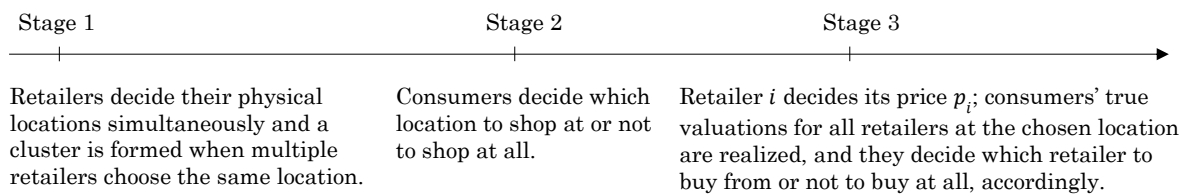

Figure 2   Timing of Events

In Stage 1 (retailer's location choice decision), retailers $i \in \{1, ..., n\}$ decide their physical locations simultaneously. Without loss of generality, we assume that at least one retailer is located at 0. This assumption eliminates the infinitely many location distributions that are equivalent under rotations

on the circle. Further, retailers' location choices lead to k clusters: cluster C_j, $j \in \{1, ..., k\}$, with cluster size s_j.

In Stage 2 (consumers' shopping location decision), given cluster size s_j, each consumer chooses a location to shop at or to not shop at all. Each consumer evaluates each cluster C_j by calculating his expected net utility from shopping at cluster C_j. Without loss of generality, we assume $C_j = \{1, ..., s_j\}$. Each consumer's valuation for retailers in cluster C_j can be modeled as $B_1 v$, $B_2 v$, ..., $B_{s_j} v$, where $B_1, B_2, ..., B_{s_j}$ are i.i.d. random variables uniformly distributed on $[0, 1]$. Following prior studies in the literature of retail clusters (Stahl 1982, Dudey 1990, Konishi 2005, Takahashi 2013), we assume that each consumer only shops at one location or not at all. The impact of consumers' potential searches at multiple locations has been studied in the consumer search literature (Wolinsky 1983). For a consumer located at x, his expected utility from shopping at cluster C_j is:

$$U_j(x) = E\left[\max_{i \in C_j}\{0, B_i v - p_i^e\}\right] - t\left\|L_j - x\right\| \tag{1}$$

where $j \in \{1, ..., k\}$ and p_i^e is the expected price of retailer i.

Consumer x's expected utility from not shopping at all is:

$$U_0(x) = 0 \tag{2}$$

Consumer x compares his expected utilities from shopping at all available locations and chooses the cluster that yields the highest expected utility. If his expected net utilities from shopping at all clusters are negative, then consumer x chooses not to shop at all. In other words, consumers' shopping location decision can be represented as:

$$j^*(x) = \arg\max_{\{0,1,...,k\}}\{U_0(x), U_j(x)\}, \text{ where } j \in \{1, ..., k\} \tag{3}$$

In Stage 3 (price competition among retailers at the same location), retailers in the same cluster simultaneously decide their prices. Once a consumer visits a cluster, his true valuations for all retailers in the cluster are realized. For each $i \in C_j$, each consumer realizes his ex post true valuation $\beta_i v$. After visiting retailer i, different consumers realize different valuations for retailer i. Consumers then become ex post heterogeneous. Across all consumers, the different realized β_i values also follow the same uniform distribution within $[0, 1]$ as B_i.[2] These heterogeneous realized true valuations (the vertical dimension), together with heterogeneous locations (the horizontal dimension), represent two dimensions of heterogeneity in our model. This modeling approach is consistent with existing

---

[2] This modeling approach is referred to as a fluid model (Huang et al. 2013, Yu et al. 2015), in which the distribution for ex post realized valuations across consumers is the same as the common valuation distribution that consumers believed ex ante.

models with two-dimensional heterogeneity in the literature (Gilbert and Matutes 1993, Desai 2001, Morton and Zettelmeyer 2004, Du et al. 2005, Zhao et al. 2012), which allows us to consider both horizontal and vertical differentiations simultaneously.

Since transportation costs are sunk, consumers make their purchasing decisions based on their net utilities (excluding transportation costs) of buying from retailers at the chosen location. For a consumer who shops at cluster C_j and retailer $i \in C_j$, his net utility from buying from retailer i is:

$$u_i = \beta_i v - p_i \tag{4}$$

His expected utility from not buying at all is:

$$u_0 = 0 \tag{5}$$

For a consumer who shops at cluster C_j, he buys from retailer i if and only if retailer i offers the maximum nonnegative utility among all the retailers in the cluster. If all retailers in the cluster yield negative utilities, then he will not purchase at all. In other words, his purchasing decision can be represented as:

$$i^* = \underset{\{0,1,\ldots,s_j\}}{\arg\max}\{u_0, u_i\}, \text{ where } i \in C_j \tag{6}$$

4. Equilibrium Analyses

We use backward induction to analyze the game. Subsections 4.1, 4.2, and 4.3 examine stages 3, 2, and 1, respectively.

4.1. Price Competition Among Stores at the Same Location

After the cluster sizes are determined in stage 1 and a cluster's market size is determined in stage 2, retailers inside a cluster decide their prices to compete with each other. Let the cluster size of cluster C_j be s_j and the market size be $2r_j$. Without loss of generality, we may assume that $1, 2, \ldots, s \in C_j$.

In stage 3, a consumer buys from a retailer i if and only if retailer i offers him the maximum nonnegative utility among all s retailers in the cluster. That is, a consumer buys from retailer i in the cluster with the following probability.

$$Prob\left[\beta_i v - p_i \geq 0 \text{ and } \beta_m v - p_m \leq \beta_i v - p_i \text{ for } m \in \{1, 2, \ldots, s\} \setminus i\right]$$

Knowing the consumer's decision making process, retailer i, where $i \in \{1, 2, \ldots s\}$, decides the optimal price to maximize her expected profit as follows.

$$\max_{p_i} \Pi_i = (p_i - c) \cdot (2r_j) \cdot Prob\left[\beta_i v - p_i \geq 0 \text{ and } \beta_m v - p_m \leq \beta_i v - p_i \text{ for } m \in \{1, 2, \ldots, s\} \setminus i\right]$$

Every retailer in cluster C_j decides her optimal price. In this simultaneous price competition game, we find that there are only symmetric equilibrium solutions. Furthermore, the symmetric price equilibrium is unique. The results are summarized in the following Lemma 1.

LEMMA 1. *For a given cluster size s, there is a unique price equilibrium p_i^*, which is symmetric with $p_i^* = p_s^*, \forall i \in \{1, 2, \cdots, s\}$, where p_s^* solves $p^s = v^{s-1}[v - sp + sc]$. Furthermore, p_s^* has the following properties:*

- *p_s^* decreases in cluster size s.*
- *p_s^* increases in consumers' maximum valuation v.*
- *p_s^* increases in marginal cost c.*

All the proofs can be found in the Appendix.

Lemma 1 shows[3] that the unique equilibrium price decreases in cluster size s. This is caused by price competition. When there are more retailers in a cluster, price competition is more intense, which drives the equilibrium price down. The equilibrium price also increases in v and c. When consumers are willing to pay more, retailers can charge a higher price. When marginal cost increases, retailers increase the price to maintain profitability.

Note that the equilibrium price is always higher than the marginal cost, no matter how intense the price competition is. This is because consumers are ex ante homogeneous but ex post heterogeneous in terms of their valuations for retailers. Thus, they realize different valuations for different retailers. Consider an example with two retailers 1 and 2: even if retailer 1 charges a very low price, retailer 2 can still charge a higher price and make positive profit because some consumers' realized valuation for retailer 2 is higher than that for retailer 1. As a result, all retailers are able to charge a markup and still make positive profits.

4.2. Market Size of a Retail Cluster

In this subsection, we analyze the second stage, in which consumers decide where to shop. A consumer chooses to go to a location that offers him the highest nonnegative expected utility or to not shop at all. Given a cluster C_j with cluster size s and location L_j, a consumer's expected utility from visiting this cluster is $U_j(x) = E\left[\max_{i \in C_j}\{0, B_i v - p_i^e\}\right] - t\|L_j - x\|$. Note that this expected utility $U_j(x)$ increases in cluster size s_j. In other words, a consumer prefers to shop at larger clusters because the retail price is lower and his highest valuation across all available retailers is higher at larger clusters.

---

[3] Konishi (2005) showed the existence and uniqueness of symmetric equilibrium under more general distributions of random variable B_is, but he did not consider the possibility of asymmetric price equilibria. Here, we prove that there is no asymmetric price equilibrium when B_is are uniformly distributed. Thus, our result is much stronger than Konishi's (2005).

Individual consumers' choices of shopping locations or not shopping at all result in different demand patterns for retailers. Depending on whether the consumer market is fully covered or not, there are two cases of demand patterns: the non-overlapping case and the overlapping case. We use subscripts n and o to denote these two cases, respectively. Figures 3 and 4 illustrate the non-overlapping and overlapping cases for an example of three clusters.
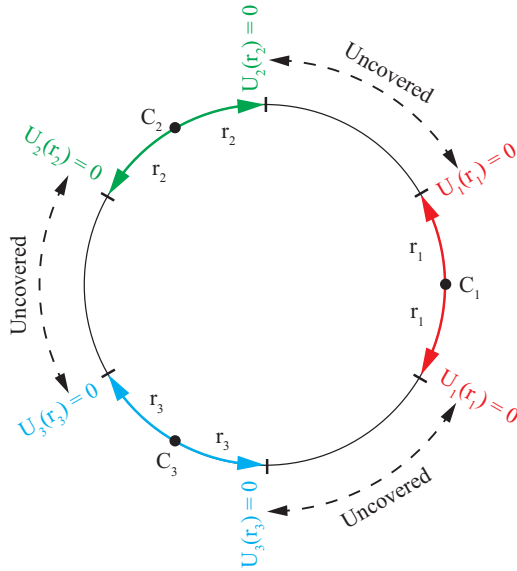

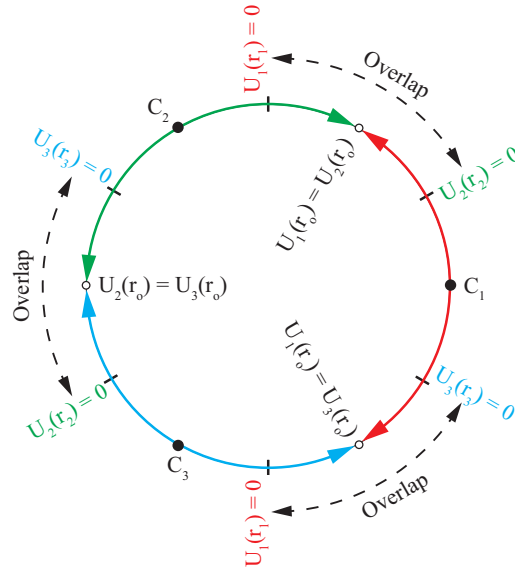**Figure 3   Non-overlapping case**   **Figure 4   Overlapping case**

The *non-overlapping* case refers to the case in which the consumer market is not fully covered, as shown in Figure 3. Let r_n denote the maximum arc distance that consumers are willing to travel to visit the cluster and obtain nonnegative expected utility. Then,

$$r_n \equiv \frac{E\left[\max_{i \in C_j}\{0, B_i v - p_i^e\}\right]}{t}.$$

All consumers within a cluster's potential market coverage should choose to visit the cluster. In other words, consumers located within the arc distance r_n to the cluster location L_j should be willing to visit the cluster C_j, unless there are other clusters offering higher positive expected utilities. Therefore, a cluster's potential market coverage is r_n arc distance from its cluster location in both directions, and the market size of a cluster is $2r_n$. In this case, no other clusters yield positive expected utilities within a cluster's potential market coverage; as a result, the clusters' potential market coverages have no overlapping areas. In Figure 3, the arrows indicate the market coverages of all three clusters. For example, for cluster C_1, only consumers located within arc distance r_1 to cluster C_1 have nonnegative utilities from visiting it. Therefore, the market coverage of C_1 is $2r_1$, as indicated by the red arrows in Figure 3.

The *overlapping* case refers to the case in which the consumer market is fully covered, as shown in Figure 4. Consumers located within the arc distance r_o to a cluster should choose to visit the cluster; thus, a cluster's market size in the overlapping case is $2r_o$. In this case, some consumers can get positive expected utilities from at least two clusters, and consumers have to compare the expected utilities to decide which cluster to visit. Thus, the clusters' potential market coverages have overlapping areas. In Figure 4, the dashed arrows indicate the overlapped areas, in which consumers can get nonnegative utilities from more than one cluster. For example, clusters C_1 and C_2 have an overlapping area located somewhere between the two clusters. Consumers in this area need to compare U_1 and U_2 to decide which cluster to visit. Let r_o denote the maximum arc distance that a consumer is willing to travel to obtain the higher expected utility from one of its neighboring clusters in the overlapping case. The indifferent consumer locates at r_o arc distance away from C_1. Therefore, the market coverage of C_1 is $2r_o$, as indicated by the red arrows in Figure 4.

We consider symmetric clusters evenly locating on the circle.[4] Accordingly, n total retailers can form $k = \frac{n}{s}$ clusters with s retailers in each cluster. Since the total arc distance on the circle is 1, the arc distance between two clusters is $\frac{1}{k}$, which is also $\frac{s}{n}$. If clusters are symmetric and evenly located on the circle, then r_o is exactly half the distance between these two neighboring clusters. Therefore,

$$r_o \equiv \frac{s}{2n}.$$

LEMMA 2 **(Market Size of a Retail Cluster)**. *For a given cluster size s, the market size of a cluster in the non-overlapping case is $2r_n$, and the market size of a cluster in the overlapping case is $2r_o$.*

- *Both r_n and r_o increase in cluster size s.*
- *r_n is independent of n but decreases in t.*
- *r_o is independent of t but decreases in n.*

Lemma 2 shows that a cluster's market size increases in its cluster size. In the non-overlapping case, this occurs because consumers are more likely to find what they want as the cluster size increases. Thus, a consumer's expected valuation from shopping at a cluster increases, thereby attracting consumers from farther away to visit the cluster. In turn, the cluster's market size increases. In the overlapping case, this occurs because the number of clusters decreases when the cluster size increases. Thus, the distance between clusters increases and competition among clusters decreases. In turn, the market size of a cluster increases as well.

---

[4] We temporarily assume that clusters have the same size s. This assumption is later validated in Proposition 2, when we show that the sizes of clusters in equilibrium are very close, with the majority having the maximum difference of 1. This is also consistent with the findings in Konishi (2005).

Three factors determine the market size of a cluster: the distance between clusters, a consumer's expected valuation from shopping, and unit transportation cost. In the overlapping case, the distance between clusters solely determines the market size while the last two factors do not affect a cluster's market size. In the non-overlapping case, a consumer's expected valuation from shopping and unit transportation cost determine a cluster's market size, and the distance between clusters does not matter.

The number of retailers n affects the distance between clusters. Given cluster size s, the distance between clusters decreases when n increases. Therefore, the market size decreases in n in the overlapping case, but is independent of n in the non-overlapping case.

Since consumer demand patterns in the second stage depend on retailers' location decisions in the first stage, the non-overlapping and overlapping demand patterns are different. In the non-overlapping case, although retailers located within the same cluster engage in price competition (i.e., intra-cluster competition), clusters do not directly compete with each other (i.e., no inter-cluster competition). In the overlapping case, neighboring clusters directly compete with each other. In other words, intra-cluster competition exists in both cases, but inter-cluster competition only exists in the overlapping case. We explore retailers' location decisions and the corresponding clustering equilibrium next.

4.3. Retailers' Location Decisions

In this subsection, we study retailers' location decisions. If multiple retailers choose the same location, then they form clusters. Consequently, retailers' location decisions are connected with the optimal cluster size that maximizes each retailer's profit.

In the profit function, the proportion of consumers that buy from a retailer is measured by $Prob\,[\,\beta_i v - p \geq 0$ and $\beta_m \leq \beta_i$ for $m \in \{1, 2, \ldots, s\} \setminus i]$. In other words, knowing that equilibrium price is symmetric, consumers compare realized valuations for all retailers in the cluster to determine which one to buy from. Eventually, a consumer buys from the retailer that offers him non-negative utility and the highest realized valuation when compared to other retailers in the cluster. We have

$$Prob\,[\,\beta_i v - p \geq 0 \text{ and } \beta_m \leq \beta_i \text{ for } m \in \{1, 2, \ldots, s\} \setminus i] = \int_{\frac{p}{v}}^{1} \beta_i^{s-1} d\beta_i = \left(\frac{1}{s}\right) \cdot \left(1 - \left(\frac{p}{v}\right)^s\right)$$

Therefore, the problem is

$$\max_s \Pi_i = (p - c) \cdot (2r) \cdot \left(\frac{1}{s}\right) \cdot \left(1 - \left(\frac{p}{v}\right)^s\right) \tag{7}$$

$$subject\ to\colon\ p^s = v^{s-1}\left[v - sp + sc\right]$$

$$r \leq r_n$$

$$r \leq r_o$$

$$p \geq 0,\ r \geq 0,\ 0 \leq s \leq n$$

In the profit function, $2r$ is the market size of a cluster. The condition $r \leq r_n$ ensures that consumers get nonnegative expected utility from visiting the cluster. The condition $r \leq r_o$ ensures consumers get the highest expected utility from visiting the cluster. Together, they specify that consumers within the arc distance r to the cluster location should visit the cluster because they get the highest nonnegative expected utility from doing so. At the optimal solution, one of the two conditions must be binding. If $r_n < r_o$, then the condition $r \leq r_n$ binds and $r = r_n$, which indicates that the clustering equilibrium is the non-overlapping case. On the contrary, if $r_o < r_n$, then the condition $r \leq r_o$ binds and $r = r_o$, which indicates that the clustering equilibrium is the overlapping case.

PROPOSITION 1. *There exists a unique s^* that solves problem (7).*

Based on s^*, which is a continuous variable, one can derive the Pareto optimal equilibrium cluster size, which is an integer. The result is summarized in Proposition 2.

PROPOSITION 2 (**Equilibrium Cluster Size**). *There are three possible Pareto optimal equilibria[5]. Either all clusters have the same size s_d^*, or there are two cluster sizes s_d^* and s_d^{**}. The optimal cluster size $s_d^* = \lfloor s^* \rfloor$ or $\lceil s^* \rceil$, depending on whichever yields the higher profit. Specifically,*

- *If n is divisible by s_d^*, i.e., $(n \mod s_d^*) = 0$, then all clusters have the same size s_d^*.*
- *If n is indivisible by s_d^*, i.e., $(n \mod s_d^*) > 0$, then there are two cluster sizes s_d^* and s_d^{**}.*
    - *If $\Pi(s_d^* + 1) \geq \Pi(n \mod s_d^*)$, then $s_d^{**} = s_d^* + 1$.*
    - *If $\Pi(s_d^* + 1) < \Pi(n \mod s_d^*)$, then $s_d^{**} = (n \mod s_d^*)$. In this case, there is only one cluster with size s_d^{**}; the rest have size s_d^*.*

Proposition 2 shows that the optimal continuous solution s^* is very close to the optimal cluster size s_d^*. Furthermore, the majority of clusters have a size of s_d^* or $s_d^* + 1$. Considering the large cluster scale in developing economies, we will use the continuous solution s^* to approximate the optimal cluster size in Section 5 and obtain the sensitivity analysis based on s^*.

---

[5] In the rest of the paper, equilibrium refers to Pareto optimal equilibrium unless explicitly specified otherwise.

5. Equilibrium Cluster Size and Clustering Equilibrium

In this section, we present two important factors that determine the equilibrium cluster size and clustering equilibrium. In Subsection 5.1, we discuss one of the two key factors, the valuation-cost ratio. We then study the other factor, the retailer density in Subsection 5.2. Finally, we show implications of our findings for developing economies in Subsection 5.3.

5.1. Impact of Valuation-Cost Ratio on Equilibrium Cluster Size

First, we define $\frac{v}{c}$ as the *valuation-cost ratio*, which increases in consumers' maximum valuation v and decreases in retailers' marginal cost c. The valuation-cost ratio indicates the profit potential for retailers. The higher the valuation-cost ratio, the higher the profit potential for a retailer.

Second, we find that the equilibrium cluster size in both overlapping and non-overlapping cases *solely* depends on the valuation-cost ratio. In other words, if both v and c increase or decrease by the same percentage, then the valuation-cost ratio stays the same and the equilibrium cluster size also remains the same. In both cases, the equilibrium cluster size monotonically decreases in the valuation-cost ratio. We summarize this result as follows.

PROPOSITION 3 (**Impact of Valuation-Cost Ratio on Equilibrium Cluster Size**). *The optimal cluster size in both the overlapping case and the non-overlapping case decreases in valuation-cost ratio $\frac{v}{c}$.*

The intuition behind Proposition 3 is as follows. A cluster has a market-size effect, which is mainly driven by consumer valuation uncertainty. In other words, because of consumer valuation uncertainty, a consumer's expected valuation from visiting a large cluster (which is the expected maximum valuation among the stores in the cluster) is higher than visiting a small cluster. Therefore, a large cluster attracts more customers and enjoys a larger market size than a small cluster.

However, a cluster also has a price-cutting effect due to price competition within a cluster. Note that the equilibrium price is still higher than the marginal production cost for retailers because of consumer heterogeneity in terms of their realized valuations.

Hence, the equilibrium cluster size is the result of balancing these two effects: the market-size effect and the price-cutting effect. We find that the equilibrium cluster size decreases in the valuation-cost ratio, which indicates the profit potential from serving one customer. When the valuation-cost ratio decreases, the profit potential decreases. For a product with low profit potential, the improvement on profit margin is limited and, thus, retailers focus more on improving sales volume. This makes the market-size effect more important than the price-cutting effect when assessing the tradeoff between them. Therefore, the cluster size increases to further improve market size when the profit potential decreases (i.e., the valuation-cost ratio decreases).

One important difference between a developed economy and a developing economy lies in their valuation-cost ratios (i.e., their profit potentials). For example, in 2016, the average net margins in the U.S. were 14.18% for computer peripherals and 14.19% for pharmaceuticals, compared to 1.83% (computer peripherals) and 9.93% (pharmaceuticals) in China (Damodaran 2016). The valuation-cost ratio in a developing economy is usually lower than the ratio in a developed economy for most products, due to either high retailers' production cost c, low consumers' valuation v, or both. Consumers are relatively richer in a developed economy than in a developing economy, and retailers in a developed economy usually sell products with well-known brands. However, retailers in a developing economy usually sell unrecognized brands. The profit potential is usually higher for well-known brands than for unrecognized brands.

Our finding in Proposition 3 suggests a larger cluster size in developing economies since the valuation-cost ratio is lower in developing economies than in developed economies. This helps to explain the phenomenon of massive clusters in developing economies, as well as the phenomenon that the majority of retailers in massive clusters sell unrecognized brands.

5.2. Impact of Retailer Density on Clustering Equilibrium

Another natural question is whether the Pareto optimal clustering equilibrium is the overlapping case or the non-overlapping case. We find that the answer depends on the ratio of the total number of retailers and unit transportation cost (i.e., $\frac{n}{t}$). We refer to the ratio of $\frac{n}{t}$ as *retailer density* since it represents the number of retailers normalized by the unit transportation cost. Retailer density $\frac{n}{t}$ measures the degree of competition intensity. When more retailers enter the market (n increases), the retail competition is more intense; when it is more costly for consumers to travel to store locations (t increases), the retail competition is less intense.

PROPOSITION 4 **(Impact of Retailer Density on Clustering Equilibrium)**. *When retailer density $\frac{n}{t}$ is lower than a threshold, the Pareto-optimal clustering equilibrium is the non-overlapping case; otherwise, the clustering equilibrium is the overlapping case.*

The intuition of Proposition 4 is as follows. When retailer density $\frac{n}{t}$ is very low (due to either very few retailers existing in the market or very high unit transportation cost), each cluster serves its own local customers and has limited market coverage. As a result, the clustering equilibrium is the non-overlapping case, and the consumer market is not fully covered. Consequently, there is no inter-cluster competition. When retailer density $\frac{n}{t}$ increases, the number of retailers in a cluster grows, which intensifies the intra-cluster competition among retailers located at the same location. Once intra-cluster competition becomes sufficiently severe, the clustering equilibrium switches from the non-overlapping case to the overlapping case. Retailers would rather endure the increased inter-cluster competition in order to reduce the intra-cluster competition. From the retailers' perspective,

inter-cluster competition is reduced in the non-overlapping case, while intra-cluster competition is reduced in the overlapping case.

When compared to travel in developed economies, traveling is usually more costly in developing economies (higher t) due to poor transportation infrastructure. However, the comparison of the total number of retailers between developing economies and developed economies varies with products. For low-tech product markets with relatively low entry costs, there is a much larger number of retailers in developing economies (e.g., small commodity markets in Yiwu, China) than in developed economies. Thus, retailer density $\frac{n}{t}$ is higher in developing economies for these products. Our finding in Proposition 4 suggests that the overlapping clustering equilibrium is more likely to be observed in developing economies for these products. On the other hand, for high-tech product markets with relatively high entry costs (e.g., automobiles), there are few retailers in both developing and developed economies. Thus, retailer density $\frac{n}{t}$ is lower in developing economies for these products due to high transportation cost, leading to the non-overlapping clustering equilibrium.

Although Proposition 4 shows that the Pareto optimal equilibrium may be either overlapping or non-overlapping depending on retailer density $\frac{n}{t}$, there may exist two possible candidate equilibria: overlapping and non-overlapping. Proposition 5 compares the cluster sizes in these two possible candidate equilibria.

PROPOSITION 5 (**Cluster Size Comparison**). *When valuation-cost ratio $\frac{v}{c}$ is lower than a threshold, the cluster size is higher in the overlapping equilibrium than the non-overlapping equilibrium; otherwise, the cluster size is lower in the overlapping equilibrium than the non-overlapping equilibrium.*

The intuition of Proposition 5 is as follows. Suppose the clustering equilibrium switches from the non-overlapping case to the overlapping case, i.e., the total market coverage increases from partial coverage to full coverage. Retailers can do so by choosing one of the following two distinct strategies: the concentration strategy (fewer clusters with larger cluster size) and the dispersion strategy (more clusters with smaller cluster size). Following the concentration strategy, more retailers co-locate together to attract more total customers by taking advantage of the increased highest valuation from more available retailers at a cluster. Meanwhile, following the dispersion strategy, retailers locate away from each other and closer to customers to cover more total customers by taking advantage of reduced price competition and reduced transportation cost.

When choosing between concentration and dispersion strategies, retailers consider the tradeoff between two effects: the price-cutting effect and the market-size effect. For a product with low profit potential, the improvement on profit margin is limited and, thus, retailers focus more on improving sales volume. This makes the market-size effect more important than the price-cutting

effect, and retailers subsequently use the concentration strategy to increase total market coverage. Therefore, for products with low profit potentials, the cluster size increases, i.e., the cluster size in the overlapping equilibrium is higher than that in the non-overlapping equilibrium. On the other hand, for a product with high profit potential, retailers focus more on profit margin than sales volume. Thus, the price-cutting effect dominates the market-size effect, and retailers subsequently use the dispersion strategy to reduce cluster size but increase the number of clusters to increase total market coverage. Therefore, the cluster size in the overlapping equilibrium is smaller than that in the non-overlapping equilibrium for products with high profit potentials.

Proposition 5 has opposite implications for developing and developed economies. Since the majority of retailers in developing economies have low profit potentials (i.e., low valuation-cost ratios), the cluster size is most likely larger in the overlapping equilibrium than in the non-overlapping equilibrium for developing economies. Conversely, because of the high profit potentials (i.e., high valuation-cost ratios) in developed economies, the cluster size is usually smaller in the overlapping equilibrium than in the non-overlapping equilibrium for developed economies.

5.3. Managerial Implications for Retail Clusters

In the previous two subsections, we identified two important determinants for the Pareto optimal clustering equilibrium: valuation-cost ratio ($\frac{v}{c}$) and retailer density ($\frac{n}{t}$). In this subsection, we compare developing and developed economies in terms of their valuation-cost ratios and retailer densities, and then discuss the managerial implications of our findings for retail clusters in developing economies.

Overall, valuation-cost ratio and retailer density jointly determine the clustering equilibrium. As shown in Table 3, there are four types of retail clusters (overlapping massive clusters, non-overlapping large clusters, non-overlapping small clusters, and overlapping mini clusters) depending on the levels of valuation-cost ratio ($\frac{v}{c}$) and retailer density ($\frac{n}{t}$). Our findings suggest that retailers form either mini or small clusters for product markets with high profit potential (high $\frac{v}{c}$) while they form either large or massive clusters for product markets with low profit potential (low $\frac{v}{c}$). Furthermore, when facing intense competition (high $\frac{n}{t}$), retailers form overlapping clusters and serve the whole consumer market. When the competition intensity is low, it is more beneficial for the retailers to form non-overlapping clusters and only serve the local consumer market. Next, we discuss these four types of retail clusters in detail.

Overlapping massive clusters form in product markets with high $\frac{n}{t}$ and low $\frac{v}{c}$. The overlapping case becomes the Pareto optimal equilibrium when retailer density $\frac{n}{t}$ is high (according to Proposition 4). The cluster size is large when valuation-cost ratio $\frac{v}{c}$ is low (according to Proposition 3). Furthermore, the cluster size in the overlapping case is even larger than the one in the

Table 3   Summary of Key Findings for Retail Clusters

		Valuation-Cost Ratio (v/c)	
		Low v/c (Low profit potential)	High v/c (High profit potential)
Retailer Density (n/t)	High n/t (High competition intensity)	Overlapping massive clusters	Overlapping mini clusters
	Low n/t (Low competition intensity)	Non-overlapping large clusters	Non-overlapping small clusters

non-overlapping case when valuation-cost ratio $\frac{v}{c}$ is low (according to Proposition 5). Therefore, massive clusters usually form in the overlapping case when valuation-cost ratio $\frac{v}{c}$ is low. For low-tech products with low entry costs (e.g., socks, zippers, and other small commodities), there are usually numerous retailers, leading to high retailer density and intense competition. In developing economies, the profit potential is low relative to developed economies due to lagging technologies, limited purchasing power, poor product quality and customer service, and so forth. In addition, most of these retailers carry unrecognized brands, which further reduces consumers' valuation and, thus, the profit potential of the product. Therefore, overlapping massive clusters are more likely to occur in developing economies than in developed economies. The small commodity markets in Yiwu, China represent a typical example of overlapping massive clusters. For example, small commodities and electronics usually have high retailer densities, which results in the overlapping clustering equlibrium. Additionally, these products usually have low profit potential in developing countries. Therefore, massive clusters have been observed for these products in developing countries, such as the small commodity markets in Yiwu, China and the Huaqiangbei electronics bazaar in Shenzhen, China. The Mega Center in the Yiwu small commodity markets occupies about 46 million square feet (Goodman 2011) and accommodates about 65,000 booths as of 2014 (Chen 2015). The Huaqiangbei electronics bazaar in Shenzhen, China is another example of overlapping massive clusters, which is referred to as the "Silicon Valley of Hardware" (Reynolds 2016). Huaqiangbei houses thousands of sellers of electronic parts, digital devices, gadgets, and their peripherals. For both the markets described here, the consumer markets (for small commodities and electronics, respectively) are fully covered by the massive clusters.

Non-overlapping large clusters form in product markets with low $\frac{n}{t}$ and low $\frac{v}{c}$. Some products require relatively more complex technology (e.g., computer keyboards, video game controllers). Such technological barriers lead to high entry costs, which limit the number of retailers in the market. These product markets usually have low profit potential due to their low consumer valuation, especially in developing economies. Retailers in such product markets are expected to form large clusters, each of which serves its own local consumer market; consequently, the consumer market is

not fully covered. Furniture clusters (e.g., Kanpur Steel Furniture cluster in Uttar Pradesh, India) represent an example of such non-overlapping large clusters.

Non-overlapping small clusters form in product markets with low $\frac{n}{t}$ and high $\frac{v}{c}$. For products with high entry cost and high profit potential (e.g., automobiles), the non-overlapping clustering equilibrium occurs since the high entry cost deters new entrants and the competition intensity is low. At the same time, these non-overlapping clusters are small because of the high profit potential. For example, in practice, we observe small automobile clusters in both developing and developed economies. However, the automobile clusters in developing economies are usually larger than those in developed economies.

Overlapping mini clusters form in product markets with high $\frac{n}{t}$ and high $\frac{v}{c}$. In these product markets, overlapping mini clusters are preferred because of both high retailer density and high profit potential. Pharmaceutical retailers (e.g., CVS, Walgreens) demonstrate an example of such overlapping mini clusters. Pharmaceutical clusters are usually very small with either a single retailer or two retailers at the same location. Also, they are usually located close to their customers, and different clusters compete with each other. Since the profit potential in developing economies is usually lower than that in developed economies, we expect to observe more such overlapping mini clusters in developed economies.

In summary, retail clusters vary across different product markets. When compared to those in developed economies, large or massive clusters are more likely to occur, and mini or small clusters are less likely to occur in developing economies, due to the lower profit potential in developing economies. For products with low entry costs, overlapping clusters are more commonly seen in developing economies than in developed economies, due to a larger number of retailers in developing economies. For products with high entry costs, non-overlapping clusters are more commonly seen in developing economies than in developed economies because transportation is more costly in developing economies.

6. Socially Optimal Cluster Size

In this section, we examine the socially optimal retail clusters from the social planner's perspective and discuss the policy implications of our findings.

6.1. Socially Optimal Cluster Size and Its Comparison to Equilibrium Cluster Size

In this subsection, we first solve for the socially optimal cluster size and then compare it to the equilibrium. In this analysis, the social planner makes location decisions for the retailers in Stage 1, and Stages 2 and 3 remain the same as in the equilibrium analysis. The social planner aims to maximize social welfare, denoted by SW, which consists of consumer surplus and the profits of all

retailers. Consumer surplus, denoted by CS, is the sum of net utilities for all consumers. For each retailer, we know from Subsection 4.3 that her profit is $\Pi = \left(\frac{1}{s}\right) \cdot \left(1 - \left(\frac{p}{v}\right)^s\right) \cdot (p - c) \cdot (2r)$. Thus, the total profit of all retailers $n\Pi = \left(\frac{n}{s}\right) \cdot \left(1 - \left(\frac{p}{v}\right)^s\right) \cdot (p - c) \cdot (2r)$. The expected surplus of consumers who shop at clusters is

$$
CS = \underbrace{\left(\frac{n}{s}\right)}_{\text{number of clusters}} \left(2r \underbrace{\int_{\frac{p}{v}}^1 s\beta_i^{s-1}(\beta_i v - p)\, d\beta_i}_{\text{a customer's expected benefit from a cluster}} - \underbrace{2\int_0^r tx\, dx}_{\text{total transportation cost to a cluster}} \right)
$$
$$
= \left(\frac{n}{s}\right)\left(\frac{2r\left(p\left(\left(\frac{p}{v}\right)^s - s - 1\right) + sv\right)}{s+1} - r^2 t \right).
$$

Therefore, the social planner's problem can be formulated as:

$$
\max_s SW = n\Pi + CS = \frac{n}{s}\left(\frac{2r\left((cs + c - ps)\left(\frac{p}{v}\right)^s - cs - c + sv\right)}{s+1} - r^2 t \right) \tag{8}
$$

subject to the same set of constraints presented in Subsection 4.3. Solving the social planner's problem yields the socially optimal cluster size s^*_{social}.

PROPOSITION 6 (**Socially Optimal Cluster Size**). *In the non-overlapping case, the socially optimal cluster size decreases in valuation-cost ratio $\frac{v}{c}$.*

Proposition 6 shows that the socially optimal cluster size is larger for products with lower profit potential (lower $\frac{v}{c}$) in the non-overlapping case. This property of the socially optimal cluster size is similar to the equilibrium cluster size.

PROPOSITION 7 (**Socially Optimal Cluster Size vs. Equilibrium Cluster Size**). *In the non-overlapping case, the socially optimal cluster size is larger than the equilibrium cluster size. The gap between these two cluster sizes decreases in valuation-cost ratio $\frac{v}{c}$.*

Proposition 7 suggests that for low-retailer-density products and, thus, in the non-overlapping case, the social planner should try to increase cluster size in both developed and developing economies. For developing economies, although the cluster sizes are usually large already, Proposition 7 indicates a bigger gap between the equilibrium and the socially optimal cluster sizes, when compared to developed economies.

6.2. Policy Implications for Retail Clusters

Our findings regarding socially optimal retail clusters presented in the previous subsection have important policy implications for retail clusters. We find that without government intervention,

retailers tend to form smaller clusters than the social optimum. Therefore, to encourage larger clusters, government agencies are advised to deploy various policies such as reduced rents, participation grants, tax benefits, etc.

Furthermore, our results indicate that the gap between the equilibrium cluster size and the social optimum is larger for products with low profit potential. Policies that offer more substantial benefits to participating retailers in clusters are recommended for these products. In addition, since the profit potential is usually lower in developing economies, there is a stronger need for such policies in developing economies than in developed economies.

7. Extended Model with Retailer Entry

In this section, we extend the model to endogenize the entry of retailers with an entry cost c_e. A retailer enters the market if and only if her profit is no less than entry cost c_e. Hence, number of retailers n on the market becomes an endogenous decision. We solve this problem in two steps. First, we explore the properties of a retailer's optimal profit as a function of n. Second, we examine the impact of entry cost c_e on equilibrium number of retailers on the market n^*.

Let $\Pi(n)$ be the maximum profit function of a retailer, given n retailers in the market. We use subscripts N and O to denote non-overlapping and overlapping cases in this extended model, respectively. Let s_N^* be the unique optimal cluster size that maximizes a retailer's equilibrium profit in the non-overlapping case.

PROPOSITION 8. *There exists a threshold $\widehat{n}$ as illustrated in Figure 5. A retailer's optimal profit $\Pi(n)$ has the following properties:*

- *If $1 \leq n \leq s_N^*$, then all retailers join one cluster, and the market is not fully covered, i.e., the non-overlapping case is the equilibrium. A retailer's profit $\Pi(n) = \Pi_N(n)$ increases in n.*
- *If $s_N^* < n < \widehat{n}$, then retailers may form more than one cluster, but the market is still not fully covered, i.e., the non-overlapping case is still the equilibrium. A retailer's profit $\Pi(n) = \Pi_N(n)$ is a section of a damped oscillating function with horizontal asymptote at $\Pi_N(s_N^*)$.*
- *If $n \geq \widehat{n}$, then the market is fully covered, i.e., the overlapping case is the equilibrium. A retailer's profit $\Pi(n) = \Pi_O(n) = \frac{\Pi_T}{n}$ decreases in n, where constant Π_T is the total profit of all retailers.*

Proposition 8 and Figure 5 show how a retailer's optimal profit changes as number of retailers n increases. When n is small (less than s_N^*), the market-size effect dominates the price-cutting effect. All retailers cluster together and the cluster size keeps increasing as n increases. When n exceeds s_N^*, the price-cutting effect dominates the market-size effect. A retailer's profit decreases as more retailers join the cluster. At a certain point, retailers are better off breaking away to form
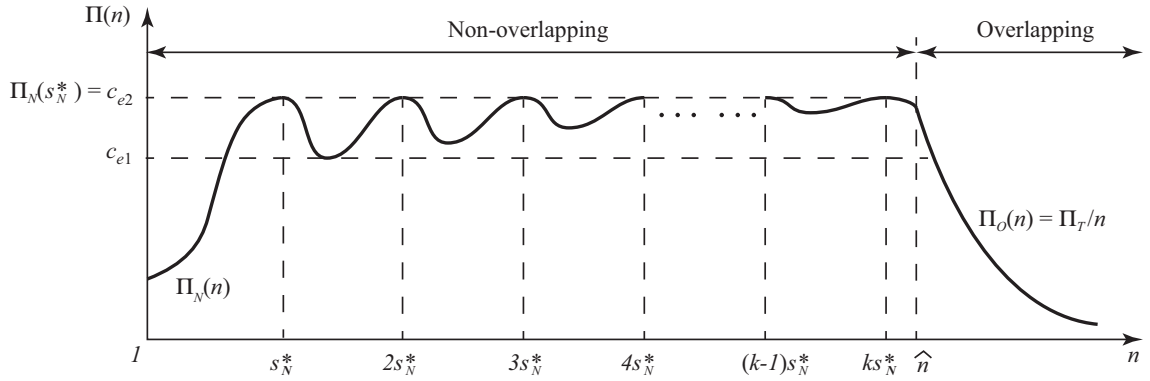
Figure 5 A retailer's profit in equilibrium as a function of n

a new cluster. In the new equilibrium, the number of clusters increases and the cluster size is less than s_N^* again. Consequently, a retailer's profit increases again as n increases. The cycle repeats until number of retailers n reaches threshold $\widehat{n}$. Then, the market is fully covered and becomes the overlapping case. A retailer's profit monotonically decreases in n in this overlapping region.

PROPOSITION 9. *There exists two thresholds c_{e1} and c_{e2} as illustrated in Figure 5. Equilibrium number of retailers n^* has the following properties:*

- *When $c_e \leq c_{e1}$, the clustering equilibrium is the overlapping case and $n^* > \widehat{n}$.*
- *When $c_{e1} < c_e \leq c_{e2}$, the clustering equilibrium is the non-overlapping case and $s_N^* \leq n^* < \widehat{n}$.*
- *When $c_e > c_{e2}$, the market is not profitable and no retailers enter the market.*

Proposition 9 implies that for products with low entry costs (e.g., small commodities), the number of retailers is high. The market is fully covered and it is the overlapping case. For products with relatively high entry costs (e.g., high tech products), the number of retailers is relatively low. The market is partially covered and it is the non-overlapping case.

8. Conclusion

Using a Salop circle location model and a multi-stage game between heterogeneous retailers and consumers, we have provided some economic explanations for the wide-spread phenomenon of retail clusters in developing economies. We have also explored the social planner's perspective to explain when governments in these economies should foster these clusters and leverage them to improve social welfare.

We identify two determining factors for retail clusters: valuation-cost ratio and retailer density. These two factors jointly determine four types of clusters in equilibrium: overlapping massive clusters, non-overlapping large clusters, non-overlapping small clusters, and overlapping mini clusters. These four types of clusters correspond to products with low valuation-cost ratio and high

retailer density, products with low valuation-cost ratio and low retailer density, products with high valuation-cost ratio and high retailer density, and products with high valuation-cost ratio and low retailer density, respectively. Developing economies are more likely to have the first two types of clusters. Below, we review our key findings.

- The equilibrium cluster size decreases in the valuation-cost ratio (i.e., decreases in profit potential). This finding helps to explain the phenomenon that clusters in developing economies are usually large because retailers usually sell cheap, unrecognized brands with low profit potentials.

- When the retailer density is higher than a threshold, retailers form overlapping clusters.

- When compared to the non-overlapping case, the equilibrium cluster size is higher in the overlapping case for products with a low valuation-cost ratio but lower for products with a high valuation-cost ratio. This finding has opposite implications for developing and developed economies. In developing economies, cluster sizes for high-retailer-density products are higher than those for low-retailer-density products. On the contrary, in developed economies, cluster sizes for high-retailer-density products are lower than those for low-retailer-density products.

- The socially optimal cluster size is larger than the equilibrium cluster size. This finding suggests that governments should encourage larger retailer clusters.

- The gap between the socially optimal and equilibrium cluster sizes decreases in the valuation-cost ratio. This finding suggests that developing economies should be more aggressive in encouraging retailers to form larger clusters.

Our research can be extended in various ways. For example, we can consider more effects from clustering, such as the bazar effect. In other words, when the cluster size increases, retailers' cost decreases because of economics of scale (e.g., lower inbound transportation cost for suppliers sending supplies into the cluster). The bazar effect can be incorporated in our model by modifying retailers' cost to $c - bs$, where b is the degree of the bazar effect. However, making retailers' cost dependent on cluster size significantly increases the complexity of the model and makes it analytically intractable. Future work may resort to numerical analyses to examine the impact of the degree of the bazar effect on clustering results. In addition, it would be interesting to introduce retailer asymmetry. For example, retailers can be classified into different types such as major retailers (e.g., Best Buy) and small retailers. Future work may explore the impact of retailer asymmetry and retailer types on the clustering equilibrium.

Acknowledgments

References

Akçay Y, Tan B (2008) On the benefits of assortment-based cooperation among independent producers. *Production and Operations Management* 17(6):626–640.

Amonkar RV (2016) *Innovative solutions for implementing global supply chains in emerging markets*, chapter 2, 28–48. Advances in Logistics, Operations, and Management Science (Hershey, PA: IGI Global), 1st edition.

An J, Cho S, Tang CS (2015) Aggregating smallholder farmers in emerging economies. *Production and Operations Management* 24(9):1414–1429.

Bagwell S (2008) Creative clusters and city growth. *Creative Industries Journal* 1(1):31–46.

Berry S, Waldfogel J (2010) Product quality and market size. *Journal of Industrial Economics* 58(1):1–31.

Chen X (2015) China's key cities: From local places to global players. *The European Financial Review* October - November, URL http://www.europeanfinancialreview.com/?p=4980.

Damodaran A (2016) Datasets for cash flow estimation: Operating and net margins by industry URL http://people.stern.nyu.edu/adamodar/New_Home_Page/data.html.

Dawson J, Findlay AM, Paddison R (2005) *Retailing environments in developing countries* (London, United Kingdom: Routledge).

Deng C, Pekec A, Song JJ (2016) Teaming up for sustained performance: A repeated-game model of voluntary horizontal collaboration, working paper.

Desai PS (2001) Quality segmentation in spatial markets: When does cannibalization affect product line design? *Marketing Science* 20(3):265–283.

Du R, Lee E, Staelin R (2005) Bridge, focus, attack, or stimulate: Retail category management strategies with a store brand. *Quantitative Marketing and Economics* 3(4):393–418.

Dudey M (1990) Competition by choice: The effect of consumer search on firm location decisions. *American Economic Review* 80(5):1092–1104.

Eaton BC, Lipsey RG (1979) Comparison shopping and the clustering of homogeneous firms. *Journal of Regional Science* 19(4):421–435.

Elbert R, Schönberger R (2009) *Rapid Modelling for Increasing Competitiveness*, chapter Logistics Clusters - How Regional Value Chains Speed Up Global Supply Chains, 233–245 (London, United Kingdom: Springer).

Ellickson PB (2006) Quality competition in retailing: A structural analysis. *International Journal of Industrial Organization* 24(3):521–540.

Ellickson PB (2007) Does Sutton apply to supermarkets? *RAND Journal of Economics* 38(1):43–59.

Fischer JH, Harrington JE (1996) Product variety and firm agglomeration. *RAND Journal of Economics* 27(2):281–309.

Fujita M, Thisse JF (2013) *Economics of Agglomeration: Cities, Industrial Location, and Globalization* (Cambridge, UK: Cambridge University Press), 2 edition.

Gehrig T (1998) Competing markets. *European Economic Review* 42(2):277–310.

Gilbert RJ, Matutes C (1993) Product line rivalry with brand differentiation. *Journal of Industrial Economics* 41(3):223–240.

Goodman D (2011) The world's largest wholesale market in Yiwu, China. URL `http://www.businessinsider.com/yiwu-china-largest-wholesale-market-2011-10`.

Granot D, Yin S (2008) Competition and cooperation in decentralized push and pull assembly systems. *Management Science* 54(4):733–747.

Huang J, Leng M, Parlar M (2013) Demand functions in decision making: A comprehensive survey and research directions. *Decision Science* 44(3):557–609.

Huang X, Boyacı T, Gümüş M, Ray S, Zhang D (2015) United we stand or divided we stand? Strategic supplier alliances under order default risk. *Management Science* 62(5):1297–1315.

Konishi H (2005) Concentration of competing retail stores. *Journal of Urban Economics* 58(3):488–512.

Krugman PR (1991) *Geography and Trade* (Cambridge, MA: The MIT Press).

Latcovich S, Smith H (2001) Pricing, sunk costs, and market structure online: Evidence from book retailing. *Oxford Review of Economic Policy* 17(2):217–234.

Lee H, Yang HM (2003) Strategies for a global logistics and economic hub: Incheon international airport. *Journal of Air Transport Management* 9(2):113–121.

Marianov V, Eiselt HA (2014) On agglomeration in competitive location models. *Annals of Operations Research* 1–25.

McCann BT, Folta TB (2008) Location matters: where we have been and where we might go in agglomeration research. *Journal of Management* 34(3):532–565.

Morton FS, Zettelmeyer F (2004) The strategic positioning of store brands in retailer-manufacturer negotiations. *Review of Industrial Organization* 24(2):161–194.

Nagarajan M, Sošić G (2009) Coalition stability in assembly models. *Operations Research* 57(1):131–145.

Porter ME (1998) Cluster and the new economics of competition. *Harvard Business Review* 76(6):77–90.

Porter ME (2000) Location, competition, and economic development: Local clusters in a global economy. *Economic Development Quarterly* 14(1):15–34.

Reynolds M (2016) Inside Shenzhen: The Silicon Valley of hardware. URL `http://www.wired.co.uk/article/shenzhen-silicon-valley-of-hardware-wired-documentary`.

Salop SC (1979) Monopolistic competition with outside goods. *The Bell Journal of Economics* 10(1):141–156.

Schmitz H (1999) Global competition and local cooperation: Success and failure in the Sinos Valley, Brazil. *World Development* 27(9):1627–1650.

Sheffi Y (2010) Logistics intensive clusters. *Época* 20(1-2):11–17.

Sheffi Y (2012) *Logistics clusters: Delivering value and driving growth* (Cambridge, MA: MIT Press).

Stahl K (1982) Differentiated products, consumer search, and locational oligopoly. *Journal of Industrial Economics* 31(1/2):97–113.

Takahashi T (2013) Agglomeration in a city with choosy consumers under imperfect information. *Journal of Urban Economics* 76:28–42.

Wolinsky A (1983) Retail trade concentration due to consumers' imperfect information. *Bell Journal of Economics* 14(1):275–282.

Wu L, Yue X, Sim T (2006) Supply clusters: A key to China's cost advantage. *Supply Chain Management Review* 10(2):46–51.

Yin S (2010) Alliance formation among perfectly complementary suppliers in a price-sensitive assembly system. *Manufacturing & Service Operations Management* 12(3):527–544.

Yu M, Kapuscinski R, Ahn HS (2015) Advance selling: Effects of interdependent consumer valuations and seller's capacity. *Management Science* 61(9):2100–2117.

Zhao X, Stecke KE, Prasad A (2012) Lead time and price quotation mode selection: Uniform or differentiated? *Production and Operations Management* 21(1):177–193.