Heterogeneous consumers, search and retail formats

Mingming Shi[*] and Xiao Yan

* Correspondence: ms@ruc.edu.cn
Business School, Renmin University of China, Room 733, Mingde Building, N0.59, Zhongguancun St. Haidian Dist., Beijing, People's Republic of China

Abstract

This paper proposes a model featuring the heterogeneity of consumer preferences and analyzes the multiple equilibrium of retail formats by building a map of the relationships between consumer heterogeneity and retail formats. The key questions analyzed in this paper are how the retailer adjusts its combination of marketing elements through repositioning and how innovation in retail formats is implemented to match consumers' heterogeneous preferences in a market with consumer search costs. Unlike recent research, our model, by taking different consumer preference structures into account, introduces consumer psychological costs into the Ehrlich-Fisher model and deduces the existence of different retail formats and their multiple equilibriums. We find that consumer heterogeneity, retailers' diversified transfer costs and economies of scale are endogenous drivers of prosperous retail formats. Accordingly, diversified retail formats with complementary functions and differentiated services can be described as the horizontal extensions and interface changes of the retailing industry.

Keywords: Consumer heterogeneity, Retail formats, Search, Psychological cost

Jel: L81, L88, M21

Introduction

The retailing industry is continuously undergoing stunning innovations. These innovations include the transformation of economic and social environments with changes in consumer behaviors and changes in production patterns from mass production, namely Fordism, to flexible and mass customization production, namely post-Fordism. Under these processes, retailers, as market transaction organizers connecting producers and consumers, respond to the market sensitively, continuously changing retail formats.

The literature recognizes that the worldwide economic paradigm has undergone fundamental changes since the 1970s in contemporary society, that is to say the "economic paradigm" has transitioned from "Fordism" to "Post-Fordism". The typical characteristics of Fordism include a refined division of labor and economies of scale. Post-Fordism describes a form of production and organization as "flexible production" and "flexible division of labor". The concept of post-Fordism at present has gone far beyond production patterns, and has become an important paradigm to describe contemporary politics, economy and culture. In the economic field, in general, post-Fordism refers to a certain production and organization form with the purpose of satisfying personal requirements, based on information and communication technology, with a

flexible production process and labor relations. The main features of Post-Fordism can be described as follows.

(i) The mobility of consumer preferences, consumer heterogeneity and consumer power. In terms of post-Fordism, innovation activities, including technological and organizational, can result in different consumer preference structures. Consumer heterogeneity and the mobility of preferences create higher requirements for production and organization. Enterprises not only need to apply advanced scientific and technological knowledge constantly to create new products and industries, but also need to innovate quickly. In this atmosphere of innovation competition that was called "Creative Destruction" by Schumpeter, on the one hand, consumers' desires are satisfied and utility is maximized. On the other hand, enterprises are guaranteed adequate profit levels through technological improvements and the efficient use of resources. However, the monopolies enterprises have enjoyed are only temporary and cannot be sustained because of competition from substitutes and potential entrants.

(ii) The flexibility of organization and industry boundaries. In the classical and neoclassical paradigm, enterpriseswhich participates in market competition, are engaged in all sectors of the industry chain. However, in the new economic conditions, "the black box" is opened and the entire industry chain "fractures", which results in the fragmentation of various production processes. On the one hand, enterprises in these chains have strong fluidity. On the other hand, because the modular value realization depends more closely on the other links, the relationship of competition and cooperation has mainstreamed and the whole industrial activity has become an open, three-dimensional economic activity system.

(iii) Mass customization and modularization. Consumer heterogeneity and the rise of consumer power which gained by consumers to access to vast amounts of information and search abilities to influence their own lives have helped to bring about "product personalization". "Diversity of product structure" and the resulting "enterprise heterogeneity" become inevitable requirements. Different kinds of enterprises must maintain their value and respond to growing demands for personalization at the same time. Mass customization is the strategy that enterprises have to adopt to not only improve efficiency but also meet personalized demands. Modularization is the key to implementing mass customization.

(iv) The competition relationship between enterprises. With the fracturing of the industry chain, enterprises seek different market capacity and technical content in industrial links based on their own core assets / abilities, so that the coupling between different industrial sectors becomes more complicated. Therefore, the relationship between enterprises breaks through traditional homogeneous competition and enterprises instead have cooperative and competitive relationships simultaneously.

Roundabout production is a typical feature of economic systems in the post-Fordism Era in the production field. With the increase in industry levels between the initial resources and final consumption, the roundabout production chain gradually extends and the ability of industrial systems to provide differentiated, high quality end products is enhanced. Based on these roundabout and modular production activities, enterprises

can combine the module components produced by mass production and then assemble them into customized products or services. This means enterprises can implement a combination of advantages between the two modes of production – "mass production" and "customized production", which results in a unique production form called "mass customization". "Mass customization" meets modern consumers' heterogeneous demands in a precise and orderly, dynamic and coordinated way.

Consumer heterogeneity is a typical feature of a post-Fordism society from the view of demand. Labor divisions or roundabout production and the resulting standardization and modularization are the key points to understand the diversity of products and mass customization. Is this framework also suitable for the interpretation of the innovations of market transaction patterns? More fundamentally, how do we understand the differences between and co-existence of a variety of emerging retail formats? In this paper, we will consider consumer heterogeneity and provide a theoretical explanation to this question.

Retail formats are academically understood as the combination form of marketing elements such as commodities, prices, shops and sales, in order to provide consumers with varied retailing services. We think the key concept to understanding the diversity and co-existence of retail formats is consumer heterogeneity. In this paper, we use the structure of consumer preference to define the heterogeneity, and then start our analysis in a market environment with consumer search costs. Retailers can adjust their combination of marketing elements through repositioning and retail format innovation to match consumer behaviors. To model the above process, in this paper, we will establish $N \times 1$, $1 \times N$ and $N \times N$ models to observe and study the comprehensive effect of the interaction between heterogeneous consumers and retail formats.

There are two main contributions in this paper. The first is that we extend the evolution model of retail formats through integrating the work of Ehrlich and Fisher (1982) and Betancourt (2006). Our model introduces consumers' reactions to the retail service environment, namely consumer psychological costs, into the Ehrlich-Fisher Model. Through the revised model, we deduce the existence of different retail formats. The second contribution is that we demonstrate the possibility of the existence of multiple economic equilibriums by building a partial equilibrium model. This paper describes the distribution of the multiple equilibriums of retail formats and analyzes the evolution of equilibrium with changes in the external environment. These conclusions reveal how retail formats response to consumer heterogeneity.

Literature review

Recent research describes the components of retailing output as a set of services, such as location, information, assortment, delivery and ambience (Betancourt and Gautschi, 1990; Betancourt, 2006, Betancourt et al., 2007, 2016). Consumers are willing to pay for the benefits from lower transportation and search costs. For retailers, services such as location and ambience are like public goods. Consumers seek to minimize the cost of buying goods. Different retail formats provide different services. They usually do not charge these services directly but will cover the cost through providing goods. Hence, Kopalle et al. (2009) conclude that retailing has unique features that affect pricing in a competitive environment. Betancourt et al. (2007) analyze retail supply on the basis of consumer satisfaction. They study the influence that consumer well-being (consumer

satisfaction) has on supermarket supply levels and provide a relatively complete literature review of the effect that customer satisfaction has on the evolution of retail formats. Some scholars have also considered the innovation retailing business model that consumers search on the internet (Sorescu et al., 2011). They suggest that a successful retailing business model focuses not only on what a retailer sells, but also more importantly on how the retailer sells.

Notably, many economists and scholars also explore the innovation and evolution characteristics of the development of retail formats by considering changes in the social environment. For example, McNair (1957) elaborate the point of view that retail formats change periodically reacting to a complex background including culture, society, economy, law and so on. White and Cundiff (1978) point out that the innovation of retail formats is closely related to the level of economic development based on an empirical study of 20 countries. Cundiff also suggest that the methods of operation and technology of retail formats is a function of the surrounding environment. Based on the theory of the wheel of retailing, Deiderick and Dodge (1983) explain the relationship between the generation of new retail formats and logistics, information flow and innovation of management technology. The vacuum hypothesis put forward by Nielsen (1966) evaluates the change of retail formats directly from the perspective of consumer heterogeneity and notes that new retail formats generate from different consumer preferences for prices and retailer service levels.

However, very few theoretical studies consider both different consumer preference structures and the unique features of retailing. The key to describing the fact that retail format innovations satisfy the requirements of consumer heterogeneity is to establish the mapping relationship between consumer preference and retail formats. Ehrlich and Fisher (1982) implement this successfully. They unify service and price variables into the study of the relationship between consumers and retail formats formally and put forward the Ehrlich-Fisher model. This model connects the concept of retail formats with consumer cost creatively, i.e.: $PI_{ij} = P_{ij} + W_j I_{ij}$. In this model, P_{ij} is the commodity price that retail format i provides to consumer j. W_j is the cost per unit of time that consumers spend on getting commodities. I_{ij} is the time consumers spend on getting commodities. In the Ehrlich-Fisher model, the consumer cost PI_{ij} is equal to the sum of commodity price and time cost that consumers spend obtaining commodities. The Ehrlich-Fisher model starts from consumer cost, translates the cost consumers spend on getting commodities into the services retail formats provide and expresses the price and service combination in a formal theoretical model. Another noteworthy study is Betancourt (2006). He uses a specific function, in which the services that retail formats provide are an inverse function of consumer cost considering different consumer preferences, to build a mapping relationship between consumer preferences in a consumption environment and a service input stream from retailing providers.

Taking into account recent research progress, especially the contributions of Betancourt (2006), Betancourt et al. (2016), it may be an appropriate time to comprehensively model the mapping between consumer heterogeneity and retail formats by considering both different consumer preference structures and the unique features of retailing. This is the basic considerations and contributions of this paper.

Model

In the literature, some researchers recognize the interaction between consumer heterogeneity and retail formats. This paper will mainly reference research results by Nielsen, Ehrlich,

Fisher and Betancourt, in the following ways. Firstly, Nielsen put forward the concept of service and price combinations and uses consumer preference to explain retail formats. All of these contribute to our paper as a fundamental step in the research of service and price variables. Secondly, the consumer cost model that integrates retail formats by Ehrlich and Fisher provides the technology base for this paper, which will create a mapping of the relationship between consumer preferences and retail formats by extending the Ehrlich-Fisher model. Thirdly, we will use the technological processing method proposed by Betancourt (2006), the function $s = \varepsilon^{-1}(\zeta)$ to map the relationship between consumer preferences for the consumption environment and the service input stream from retail providers. Psychological factors play an important role in modern consumer behavior. For example, Urbany et al. (2000) and other scholars think when consumers are searching for something, they can obtain non-economic benefits called social psychological compensation such as feelings of happiness and a sense of satisfaction in addition to economic benefits. Based on the literature review, we construct a retail formats model from the consumer perspective, the core of which includes two core technical assumptions.

Consumer preference assumption

We assume that there is a typical consumer whose commodity portfolio preference is $\mathbf{x}$. The inner product of the commodity portfolio $\mathbf{x}$ and the price vector $\mathbf{p}$ denoted as $\mathbf{p} \cdot \mathbf{x}.j$ is the type of commodities; δ is the distance for accessing the commodities (or it can also be understood as the necessary cost per unit of time for purchasing the commodities). ζ is consumers' reaction parameter to the service environment, namely the consumers' psychological cost. This is the inverse function of the retail service environment or atmosphere, expressed as $\zeta_i = \varepsilon(s_i)$, which means the more services a retailer offers, the less psychological cost to consumers. Therefore, the function of consumers' real expenditure can be expressed as $e(\mathbf{p}, \mathbf{x}, j, \delta, \zeta)$. Obviously, we demand that $\delta \geq 0$, the cost of transportation paid by consumers to go to retail stores $c_c^i(\delta^i) \geq 0$, the longer the distance, the higher the cost to consumers $\frac{\partial c_c^i(\delta^i)}{\partial \delta^i} > 0$, $\frac{\partial^2 c_c^i(\delta^i)}{\partial \delta^{i2}} < 0$. When the psychological cost $\zeta^i < \zeta^j$, we say the transaction environment of i is more comfortable than j. Generally, we can assume that the environment satisfaction for initial trading is 0, so convenience can be expressed as ζ_-^i and inconvenience as ζ_+^i.

Transaction space assumption

We divide market space into two parts on the premise of not affecting the analysis. One part is consumer space and the other part is retailer space. The basic pattern is shown in Fig. 1. As shown in Fig. 1, the solid points C_1 to C_m on the left side represent consumers and on the right side, the solid points S_1 to S_n represent retailers. The solid lines represent transactions between consumers and retailers. For example, the line connecting C_1 and S_n represent a deal that consumer i might make with retailer n.

Diversity of retail formats

We firstly assume that consumer i's expenditure function is expressed in a linear form:

Fig. 1 Transaction space hypothesis
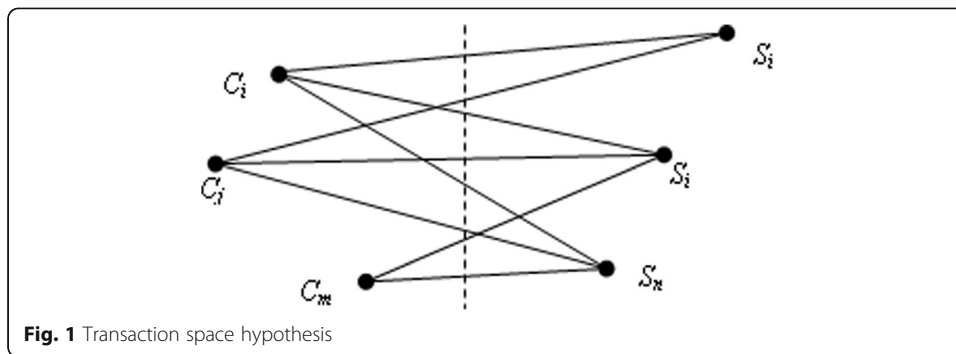
$$e(\mathbf{p}, \mathbf{x}, j, \delta, \zeta) = \mathbf{p} \cdot \mathbf{x} + \sum_{1}^{n} c(\delta_j) + \sum_{1}^{n} \zeta_j.$$

The initial transaction mode is cash on delivery among spatially dispersed individuals, in which we set the level of retail formats as zero. Under this assumption, the function of a consumer's expenditure can be written as $e_0 = \mathbf{p}\mathbf{x} + \sum_{1}^{n} c(\delta) + \sum_n \zeta_i$. At this time, we can obtain the condition for the existence of retail formats as follows:

If there exists a solution $\tau = (\mathbf{p}^*, \mathbf{x}^*, k^*, \delta^*, \varepsilon^*)$, which satisfies the condition

$$(\mathbf{p}_k - \mathbf{p}) \cdot \mathbf{x}_k + \left(\sum_k c - c \right) + \left(\sum_k \zeta - \zeta \right) \geq 0 \tag{1}$$

then we can say that there is a retail format existing that satisfies the condition. If the solution is not unique, we can say that the retail formats are diverse and the solution space, which satisfies the above formula, is the format space. Among them, $k \in (0, n]$ is the commodity portfolio which can be one-stop purchased, $\mathbf{x}_k$ is the proper subset of commodity space X. According to this definition, we get proposition 1.

Proposition 1

Proposition 1 deduces the supply conditions of existence for retail formats. If consumers need to purchase n commodities, then the sufficient condition of existence of some retail formats is that the purchase cost of k commodities is less than the decentralized purchase cost.

We can give a simple proof. Let $\zeta = 0$, then:

$$e_n = \mathbf{p}_n \cdot \mathbf{x}_n + \sum_n c = (\mathbf{p}_k \cdot \mathbf{x}_k + \mathbf{p}_{-k} \cdot \mathbf{x}_{-k}) + \left(\sum_k c + \sum_{-k} c \right) \qquad .$$

$$= e_k + e_{-k} \geq (\mathbf{p}_k \cdot \mathbf{x}_k + \mathbf{p}_{-k} \cdot \mathbf{x}_{-k}) + c(\delta) + \sum_{-k} c(\delta) = e_k' + e_{-k}$$

Simplifying the inequality above, we can know that $e_k \geq e_k'$ is permanent established.

Similarly, we can deduce the supply conditions of existence for retail formats as follows. If there exist organization forms which satisfy $t = \left(\mathbf{p}_{k^*}^*, \mathbf{x}_{k^*}^*, k^*, D^*, \varepsilon^* \right)$, then it means retail formats exist. Corresponding to consumers' surplus functions, as long as the consumer's willingness to pay is more than the cost of the purchase of products and the operation of organizations, we generally assume that there always exist organizations which can satisfy the cost constraints they have established under the condition of

ignoring interest rates. Such organizations only need to make entering and positioning decisions. Assuming that the decision function of enterprises is maximized, then the surplus function is $\pi = (\mathbf{p} - \mathbf{p}_0)\mathbf{x} - i - g(D)$. We assume that: (1) the entering function is $t = \left(\mathbf{p}_{j^*}^*, (\mathbf{p}_{st}, \mathbf{r}), \mathbf{x}_{j^*}^*, j^*, \delta^*, s^*\right)$ for the organization and its decision space is $t = (\mathbf{p}(\mathbf{p}_s, \mathbf{r}),$ $\mathbf{x}, \mathbf{j}, \boldsymbol{\delta}, \mathbf{s})$; (2) s^* is the level of services provided, which depends on the investment i; consumers' psychological cost $\zeta^* = \zeta(s^*)$, so $s^* = \zeta^{-1}(s^*)$; (3) the necessary conditions for the existence of organizations is $\pi \geq 0$, that is $(\mathbf{p} - \mathbf{p}_0)\mathbf{x} \geq i + g(D)$. At this time, commodities' selling even at a loss is permitted under the condition of overall profitability.

If the demand and supply conditions are satisfied, then the innovation of new retail formats is a typical Pareto improvement. If the corresponding price of the above commodity space $\mathbf{x}_j$ of retail formats is $\mathbf{p}_j^*$, then we name $\mathbf{p}_j^*$ as the overall price level of the retail format. When $\left(\mathbf{p}_j^* - \mathbf{p}_j\right) \cdot \mathbf{x}_j \geq 0$, we say the overall price level of the retail format is higher than the general shopping level. Vice versa, when $\left(\mathbf{p}_j^* - \mathbf{p}_j\right) \cdot \mathbf{x}_j \leq 0$, we say the overall price level of the retail format is lower than the general shopping level.

In order to simplify and focus the discussion, we assume the satisfaction of supply conditions from beginning to end in this section. Now we discuss the diversity of retail formats formally. First of all, adjusting formula (1), we can get $(\mathbf{p} - \mathbf{p}_0) \cdot \mathbf{x}_j \leq \sum_{i=1}^{k} c\left(\delta^i\right) - c(\delta^*)$. If $(\mathbf{p}^* - \mathbf{p}) \cdot \mathbf{x} \geq 0$, when the inequality below is satisfied:

$$\left|\left(\mathbf{p}_{j^*}^* - \mathbf{p}_j\right) \cdot \mathbf{x}_{j^*}^*\right| \leq \sum_{i=1}^{j} c_c\left(\delta^i\right) - c_c(\delta^*) - \zeta^* \tag{2}$$

then retail formats always exist. Moreover, the solution is not unique according to the continuity of the function. Among them, the value space of ς^* is:

$$\left[-\infty, \sum_{i=1}^{j} c_c\left(\delta^i\right) - c_c(\delta^*) - \left|\left(\mathbf{p}_{j^*}^* - \mathbf{p}_j\right) \cdot \mathbf{x}_{j^*}^*\right|\right].$$

Thus, we can deduct the following proposition:

Proposition 2
Proposition 2 deduces under which conditions retail formats will be diverse. When the overall price level of the retail formats is higher than the general price level, if condition (2) is established, there are diverse retail formats.

If $(\mathbf{p}^* - \mathbf{p}) \cdot \mathbf{x} < 0$, when the following formula is established:

$$-\left|\left(\mathbf{p}_{j^*}^* - \mathbf{p}_j\right) \cdot \mathbf{x}_{j^*}^*\right| \leq \sum_{i=1}^{j} c_c\left(\delta^i\right) - c_c(\delta^*) - \zeta^* \tag{3}$$

then the retail formats always exist. At this time, the value space is.

$$\left[-\infty, \sum_{i=1}^{j} c_c\left(\delta^i\right) - c_c(\delta^*) + \left|\left(\mathbf{p}_{j^*}^* - \mathbf{p}_j\right) \cdot \mathbf{x}_{j^*}^*\right|\right].$$

Proposition 3
Proposition 3 further discusses two situations to explain how retailer commodity categories and consumer psychological costs create diversified retail formats. When the

overall price level of the retail format is lower than the general price level, if condition (3) is established, under this condition retail formats will be diverse.

This can be divided into two situations: one is where the overall price is higher than the general price level. For the convenience of the following discussion, we assume the function of distance cost as a linear function, that is $c_c(\delta) = \alpha_0 \cdot \delta$ ($\frac{\partial c_c(\delta)}{\partial \delta} = a_0 > 0$ is satisfied). Because $\left| \left(\mathbf{p}_{j^*}^* - \mathbf{p}_j \right) \cdot \mathbf{x}_j^* \right| \le a_0 \sum_{i=1}^{j} \delta^i - a_0 \delta^* - \zeta^*$, let $\Delta = \left| \left(\mathbf{p}_{j^*}^* - \mathbf{p}_j \right) \cdot \mathbf{x}_{j^*}^* \right| / a_0$, then $\delta^* \le \sum_{i=1}^{j} \delta^i - \frac{\zeta^*}{a_0} - \Delta$.

Another is when the psychological cost is less than 0, we can conclude that $\delta^* \in \left[0, \sum_{i=1}^{j} \delta^i + \left| \frac{\zeta^*}{a_0} \right| - \Delta \right]$. Among them, the bigger $|\zeta^*|$ and j are, the bigger δ^* and the value space δ^* will be. Thus, we can deduce the following proposition:

Proposition 4

Proposition 4 points out why franchised stores and specialty stores conform to the derivation conditions. There exists a type of retail format which satisfies the following conditions: relatively complete types of merchandise, a relatively high price level, service of high quality and remote distance (longer than the whole distance of decentralized purchasing). If the service level that the format provides is higher and more convenient for consumers, then the space and business district for the format will be bigger.

This explains the existence of urban shopping centers, supermarkets and hypermarkets, etc. Under the condition that the overall price level of the retail format is higher than the general price level, when types of merchandise $j \to k$ (i.e., one-stop shopping), the leading solution has two feature variables: one is the difference in the price level of commodities Δ, and the other is service supply ζ^*. The higher the service supply is, the higher the overall price level that the format can provide will be. When the service supply is low, the overall price level of the format also has to be lower.

When the types of merchandise $j \to 1$, the above value space tends to $\left[0, \delta + \left| \frac{\zeta^*}{a_0} \right| - \Delta \right]$, which means that there exists a type of format that satisfies the following conditions: relatively fewer types of merchandise, a relatively high price level, service of high quality and close distance (shorter than the whole distance of decentralized purchasing), which explains the existence of franchised stores and specialty stores.

When the psychological cost is bigger than 0, we can conclude that $\delta^* \in \left[0, \sum_{i=1}^{j} \delta^i - \left| \frac{\zeta^*}{a_0} \right| - \Delta \right]$. We can deduce the following proposition:

Proposition 5

Proposition 5 shows why grocery stores, automatic kiosks, mom and pop store and convenience stores conform to the derivation conditions.

(1) When $\zeta^* \ge a_0 \sum_{i=1}^{j} \delta^i$, the retail format will not exist.

(2) When $\zeta_+^* \le a_0 \sum_{i=1}^{j} \delta^i$, the existence of the retail formats depends on both types of merchandise j and ζ_+^*.

When $j \to k$, there exists a type of retail format that satisfies the following conditions: relatively complete types of merchandise, a relatively high price level, low quality service

and close distance, which explains the existence of the grocery store. When $j \to 1$, there exists a type of retail format that satisfies the following conditions: relatively fewer types of merchandise, a relatively high price level, low quality service and close distance, which explains the existence of automatic kiosks, mom and pop stores and convenience stores.

The second situation is when the overall price level of the retail format is lower than the general price level. At this time, according to condition (1), we can conclude that:

$$-\left| \left(\mathbf{p}_{j^*}^* - \mathbf{p}_j \right) \cdot \mathbf{x}_{j^*}^* \right| \le a_0 \sum_{i=1}^{j} \delta^i - a_0 \delta^* - \zeta^*,$$

i.e., $a_0 \delta^* - a_0 \sum_{i=1}^{j} \delta^i + \zeta^* \ge -a_0 \Delta$, at which time $\delta^* \ge \sum_{i=1}^{j} \delta^i - \frac{\zeta^*}{a_0} - \Delta$ and there is a negative correlation between the value of δ^* and the quality of service and price level of the retail format. We can deduce the following proposition:

Proposition 6

Proposition 6 describes why factory outlet centers, suburban storage centers and brand discount stores conform to the derivation conditions.

(1) When $j \to k$ (one-stop shopping), there exists a type of retail format that satisfies the following conditions: relatively more types of merchandise, a relatively low price level, medium quality service and remote distance.

This can explain the existence of the factory outlet center, which is a kind of shopping center. Factory outlet centers collect brand stores of various manufacturers and sell brand commodities at a discount.

(2) When $j \to 1$, there exists a type of retail format that satisfies the following conditions: relatively fewer types of merchandise, a relatively low price level, service of medium quality and remote distance, which explains the existence of the suburban storage center.

(3) When $j \to 1$ and the service is of high quality, there exists a type of retail format that satisfies the following conditions: relatively fewer types of merchandise, a relatively low price level, high quality service and remote distance, which explains the existence of the brand discount store.

Expanding analysis I

More complex retail formats with economies of scale

We have shown that there exists a variety of retail formats in response to the constraint of consumers' budgets. In this section, we will consider the situation of n consumers who purchase together which means that the types of existing retail formats will be constrained by consumer preference structures and the number and types of commodities.

We define that $C_1, C_2 ..., C_n$ represent n consumers, and S represents the seller. Consumers can only purchase commodities or services from S. At this time, we can express the purchase cost of consumers with the same preference structure as:

$$\sum e = \sum_{i=1}^{n} \mathbf{p} \cdot \mathbf{x} + \sum_{i=1}^{n} c(d) + \sum_{i=1}^{n} \varsigma = \mathbf{p} \cdot \sum_{i=1}^{n} \mathbf{x} + \sum_{i=1}^{n} c(d) + n\varsigma.$$

Therefore, we find that although consumers' preference structures are different, the retailer may have an economy of scale, namely the efficiencies formed by volume, because the overall commodities consumers purchase are $\sum_{i=1}^{n} \mathbf{x}$. Here, we study the decision function of the retailer:

$$\text{Max } \pi = (\mathbf{p} - \mathbf{p}_0)\mathbf{x} - i - g(D) = \mathbf{p} \cdot \mathbf{x} - [\mathbf{p}_0 \cdot \mathbf{x} + i + g(D)] = R - C.$$

We assume that there are two standardized consumers, who choose to purchase $\mathbf{x}_1$ and $\mathbf{x}_2$. Assuming that the two satisfy $\mathbf{x}_1, \mathbf{x}_2 >> 0$, the cost function of the enterprise is expressed as:

$$C^0(\mathbf{p}, \mathbf{x}_1 + \mathbf{x}_1, i, g) = \mathbf{p}_0(\mathbf{x}_1 + \mathbf{x}_2) + i + g$$

$$\leq \mathbf{p}_0(\mathbf{x}_1 + \mathbf{x}_2) + 2i + 2g = (\mathbf{p}_0\mathbf{x}_1 + i + g) + (\mathbf{p}_0\mathbf{x}_2 + i + g)$$

$$= C^1(\mathbf{p}_0, \mathbf{x}_1, i, d) + C^2(\mathbf{p}_0, \mathbf{x}_2, i, d).$$

That is, when there exist two consumers, to the enterprise, $C^0(\mathbf{p}, \mathbf{x}_1 + \mathbf{x}_1, i, g) \leq C^1(\mathbf{p}_0, \mathbf{x}_1, i, d) + C^2(\mathbf{p}_0, \mathbf{x}_2, i, d)$. We can naturally extend the above conclusion to the situation with N consumers. This means the enterprise has cost sub-additivity to some extent, which means it has an economy of scale as well according to the definition. So we have the following proposition:

Proposition 7

When there are n consumers, the retailer has an economy of scale.

At this time, $C^0\left(\sum_{i=1}^{n} \mathbf{x}_i, D\right) \leq \sum_{i=1}^{n} C^0(\mathbf{x}_i, D)$. We can see that the economy of scale, which the retailer has, is mainly due to sunk cost i (retail environment) and g (location), which were invested in to improve the transaction conditions. Nevertheless, when consumers are heterogeneous ones, we also find the following proposition:

Proposition 8

The retailer has not only an economy of scale but also an economy of scope and economy of network.

A simple proof is given as following. We assume that the consumer assemblage consists of two consumers, whose demand for commodity 1 and 2 is $(x_1, 0)$ and $(0, x_2)$. At this time, the retailer's decision of extending the sale of only commodity 1 to both commodity 1 and 2 will be lucrative. Then his category portfolio is (x_1, x_2), namely $C^0(x_1, x_2, D) \leq C^0(x_1, 0, D) + C^0(0, x_2, D)$. This conclusion can also be extended to the situation with n consumers, i.e.:

$$C^0(x_1, x_2, ..., x_n, D) \leq C^0(x_1, 0, ..., 0, D) + C^0(0, x_2, ..., 0, D) + ... + C^0(0, ...x_n, D).$$

Obviously, the above result is due to the economy brought by the enterprise's increase in categories of merchandise. This type of economy is defined as the economy of scope because the efficiencies are formed by variety, not volume. In fact, another side of the above result is the improvement of consumer welfare. We add a marginal consumer on

the base of the above two consumers, whose demand for commodity 1 and 2 is (x_3, x_4). If he purchases the two types of commodities separately now, the potential cost is expressed as:
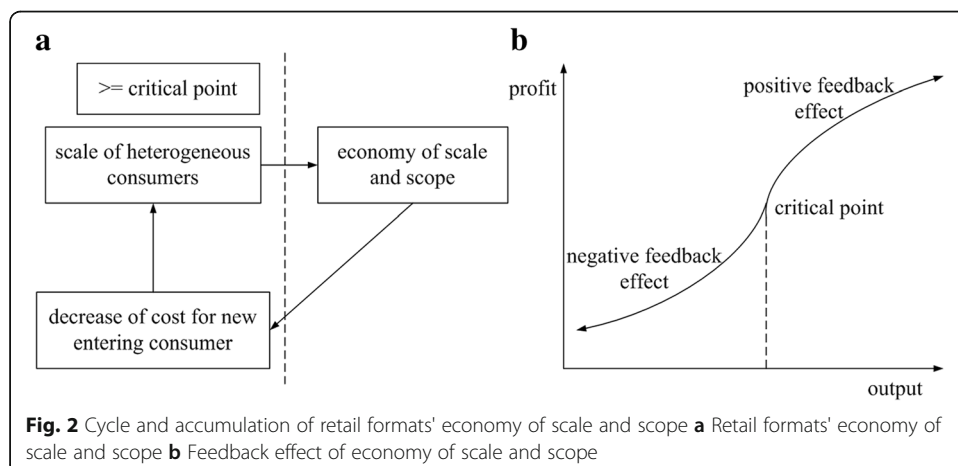
$$e_3 = (p_1 p_2)\begin{pmatrix} x_3 \\ x_4 \end{pmatrix} + [c(d_1) + c(d_2)] + (\varsigma_1 + \varsigma_2).$$

At this time, if the retailer increases the supply of commodity 1 and 2 to (x_1', x_2'), among which $x_1' = x_1 + x_3, x_2' = x_2 + x_4$, we can find that the third consumer only needs to purchase from S to meet his demand when he accesses the market. At this point, for the third consumer there exists:

$$\begin{aligned} e(x_3, x_4) &= (p_1 p_2)\begin{pmatrix} x_3 \\ x_4 \end{pmatrix} + c(d_1) + \varsigma_1 \le (p_1 p_2)\begin{pmatrix} x_3 \\ x_4 \end{pmatrix} + [c(d_1) + c(d_2)] + (\varsigma_1 + \varsigma_2) \\ &= [p_1 x_3 + c(d_1) + \varsigma_1] + [p_2 x_4 + c(d_2) + \varsigma_2] = e(x_3, 0) + e(x_4, 0). \end{aligned}$$

The essence of the above inference is if the retailer increases its categories, it will attract new consumers to enter, the net effect of which is both the realization of economies of scale and scope for the manufacturer and the decrease of costs for consumers. On the side of the consumers, we can express the result as: due to the increase in categories of merchandise that consumers demand and the formation of a certain scale which break the critical point, the retailer may be willing to expand its scope of business. Expanding the scope of business makes consumers narrow the search range and can purchase commodities that previously needed to be bought in many shops in only one store now, which ultimately reduces the purchase cost. The involvement of consumers further enlarges the advantage of economies of scale and increases sales for related products, which improves the performance of the enterprises. This creates a positive feedback loop and circular and cumulative effects, which ultimately appear as an economy of scale on the consumer side, namely the effect of a network economy. This cumulative effect of this cycle can be expressed in Fig. 2.

As shown in Fig. 2 (a), the result of the above process is that enterprises with good category structure and large scale will form benign interactions with consumers. Once breaking a critical point, it will form constant positive feedback, which leads to lower costs for consumers and more competitive advantages for retailers. At the same time,

Fig. 2 Cycle and accumulation of retail formats' economy of scale and scope **a** Retail formats' economy of scale and scope **b** Feedback effect of economy of scale and scope

the competitive disadvantages for enterprises with poor category structure and small scale will be more obvious because of being unable to realize economies of scale and scope. Above is a typical process of "winner-take-all" (typical market results shown in Fig. 2 (b)). This can explain how modern hypermarkets and shopping malls often have a network effect, thus demonstrating "positive feedback" and "negative feedback" effects, which also means that this type of format has a very strong "polarization effect" to prevent similar types of shops from surviving around the supermarket. Even more, it is possible to form a "shop in shops" in this kind of store in order to draw support from economies of scale of consumers.

### Spatial agglomeration and the shopping street

Synthesizing both of the sections above, we further consider the situation with n consumers and n retailers. In this situation, the types of retail formats are more complicated and there is another form of retail spatial agglomeration, namely the shopping street. In the first section, we find that there exist diversified retail formats to satisfy the requirements of consumers who have fixed expenditures or balanced budgets, on which their preference structure depends. In the second section, we find that there may be a positive feedback effect between retailers and consumers. The situation considering both effects above are: if n is big enough, then all of the different retail formats may coexist. Some retailers can take advantage of not only economies of scale but also the increase in categories. In these circumstances, consumers can enjoy the benefits not only from retailers' economies of scale and scope, but also from a network economy. At this point, a positive feedback effect between aggregated retailers and consumers can also occur.

Expanding analysis II

Search costs and the multiple equilibrium of retail formats

Search and multiple equilibrium

From the sections above, we can see the complex mapping relations between consumer heterogeneity and retail formats. Consumer heterogeneity is an important foundation for the existence of retail formats. We can apply the standard supply-demand framework to express the equilibrium feature of retail formats concisely. First of all, from the consumer's point of view, we analyze consumer behavior $e_0 = \mathbf{p} \cdot \mathbf{x} + c(\delta) + \zeta$ in a situation where total real spending is established, then we can deduce that $\zeta = e_0 - \mathbf{p} \cdot \mathbf{x} - c(\delta) = e_0 - m - c(\delta)$, which is $\zeta(m) = e_0 - m - c(\delta)$ in the functional form. Among them, m is consumer expenditure. When the consumer's total budget is balanced, he can weigh between expenditure, distance cost (or time cost) and psychological cost. We express the consumer's trade-off between distance cost, psychological cost and monetary expenditure in Fig. 3 (a). As is shown in Fig. 3 (a), with the preference of less distance cost, if the consumer purchase item with the value M1, he can only enjoy (poor) environment Z1. However, with the preference of more distance cost, he can enjoy (more elegant) environment Z2. Similarly, if the consumer does not mind environment Z1, but wants to reduce the distance cost (time cost) for searching, then the money of his willingness to pay to purchase the item is M2. A and C show the consumer's trade-off among distance cost, psychological cost and monetary expenditure with the total expenditure unchanged.

Fig. 3 Consumers' trade-off and multiple equilibrium **a** Consumers' trade-off **b** Multiple equilibrium of retail formats
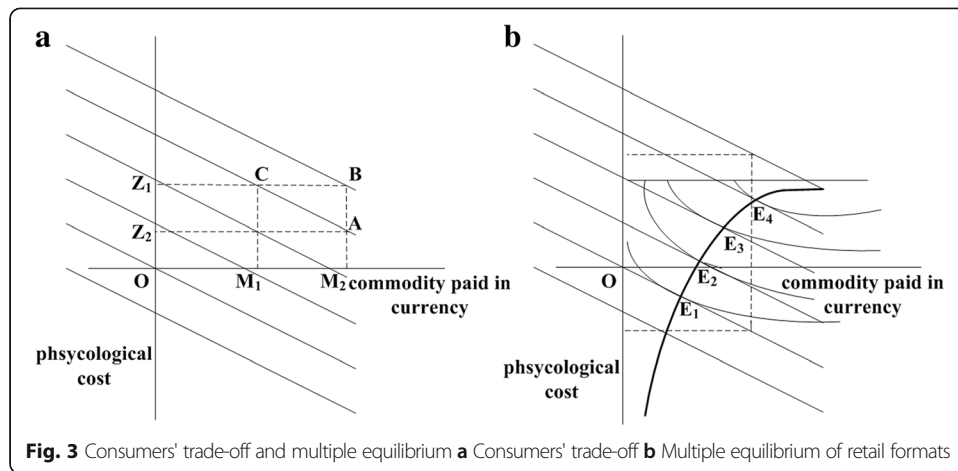
Similarly, for retailers, because $\pi = (\mathbf{p} - \mathbf{p}_0)\mathbf{x} - i - g(D)$, in the case of profit, is established, we assume the prime cost $\mathbf{p}_0 = 0$ without affecting the analysis. Then the above formula can be transformed into: $i = m - g(D) - \pi_0$. At this time, the relationship between investment in environment and sales is a kind of positive correlation, that is, the more investment in improving the environment, the more the sales will be. Further considering the relationship between retailers' investment environment and the feeling consumers have about the environment, due to the existence of the utility diminishing effect of the investment, we can set them into a binary relation, namely $i(\varsigma) = (a - \varsigma)^2 = m - g(D) - \pi_0$ (a > 0). We express the above relationship between consumers and retailers in Fig. 3 (b). In Fig. 3 (b), we find that retail format E1 means consumers choose a specific consumption environment, purchase amount and distance under the condition that the consumer's budget is balanced and the retailer's expected profit is fixed. E4 means that consumers prefer to purchase more commodities but have to endure a poor trading environment, close distance, etc. This time space (E1, E2, E3, E4, …) makes up a multiple equilibrium of consumers' and retailers' decisions. Because every point in the space represents a group of choices (environment, distance, commodity), and this forms a format space according to the concept.

The corresponding relationship between consumers' total real expenditure structure and retailers' profit structure embodies a basic cost transfer in the above equilibrium. Many scholars have conducted similar studies, such as the pioneering research conducted by Salop and Stiglitz (1977), which says that an insufficiency of information leads to higher or lower search costs for consumers, which causes the formation of different equilibrium prices. Another example is Oi (1992) who construct the model of choice of location and store from retailers' point of view. He finds that in a situation with the same full price and competition of stores, consumer heterogeneity in locations and other purchase costs may be the main source of the diversity of merchandise categories. He also speculates that there may be an equilibrium solution. This paper makes clear that consumer heterogeneity is an important reason for the diversity of retail formats and multiple equilibrium and different prices, locations and diversity of merchandise categories in different retail formats are all equilibrium characteristics. Consumer preference heterogeneity and distribution, retailers' differentiated cost transfer and its spillover effect of scale (economies of scale and scope) are the

endogenous drivers that promote the prosperous development of diversified retail formats.

Evolutionary equilibrium of retail formats

We further consider how changes in consumer preference and technology elements affect the evolution of trading mechanisms under the condition of a transition economy based on the above model. Firstly, we introduce structural preference parameters to explain the phenomenon better. According to the above model of spending function under consumers' trading expenditure, we can set consumers' structural preference parameters as γ_1, γ_2 and γ_3, i.e.:

$$E = \gamma_1 \mathbf{p} \cdot \mathbf{x} + \gamma_2 \sum_1^n c(\delta) + \gamma_3 \sum_1^n \zeta = (\gamma_1, \gamma_2, \gamma_3)\left(\mathbf{p} \cdot \mathbf{x}, \sum_1^n c(\delta), \sum_1^n \zeta\right)^T.$$

Among them, $\theta = (\gamma_1, \gamma_2, \gamma_3)$ represents a structural preference matrix, in which γ_1 represents the consumer preference on the unit of currency, γ_2 is on behalf of the consumer preference on the distance cost, and γ_3 represents the consumer preference on the trading environment. Since consumers' own endowment will affect the relative magnitude, we can use $(\gamma_1/\gamma_2, \gamma_2/\gamma_3, \gamma_3/\gamma_1)$ to represent the relative strength or alternative relationship between them.

In the historical context, different changes in parameters will deeply influence the types, properties and distribution of retail formats. For example, at certain points in time, some people may prefer to have cash on hand and have low requirements for trading environments influenced by their culture or family's low-income status. Therefore, consumers will be more sensitive to the price and quantity of commodities and pay little attention to the trading environment. At this time, retail formats with a simple layout, complete categories and long distance will exist. Similarly, when at a certain point in time, people may prefer the elegance and comfort of an instant trading environment and hate the long distance influenced by different cultures, then retail formats within short distances with attractive layouts will develop.

In a transition economy, dramatic change in social preference structures is a common phenomenon. Economic transition is also accompanied by political, cultural and social transformation. People's internal preferences and values, such as hobbies, tastes, habits, behaviors, etc. experience considerable changes in a relatively short period, which have a lasting and significant influence on retail forms and impact the stability of the basic format structure.

The above model reflects the change in technical parameters $c(\cdot)$ and ζ at the consumer level. At a certain point in time, because of the underdevelopment of traffic technology, the sensitive coefficient of characterization approaches infinity and people's sensitive coefficient of consumption environment is low. Assuming that it is close to 0, we can see that retail space is mainly distributed in areas close by and the main determinants of retail formats are type and quantity, linked to the homogeneity of people's preferences. Retail space expands rapidly along with the improvement of transportation conditions. In the transition economy of China, car ownership per person has increased rapidly with the swift rise in national income. At this time, trading activity has less and less dependence on $c(\cdot)$ and has increased demands on the environment ζ. The import of information technology allows the retailers to make better use of huge dynamic

database systems to collect, handle and interpret information on consumer preference, so as to provide more diversified and personalized services.

Conclusion

Consumer heterogeneity is one of the basic features of contemporary society, and has a profound impact on all aspects of the economy. Scholars have fully realized that roundabout and flexible production based on modularization is an important response to consumer heterogeneity in the post-Fordism Era. In this paper, we use the research of Nielsen, Ehrlich and Fisher, Betancourt and other scholars as reference to try to map the relationships between consumer preferences and retail formats. Carefully observing and understanding this mapping relationship, we find that consumer heterogeneity also plays a key role in the prosperous development of retail formats.

The study shows that diversified retail formats and multiple equilibria are important ways for the retail industry to respond to consumer attributes in the post-Fordism Era. The prosperity, diversity and complexity of retail formats and the flexibility and roundabout of the production field are essentially different aspects of the same process of adapting to consumer heterogeneity. However, due to the different characteristics of the production and transaction fields, their performance has different characteristics. In a certain sense, a natural production process addresses demand via labor division, cooperation and business value combination, described as vertical changes in production chains and networks. Accordingly, diversified retail formats, complementary functions and differentiated services, can be described as horizontal extensions and changes of interface. Thus, we can understand that the rich development of contemporary retail formats is not only an important economic but also an important social and historical phenomenon.

This research has multiple policy implications. The first implication is that the government's retail industry regulation policies should actively guide the diverse, orderly and healthy development of retail formats. Consumer heterogeneity is the endogenous driver of the diverse formats and there is a dynamic equilibrium relation between varieties of retail formats to some extent. The second implication is that we should fully understand the economic effect of different formats and develop different regulatory policies for different cases. For example, large general retail stores may have characteristics such as scale, scope and network economies and display a superimposed effect of "winner take all", which means that the government's policy should follow the principle of classified regulation and explore targeted regulations, and strengthen the relevance and effectiveness of regulation policy according to the characteristics of different formats.

This research is also important for retailing industry practices. Consumer heterogeneity makes value created by traditional retail for customers including price, category diversity, service, experience and entertainment. Among them, price and category diversity are the advantages of online purchase; however, service, experience and entertainment are the advantages of existing physical stores. This study also concludes that traditional retail enterprises' innovation should create value for consumers based on new business models. For example, the advantage of department stores is that they have offline stores, and most of them are in the commercial center and core business district of a community. The way to enhance the customers' satisfaction will also help to increase customers' loyalty. "Small and beautiful" means to optimize the

adjustment of the department store category, through the selection of commodity categories, selection of brands, and the positioning of high-end fashion experience shopping.

Availability of data and materials
Not applicable.

Authors contributions
All authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References
Betancourt, R., & Gautschi, D. (1990). Demand complementarities, household production, and retail assortments. *Marketing Science, 9*(2), 146–161.
Betancourt, R. R. (2006). *The economics of retailing and distribution*. Northampton, MA: Edward Elgar Publishing Ltd..
Betancourt, R. R., Chocarrob, R., Cortiñasb, M., Elorzb, M., & Mugica, J. M. (2016). Channel choice in the 21st century: The hidden role of distribution services. *Journal of Interactive Marketing, 33*, 1–12.
Betancourt, R. R., Cortinas, M., Elorz, M., & Mugica, J. M. (2007). The demand for and the supply of distribution services: A basis for the analysis of customer satisfaction in retailing. *Quantitative Marketing and Economics, 5*(3), 293–312.
Deiderick, E. T., & Dodge, H. R. (1983). The wheel of retailing rotate and moves. In J. Summery et al. (Eds.), *Marketing: Theories and concepts for an era of change*. Carbondate: Proceedings Southern Marketing Association.
Ehrlich, I., & Fisher, L. (1982). The derived demand for advertising. *American Economic Review, 72*(3), 366–388.
Kopalle, P., Biswas, D., Chintagunta, P. K., Fan, J., Pauwels, K., Ratchford, B. T., & Sills, J. A. (2009). Retailer pricing and competitive effects. *Journal of Retailing, 85*(1), 56–70.
McNair, M. (1957). Significant trends and developments in the Postwar period. In A. Smith (Ed.), *Competitive distribution in a free high level economy and its implications for the university*. Pittsburgh (PA): University of Pittsburgh Press.
Nielsen, O. (1966). Developments in retailing. In M. K. Hansen (Ed.), *Readings in Danish theory of marketing*. Amsterdam: North Holland Publishing Company.
Oi, W. (1992). Productivity in the distributive trades: The shopper and the economies of massed reserves. In Z. Griliches (Ed.), *Output measurement in the service sectors*. Chicago: University of Chicago Press.
Salop, S., & Stiglitz, J. (1977). Bargains and rip-offs: A model of monopolistically competitive price dispersions. *Review of Economic Studies, 44*(3), 493–510.
Sorescu, A., Frambach, R. T., Singh, J., Rangaswamy, A., & Bridges, C. (2011). Innovations in retail business models. *Journal of Retailing, 87*(s1), S3–S16.
Urbany, J. E., Dickson, P. R., & Sawyer, A. G. (2000). Insights into cross- and within-store price search: Retailer estimates vs. consumer self-reports. *Journal of Retailing, 76*(2), 243–258.
White, P. D., & Cundiff, E. W. (1978). Assessing the quality of industrial products. *Journal of Marketing, 42*(1), 80–86.