1. INTRODUCTION

Segmentation and its benefits can never be undermined. An article as early as in 1956 (Smith.W.R, 1956) propagated the idea of segmenting the market to effectively design marketing strategies. In order to sharpen the marketing mix, companies have moved from marketing at the aggregate level, called Mass Marketing, to Mass Customisation, the level of 1-to-1 marketing, where, each customer is treated as a separate segment. The factors that have weighed heavily on this decision are, the need to fine-tune the marketing strategies, and the realisation that the cost of retaining (read servicing) an existing customer is much lower than acquiring a new customer. Customer acquisition and retention for a retailer are more critical and difficult for the following reasons:

1. Most retailers have a defined physical catchment area, defined by the distance that most of its shoppers would travel to visit the store. Although it differs with the kind of product bought by the shopper, there is a significant impact on the frequency of visit to the store. Thus, a store has to work within a given geographical area with a radius of not more than 5 – 6 kilometres.

2. Not every retailer is able to offer merchandise that is very different, nor is the positioning of the store differentiated; and

3. Every visit of the shopper is an encounter and a moment of truth. Unless the interaction has been satisfactory, the next visit is not guaranteed. In case the store

does not provide a compelling reason for visit, the amount of purchase per visit is likely to go down.

Under this condition, it is imperative that the store understands the shoppers it wants to attract and keep. Since retailing is a service business, and is not just distribution, there are more "soft" than "hard" variables that would lead to creation of a customer franchise. It is not only important as to what is being retailed, it is equally, may be more important as to how it is being retailed. This condition shifts the focus from the merchandise being handled, to the processes that make shopping a memorable experience. All the retailers that believe in this paradigm would require that shoppers be segmented on the basis of attitude and behaviour than on demographic. The latter is a pre-condition that works as eliminator, and not describing the segment the retailer would serve. Also, a retail outlet is likely to be patronised more when it becomes a part of the life of the shopper. A major asset of store retailers is their ability to appeal to consumers using the physical environment. Retailers could modify or enhance their physical environments, while providing needed goods and services, to attract consumers to their stores. Often the retailer appeals to all five senses (i.e., sight, sound, scent, touch, taste) to entice consumers and to create a favourable store image (Lewison, 1994). Regardless of the sensory appeals used, store retailers should create images that will be suitable for their target consumers and that will influence consumer patronage. Store retailers can exploit these sensory appeals to compete with non-store retailers that lack access to these consumer attributes, as well as have a competitive advantage over stores that fail to meet the needs of their consumers.

Stores that present pleasing and attractive physical environments through sensory appeals will create a competitive environment among stores and are most likely to entice more consumers to patronize their stores. The decision to patronize a store usually starts with a set of characteristics or attributes that consumers consider important. Consumers often use these attributes to determine which stores can cater to their needs. Patronage studies in the past have attempted to identify determinants of store choice in relation to attributes, such as, price, quality, variety of merchandise, credit availability, return policies, and well-known labels/brands (e.g., Jones of New York, Liz Claiborne). Physical facilities or attributes have been included among other attributes in some studies, but they have not been fully addressed in these studies. In general, no research to date has focused on which of those environmental dimensions are important in choosing a retail store and how the physical environment affects patronage behaviour. According to Engel, Blackwell, and Miniard (1995) environmental dimensions such as air quality, lighting, layout, carpeting, aisle width and placement are physical store attributes used to project store image, and influence store choice. Today‘s retail market is characterized as being more competitive; thus, store retailers must develop effective strategies to gain a competitive advantage. To succeed in a competitive market, retailers must understand that, as the market changes, so will their target consumers. Segmenting consumers into patronage groups, based on one particular store-type for all purchases is not logical. Consumers may shop at several retail store types depending on their behaviour towards various store images. Therefore, creating a new method of segmenting the retail consumer, based on the

behavioural pattern of shoppers towards the overall store images, with specific shopping scenarios, is a significant area of research.

1.1 SIGNIFICANCE OF THE STUDY

This study develops an enhanced method of segmenting shopping tourism retail market based on behavioural pattern, using fuzzy clustering method. This is important for several reasons. First, retailers attempting to determine why consumers shop at their store will find this research important. Empirical studies such as this will provide retailers with the tools necessary to attract, target and retain consumers. Second, this research will contribute to a better understanding of the store environmental needs of consumers by store type. Third, the cluster validity functions help retailers to understand and formulate the market boundaries, that is, the number of segments can be determined. Fourth, the fuzzy cluster solutions will reflect the real market situation in several means; the hard cluster solution will result in crisp partitioning of members, which means, a member who belongs to one cluster will not belong to the other cluster or segment, whereas the results of fuzzy cluster solution will reveal the true and real market condition, that is, the results of fuzzy cluster will portray the membership value of the subject or sample or member belonging to a particular segment or cluster. The membership value of the fuzzy cluster will be between 0 and 1. The value closest to 1 will be the highest membership value of the cluster function, and the value closest to 0 will be the lowest membership value of the cluster function. So, the results of fuzzy cluster membership value will not be a crisp number, rather, it will be in fuzzy in nature. The fuzzy membership value will show the real strength of belongingness of a

member in a particular segment. The solution of fuzzy cluster can be compared with the hard cluster solution to measure the consistency of the segmented market based on the market size. The effectiveness of the segment can be measured with the help of the consistency score of various segments. The information on the consistency of segmented market will facilitate the retailers to make decisions on marketing investment plans. Fifth, this research attempts to build a priori predictive model of segment which helps the retailer to classify shoppers. Sixth, Answer-tree models will help retailers to target the right shopper to their shop. Finally, this research will help retailers gain a competitive advantage over other competitors by effectively segmenting and targeting them.

1.2 NEED FOR THE STUDY

Most of the investment in a retail outlet is of fixed nature. It is more like a sunk cost, as is in the case of an airline. Of all the investments, the real estate costs are the highest, and the store needs to improve on its productivity by utilising the space optimally. This can be achieved by ensuring that there is a high level of conversion rate from visitors to buyers. This conversion is possible by impacting the behaviour, using variables such as merchandise, display, layout and other communication, within the control of the retailer. It is, thus, necessary for the retailers to know how shoppers behave and whether there are any patterns that can be used to design the retail mix for a specific set of shoppers. It can even be used for targeting, and developing the positioning strategy for the store.

Crossword is a chain of bookstores in India. It has redefined the way books are retailed in India. A large number of changes have been brought in, by taking store decisions, based on the behaviour of the Indian book buyers at the store. For instance, the racks have a lower height than in most stores. It found that the average height of the Indian consumer is lower than those in western countries, while the racks were designed based on the international standards. This not only caused discomfort but also affected the sale of the books. The lowering of the racks not only increased sales, the store looked more spacious as the shoppers could see all through to the end of the store. Similarly, it found that the shoppers would squat on the floor to browse the books on the lower shelves of the racks. It changed the design of the rack so that the lower shelf was an incline and the shopper could see the books clearly. This increased the sale of the books even when they were placed on the lowest shelves.

Consider this scenario at a grocery store. A shopper walks in and asks for bread. The shopkeeper hands over a loaf. The customer checks the softness and asks for another piece. He checks it again, comparing the softness of the two loaves and chooses one of them. In another store, loaves of bread are placed on the counter. A shopper walks in and starts choosing the loaf on the basis of softness. He picks one, pays and leaves. The two situations elicit the extent of selling as well as buying efforts required. It is very likely that in the latter case, the shopper is more satisfied and there is lesser effort from the shopkeeper as well. It is a win-win situation. The merchandising in the second store is based on the behaviour of the shopper, where as in the other it is based on the retailer‘s stocking pattern.

Shopping is an activity aimed at collecting information. The search processes give shoppers an opportunity to ensure that they take the right decision. In addition, they also derive emotional satisfaction (Tauber.E, 1972). It has been found that a high level of brand awareness need not always translate into sales. Shoppers do take into consideration the information they acquire in stores, rather than just relying on out of store communication (Underhill.P., 1999). In a study conducted by (POPAI, 2001), it was found that the sales of some product categories increased by almost 60% due to effective communication at the store. Also, every visit of the shopper is a distinct encounter and a moment of truth. Unless the interaction is satisfactory, the next visit is not guaranteed. If the store does not provide a compelling reason to visit, the amount of purchase per visit is likely to go down (Zeithaml.V., 1988).

Shopping involves a ―see-touch-feel-select‖ sequence. The degree to which the shoppers follow the whole or part of this process, varies with brand, product category, and other elements of the marketing mix (Connolly A, 1998). In an exploratory study conducted in India (Sinha P K, 2000), it was found that the extent also depended on the association that the shopper had with the store. Shoppers, who were new to the store or were considering buying a brand for the first time, showed a higher level of information search. Those who were frequent buyers of the store would either go straight to the stack and pick up the product, or would ask the retailers when trying a new brand. In some cases they would pick up the product. In many cases they would buy the product recommended by the

retailers. In any case, they would not do a very elaborate information search. Such behaviour has been explained as routine, limited or complex buying behaviour in marketing literature.

It was also found that the shoppers changed their information search process as per the stores, even when the same product and in some cases even the same brand was being bought. It was found that shoppers would ask for a brand in a kirana (general/provision) store and resist a change in many cases. But in a self-service store, they would look at more than one brand before deciding. In some cases, they would buy only from the shops they patronise, even when it was located at a distance and there were other stores in vicinity. This behaviour is very evident among cigarette buyers who have a fixed store, either because of long association or because it is in on their way to work. It is evident that, shopping, as an information search process, signifies behaviour that may differ with type of store, association with store and type of product bought. Therefore, the other segmentation variables, such as demographic or psychographic, may not help the retailers understand the shoppers in their true form. Behaviour may be a better base for profiling the shoppers and hence taking decisions that would improve the profitability of the store. It is also possible to mould shoppers‘ behaviour to suit the retailer‘s requirements, especially in the Indian context where retailing is opening up as an organised activity.

The evolving nature of Indian retailing has created excitement among shoppers. India has the highest number of retail outlets (more than 12 million) in the world,

though the per capita retail space is the lowest (CII/McKinsey, 2000). The industry is estimated to be about $180 billion. The organised (new format) stores represent a mere 2% share of this market (Business Today, 1999) It is very low compared to other countries including those in Asia. Up to 80% of all retail sales in the United States is controlled by the organised retail sector. The corresponding number in Western Europe is 70%. The scenario of organised retailing in Asia is lopsided. In Taiwan, the share of the organised sector is 81%, followed by Malaysia and Thailand, at 45% and 40%, respectively, whereas China and India stand at 15% and 2%, respectively. (Fernandes et al., 2000). As new formats and retail concepts are being introduced at a fast rate, the shoppers are not able to build an attitude towards shopping (Sinha, 2003). It prompted the author to explore the case of behaviour as a basis to profile and segment shoppers. So far, across the world, most of the segmentation efforts have been based on demographic or psychographic dimensions. Also, they have used statistical methods. This research is an attempt towards using shoppers‘ in-store behaviour as a basis for segmenting shoppers, using both qualitative and quantitative methods.

Shopping, as an information search process, signifies that behaviour may differ with type of store, association with store and type of product bought. Therefore, while other segmentation variables such as demographic or psychographic may help retailers understand shoppers‘ attitudes, they may not explain the behaviour of shoppers directly. It has also been found that even when a shopper has a list, there is a difference between the list and the product bought (Kollat and Willet, 1967). Thus, it is safe to assume that such behaviour is caused by the variables in

the store. The hypothesis can be extended to assume that it is possible to mould shopper behaviour to suit the retailers‘ requirements. Under this assumption, the shoppers‘ behaviour at the store may prove to be an alternate segmentation basis.

Several attempts have been made to develop a typology of shoppers. Table 1.1 provides a summary of these (Brown and Reid, 1997; Westbrook and Black, 1985). The typologies have been developed based on attitude towards shopping and the motives/gratification derived. Perhaps it stems from the fact that an activity like shopping can best be explained by attitudes that lead to a certain behaviour. Some of them are based on psychographics. However, they suffer from a few limitations.

These typologies do not consider the impact of the context. They are presented as stable shopper orientations and are not sensitive to situational influences (Reid and Brown, 1997; Hibbert and Tagg, 2001). It is known that context affects shoppers‘ behaviour. Tai and Fung (1997) found that environment-induced emotional states have a positive association with in-store behaviour of the shoppers, which in turn, impacts upon the pleasure felt in the store and the in-store rating of the environmental stimuli. By inducing shoppers to stay for a longer duration, a retail outlet may lead them to increase spending (Donovan et al., 1994). Falk (1997) stresses that retailers provide an opportunity for shoppers to interact at close quarters with the scopics and derive aesthetic stimulation.

The format and ambience of a store has its own impact on shoppers. Its effect is seen in dressing patterns, language, and interaction with store personnel as well as response to communication within the store. It is likely that even a shopping apathetic would behave in the same manner as a shopper who has a liking for shopping, in a given store environment. It was found in a study that even a stereotype‘s behaviour changed when the store context changed (Otnes andMcGrath, 2001). Since retailing is a service business, and is not just distribution, there are many __soft‘‘ variables that lead to creation of a customer franchise (Baker et al., 2002).

The methodology used in all but one of the aforementioned studies, is statistical. It is quantitative in nature using scales to measure attitude. In India, where retailing is still evolving, it is very likely that the expectations of shoppers would be comparatively basic and limited to tangible aspects only. Also, retailing in India is at a flux and hence attitude measurement may not yield the required results. Behaviour at the store, as factual information, may be a more reliable measurement. It is posited that a contextual stance increases understanding of the behaviour of a shopper. It is useful because the intentional stance may not always predict accurately or may do so only when situational correspondence is extremely high. It is also indicated that such an approach prompts innovation in data collection and analysis that has not been generated by previous, non-contextual, consumer research (Foxall,2000). This methodology becomes especially appropriate for store-level research, as the context of a store is different from the closed setting experiments that have been used to understand operant

behaviourism (Foxall, 1998), Behaviour, in this case, is defined as the act of information search, as well as reaction to the cues at the store, such as layout and displays, salespeople, POP Communication and other facilities such as carts, bins, etc (operationally defined as store images). It encompasses activities that may not always lead to purchase, such as browsing of a magazine or use of a listening post in a music store. There have been attempts to study consumer behaviour with regard to product usage, but none to segment shoppers. Thus, segmenting the shoppers based on their behavioural aspect at the store (store images) is not dealt with in previous studies. Moreover, an alternative method, developing the segmentation descriptors (both qualitative and quantitative) based on behaviour at store is not developed in the earlier studies.

Another critical aspect of the retailers‘ ability to maintain and develop their market position is, in the development and management of a favourable store image (Nevin and Houston 1980, Samil 1989). Store image may be defined as the overall attitude towards the store based upon the perceptions of relevant store attributes (Bearden 1997; Doyle and Fenwick 1974-75; James, Durand, and Dreves 1976; Korgaonkar, Lund, and Price 1985; Marks 1976). Image considerations are important aspects in the development of an integrated marketing strategy for individual stores, store chains, and shopping centers. Store image has been found to be related to such key concepts of retail success as store patronage (Stanley and Sewall 1976; Korgaonkar, Lund, and price 1985), store loyalty (Lessig, 1973) (Srigy and Samli, 1989), and the share of the household budget spent in the store (Hildebrandt 1988). Stores that have a favourable image can draw customers from

larger distances, and thus mitigate possible location inconveniences (Stanley and Sewall 1976). A unique store image is one of the retailers‘ most valuable marketing assets, creating a competitive advantage that is not easily duplicated by other retailers (Rosenbloom, 1983). In as early as the study of (Martineau, 1958), however, it has been pointed out that no store can be all things to all people. Different groups of consumers might place different importance on the various store image attributes. Stores may emphasize different image attributes as part of their marketing strategy, and ideally, the important attribute stressed by the store should be those to which the target segment attaches the most importance. The importance of segmenting retail markets on the basis of store image attributes and the development of an image that conforms to the needs of the retailer‘s target group of consumers have been repeatedly stressed in the literature (Doyle and Fenwick 1974-1975; Hansen and Deutscher 1977-1978; James, Durand, and Dreves 1976; Rosenbloom 1983; Samil 1989). Store Image segmentation provides guidelines for a retail firm‘s marketing strategy and can increase profitability (Frank, Massy, and Wind 1972; Wind 1978; Samil 1989). Despite its importance, however, it appears that the managerial potential of the store image segmentation research is still unfulfilled (Berkowitz, Deutscher, and Hansen 1978; King, Tigert, and Ring 1979), an important reason being that ― the research methods in segmenting the market have not always been as sound as they might be‘ (Peterson and Kerin, 1983). In this research, various approaches to store image segmentation are deeply reviewed. Several analytical approaches have been used in store image segmentation research. One stream has examined differences in store image attribute importance for a priori defined segmentation variables (e.g.. Hansen and

Deutscher, 1977-1978), (Gentry and Burns, 1977-1978), (Schiffman, Dash, and Dillon 1977), (Bearden, 1977). Others have clustered subjects on the self-stated importance of store image attributes (Howell and Rogers 1983), (Malhotra, 1986), (Tantiwong and Wilton 1985) or on statistically estimated importances (Verhallen and DeNooy, 1982). Another approach is to cluster individuals on consumer characteristics and subsequently to estimate the importance of the various image attributes in each segment (Hortman et al, 1990) Of late, a number of new judgmental methods of market segmentation have been proposed that may be fruitfully applied to the domain of the store image segmentation. (Kamakura, 1989) has developed a segmentation method within the context of conjoint designs, and (DeSarbo, Oliver and Rangaswamy, 1989) and (Wedel and Steenkamp, 1989) have proposed cluster-wise regression approach to market segmentation.

Table 1.2 proves a comparison of seven approaches of store image segmentation. It includes reference, where appropriate, to studies in retailing employing the approach. The column of Table 1.2 corresponds to a set of six features of segmentation methods that are briefly discussed below.

1. Design: Design concerns whether the market segments are defined by the management or the researcher (a priori design) or are based on the responses of the subjects (clustering- based design). A priori segmentation is especially valuable when the management has a clear idea about the relevant basis of segmentation. However, it is often difficult to establish a relevant basis for segmentation a priori

and to capture the complexities of the market in one or two variables that are typically used (Samli, 1989), (Wind, 1978), because of which the retail firm may miss important marketing opportunities.

2. Store-image-attribute importance. The two basic approaches to obtain store image attribute importance are: (1) subject rate the importance of the various attributes (self-stated importance); and (2) attribute importance is statistically estimated based on the subject‘s overall evaluation of stores (derived importances). There is considerable evidence that statistically derived importances represent people‘s actual weight more accurately than their self-stated weights (Fishbein and Ajzen, 1975), (Slovic and Lichtenstein, 1971), See (Gentry and Burns 1977-1978) for criticism on self-stated weights in store image research).

3. Types of clustering procedure: The clustering procedures can be distinguished according to the type of partitioning obtained: non-overlapping, overlapping, or fuzzy (Hruschka, 1986). In non-overlapping clustering, subjects belong to a single segment only, while in overlapping clustering, subjects may belong to multiple segments. In fuzzy clustering, the hard membership or non-membership for a subject in one (non-overlapping) or multiple (overlapping) clusters is replaced by gradual membership, indicating the nearness of the subject to the cluster. Overlapping and fuzzy clustering approaches are consistent with the notion that consumers may belong to different segments when they desire several aspects in a store, possibly in relation to different buying and consumption situations (Howell and Rogers, 1983), (Srivastava, Alpert, and Shocker, 1984). There is a risk of

oversimplification and loss of explanatory power when clusters are assumed to be mutually exclusive in store image segmentation research (Arabie et al, 1981).

4. Number of stores: Some approaches require the number of stores to be evaluated by the subject, to exceed the number of predictor variables (i.e., the number of store image attributes or the number of dummy variables in conjoint designs) and /or require that each subject evaluates an equal number of stores.

5. Optimization criterion: Some methods cluster the subjects based on similarity of variables of interest (e.g., attribute importances), while other maximize the (within-segment) ability of the store images attributes to predict overall store image. The former type of method does not necessarily lead to a cluster whose store –image attribute importances best explain the overall evaluation of the stores of each individual in the sample (Kamakura, 1988). One may obtain a good cluster solution (in terms of the homogeneity of estimated-image-attribute weights) without any appreciable increase in the predictive power over the unsegmented model, as was indeed found in the context of store image by (Howell and Rogers, 1983). It has been argued that predictive fit of the estimated store image functions should be maximized, as it is a key measure for evaluating market segmentation results and for developing a marketing strategy (Hauser and Urban, 1977), (Kamakura, 1988).

6. Statistical test: Does the procedure incorporate significance tests for the effect of the various store images attributes on the overall evaluation and if so, which type of significance test is used?

By examining the Table 1.2, it indicates that the hierarchical least-square procedure for conjoint design of (Kamakura, 1988) and the clusterwise regression procedures of (DeSarbo, Oliver, and Rangaswamy, 1989) and (Wedel and Steenkamp, 1989) have importance advantage over the traditional methods employed in store image segmentation research. All three procedures adopt the clustering-based design, estimate the store image attribute importances rather than employing self-stated weights, and maximize predictive fit. The two clusterwise regression procedures are more versatile than Kamakura‘s method in that, they incorporate significance tests, allow subjects to belong to different segments, and most importantly, can be used even in situations where the number of images attributes exceeds the number of stores. Research designs where the model is under identified at the individual level are not uncommon in retailing, given that, the number of stores in a certain category the consumer is aware of or patronize, is typically rather small (Goldman 1977-1978), (Howell and Rogers, 1983). The set of stores a consumer is aware of or patronizes may vary considerably across consumers (Finn and Louvire, 1990). Fuzzy-clusterwise regression analysis (FCR) (Wedel and Steenkamp, 1989) imposes no restrictions on the number of stores to be evaluated by each subject, while overlapping clusterwise regression analysis (OCR) (DeSarbo, Oliver, and Rangaswamy 1989) requires an equal number of stores for all subjects. When the set of stores evaluated is equal for all subjects,

FCR and OCR may be expected to produce similar results, as evidenced by an empirical comparison of the two methods, where it appears to be that FCR is applicable to a wider variety of retail problems than OCR. (Jan Benedict, Steenkamp, and Wedel, 1991) revealed that the potential of clusterwise regression procedures for store image segmentation may not be fully realized by researchers in retailing as no applications have been reported hitherto.

On the other hand, the market segmentation literature proffers many procedures to segment the market. Michel Wedel and Kamakura (1998) describe four classes of segmentation methods:

1. A priori descriptive methods (e.g. through cross-tabs. OLAP cubes).
2. A priori predictive methods (e.g. discriminant analysis).
3. Post hoc descriptive methods (e.g. clustering methods and mixture models).
4. Post hoc predictive methods (e.g. clusterwise regression, CHAID, CART, mixture regression modes).

Researchers often use clustering analysis as a tool in studies of market segmentation (Post hoc descriptive method). Usually most researchers use Hierarchical and Non- Hierarchical clustering procedures to segment the market which typically results in crisp partitioning form; in other words, where one sample cannot belong to two or more groups. According to Cattell (1978), the boundary between segments are fuzzy in the real market which indicates that different segments overlap with each other and a sample can belong to more than

one market segment, and has a tendency to be identified with a single, stronger segment. This study attempts to utilize the fuzzy cluster method to find the membership grade to describe each segment and therefore the real market situation will be presented. Thus, this research adds value to the existing method of segmenting the retail market based on behavioral patterns on store images by using fuzzy clustering approach.

In this research it is argued that integration of subtractive cluster analysis and other cluster validity techniques to define the number of clusters in fuzzy-c means that cluster analysis (to find fuzziness) holds high potential to segment retail markets based on store image. The usefulness of fuzzy-c means clustering is empirically investigated in the context of various store images of organized retail organizations. A priori predictive model of segment is also formulated by using discriminant analysis and the model is compared with fuzzy discriminant analysis. Answer-tree based classification model is developed for defuzzified segments in tune with demographic variables for the purpose of gridding the segmented market.

Table 1.2. Approaches of Store Image Segmentation

Approach	Design	Image attribute importance	Clustering	# store	Optimization criteria	Significance test for benefit importance	Example
A priori segmentation	a priori	Self-stated	n.a	No restriction	n.a	No	Hausen and Deutscher (1977-1978)
Benefit segmentation (Self-stated)	Cluster based	Self-stated	Non-Overlapping	No restriction	Similarity in attribute importance	No	TantiwonWilton (1985)
Benefit segmentation (Derived)	Clustering Based	Derived	Non-Overlapping	Exceeds number of predictor variables	Similarity in attribute importance	t-test	Verhallen and De Noooy (1982)
Segmentation on consumer characteristics	Clustering Based	Derived	Non-Overlapping	No restriction	Similarity on consumer characteristics	t-test	Hortman et al. (1990)
Hierarchical Least square	Clustering –Based	Derived	Overlapping	Equal for all subjects; exceeds number of predictor variable	Predictive fit	no	-
Overlapping cluster-wise regression (DeSarbo et al. 1989)	Clustering based	Derived	Overlapping	Equal for all subjects	Predictive fit	t-test	-
Fuzzy clusterwise regression (Wedel and Steenkamp, 1989)	Clustering – Based	Derived	Fuzzy	No restriction	Predictive fit	Monte- Carlo	-

2. REVIEW OF LITERATURE

The purpose of this research is to create an enhanced method of segmenting the Indian shopping tourism retail market based on the behaviour pattern of shoppers towards various store images by integrating various tools and techniques of market segmentation methods. More specifically, firstly, the purpose of this study is to determine the influence of store image dimensions on store patronage and find the importance of various store images relative to shopping and behavioural orientations. Secondly, the next purpose is to determine the exact market boundaries to define various segments. In other words, the purpose is to find the optimum number of clusters or segments to be specified for initial iteration, for defining the market segment. Thirdly, the following action is to segment the shoppers based on their behavioural pattern towards store images by using hard and soft clustering methods. Fourthly, the purpose is to measure the consistency of behaviour pattern of shoppers towards store images; the results of hard cluster solutions and soft cluster solutions have to compared and analysed. Besides, a stability index is to be created for comparing the stability among the various segments. Fifthly, the next purpose is to create an enhanced predictive model for segmenting the shoppers based on their behavioural pattern towards store images by using the fuzzy method. Sixthly, the purpose is to build an answer tree model for gridding the various segmented market. Finally, the purpose is to derive various cues for decision making from the results of various approaches of segmentation. In order to accomplish the basic purpose of this research, the literature search is carried out on various aspects. The literature review includes following sections.

1) Market segmentation, 2) Retail in India, 3) Retail formats in India, 4) Choice of a Retail Store and Retail Store types, 5) Cluster Analysis and Cluster Validity Function, 6) Stability of segmentation .

2.1 Market Segmentation: Literature review

The nuanced scrutiny of the literature review concentrates on the following issues related to marketing segmentation.

Market segmentation – Definition and Concepts

The role of market segmentation in marketing strategy

What criteria need to be satisfied for successful market segmentation

The factors that determine success, satisfying these criteria

An overview of cluster analysis algorithms. Although better methods are available, traditional cluster methods are still widely applied and it is useful to understand which of the many choices you can make when you perform cluster analysis, are likely to be the best.

Practical issues you may run into when doing market segmentation analysis

Latent class segmentation methods. This is currently probably the most advanced and powerful method available for market purposes. The general statistical framework is described, as well as, some possible applications.

It provides guidelines and recommendations to researchers when performing market segmentation analyses

Marketing

Marketing is the process for: defining markets; qualifying the needs of the customer groups (segments) within these markets; determining the value propositions to meet these needs; communicating these value propositions to all those people in the organization responsible for delivering them and getting their buy-in to their role; playing an appropriate part in delivering them and getting their buy-in to their role; playing an appropriate part in delivering these value propositions (usually only communication); monitoring the value actually delivered. For this process to be effective, organizations need to be consumer/customer driven.

Market Segmentation

Market segmentation is the process of splitting customers, or potential customers, in a market into different groups or segments, within which customers share a similar level of interest in the same, or comparable, set of needs satisfied by a distinct marketing proposition. Market segmentation is a concept in economics and marketing. A **market segment** is a sub-set of a market made up of people or organizations with one or more characteristics that cause them to demand similar product and/or services based on the qualities of those products such as price or function. A true market segment meets all of the following criteria: it is distinct from other segments (different segments have different needs), it is homogeneous

within the segment (exhibits common needs); it responds similarly to a market stimulus, and it can be reached by a market intervention. The term is also used when consumers with identical product and/or service needs are divided up into groups so that they can be charged different amounts. The people in a given segment are supposed to be similar in terms of the criteria by which they are segmented and different from other segments in terms of these criteria. These can broadly be viewed as 'positive' and 'negative' applications of the same idea, splitting up the market into smaller groups. Examples: Gender, Price, Interests. While there may be theoretically 'ideal' market segments, in reality every organization engaged in a market will develop different ways of imagining market segments, and create product differentiation strategies to exploit these segments. The market segmentation and corresponding product differentiation strategy can give a firm a temporary commercial advantage. Market segmenting is dividing the market into groups of individual markets with similar wants or needs that a company divides into distinct groups which have distinct needs, wants, behavior or which might want different products & services. Broadly, markets can be divided according to a number of general criteria, such as by industry or public versus private. Although industrial market segmentation is quite different from consumer market segmentation, both have similar objectives. All of these methods of segmentation are merely proxies for true segments, which don't always fit into convenient demographic boundaries. Consumer-based market segmentation can be performed on a *product specific* basis, to provide a close match between specific products and individuals. However, a number of generic market segment systems also exist, e.g. the system provides a broad segmentation of the population of the

United States based on the statistical analysis of household and geo-demographic data. The process of segmentation is distinct from positioning (designing an appropriate marketing mix for each segment). The overall intent is to identify groups of similar customers and potential customers; to prioritize the groups to address; to understand their behavior; and to respond with appropriate marketing strategies that satisfy the different preferences of each chosen segment. Revenues are thus improved. Improved segmentation can lead to significantly improved marketing effectiveness. Distinct segments can have different industry structures and thus have higher or lower attractiveness.

When it comes to marketing strategies, most people spontaneously think about the 4Ps (Product, Price, Place, Promotion) – maybe extended by three more Ps for marketing services (People, Processes, Physical Evidence). Market segmentation and the identification of target markets, however, are an important element of each marketing strategy. They are the basis for determining any particular marketing mix. Literature suggests the following steps:

Step 1: Market Segmentation

a. Identification of customers‘ needs and market segments

b. Developing profile of resulting market segment

Step 2: Identification of Target Markets

a. Evaluation of attractiveness of each segment

b. Selection of target segments

Step 3: Positioning

a. Identification of differential advantages in each segment

b. Development and selection of positioning

concepts Step 4: Marketing Planning

The importance of market segmentation results from the fact that, the buyers of a product or a service are no homogenous group. Actually, every buyer has individual needs, preferences, resources and behaviors. Since it is virtually impossible to cater for every customer‘s individual characteristics, marketers group customers to market segments by variables they have in common. These common characteristics allow developing a standardized marketing mix for all customers in this segment. There are a huge number of variables that could be used for market segmentation in theory. They comprise easy-to-determine demographic factors, as well as variables on user behavior or customer preferences. In addition, there are differences between private customers and businesses. Since customer orientation of organizations is growing, segmentation as the basis for establishing customer relationships and customer loyalty gains importance. In this context, the elements of the loyalty ladder model could be used as segmentation variables.

Figure 2.1 The Loyalty Ladder Model

Marketers have to choose those variables that are relevant for segmenting the market for a particular product. The basic rule is to focus on a limited number of important variables. To segment the market into too many small, slightly distinct segments would require splitting up the marketing budget into too many ineffective chunks. Such varied marketing activities in the diverse segments could confuse customers and would lead to cannibalization effects.

Kotler mentions five criteria for an effective segmentation:

> **Measurable**: It has to be possible to determine the values of the variables used for segmentation with justifiable efforts. This is important especially for demographic and geographic variables. For an organization with direct sales (without intermediaries), their own customer database could deliver valuable information on buying behavior (frequency, volume, product groups, mode of payment etc).

Relevant: The size and profit potential of a market segment have to be large enough to economically justify separate marketing activities for this segment.

Accessible: The segment has to be accessible and servable for the organization. That means, for instance, there are target-group specific advertising media, as magazines or websites the target audience likes to use.

Distinguishable: The market segments have to be so diverse that they show different reactions to different marketing mixes.

Feasible: It has to be possible to approach each segment with a particular marketing program and to draw advantages from that.

2.1.1 Reasons for Market Segmentation

As already stated, segmentation is the basis for developing targeted and effective marketing plans. Furthermore, analysis of market segments enables decisions about the intensity of marketing activities in particular segments.

A segment-orientated marketing approach generally offers a range of advantages for both businesses and customers.

Better serving customers needs and wants

It is possible to satisfy a variety of customer needs with a limited product range by using different forms, bundles, incentives and promotional activities. The computer manufacturer Dell, for instance, does not organize its website by product groups (desktops, notebooks, servers, printers etc), but by customer groups (privates, small businesses, large businesses, public/state organizations). They

offer the same products to all customer groups. Nevertheless, they suggest product bundles and supporting services that are individually tailored for the needs of each particular group. As an example, Dell offers to take on all IT-administration for companies. This service provides a huge potential for savings for corporate customers. However, it would be absolutely useless for private customers. Thus, segment-specific product bundles increase chances for cross selling.

Higher Profits

It is often difficult to increase prices for the whole market. Nevertheless, it is possible to develop premium segments in which customers accept a higher price level. Such segments could be distinguished from the mass market by features like additional services, exclusive points of sale, product variations and the like. A typical segment-based price variation is by region. The generally higher price level in big cities is evidence for this.

When differentiating prices by segments, organizations have to take care that there is no chance for cannibalization between high-priced products with high margins, and budget offers in different segments. This risk is higher, the less distinguished the segments are.

Opportunities for Growth

Targeted marketing plans for particular segments allow to individually approach customer groups that otherwise would look out for specialized niche players. By segmenting markets, organizations can create their own ‗niche products' and thus attract additional customer groups.

Moreover, a segmentation strategy that is based on customer loyalty (see loyalty ladder model) offers the chance to attract new customers with starter products and to move these customers on to premium products.

Sustainable customer relationships in all phases of customer life cycle Customers change their preferences and patterns of behavior over time.

Organizations that serve different segments along a customer‘s life cycle can guide their customers from stage to stage by always offering them a special solution for their particular needs.

For example, many car manufacturers offer a product range that caters for the needs of all phases of a customer life cycle: first car for early twenty, fun-car for young professionals, family car for young families, etc. Skin care cosmetics brands often offer special series for babies, teens, normal skin, and elder skin.

Targeted communication

It is necessary to communicate in a segment-specific way even if product features and brand identity are identical in all market segments. Such a targeted communications allows stressing those criteria that are most relevant for each particular segment (e.g. price vs. reliability vs. prestige).

Stimulating Innovation

An undifferentiated marketing strategy that targets at all customers in the total market necessarily reduces customers‘ preferences to the smallest common basis.

Segmentations provide information about smaller units in the total market that share particular needs. Only the identification of these needs enables a planned development of new or improved products that better meet the wishes of these

customer groups. If a product meets and exceeds a customer‘s expectations by adding superior value, the customers normally is willing to pay a higher price for that product. Thus, profit margins and profitability of the innovating organizations increase.

Higher Market Shares

In contrast to an undifferentiated marketing strategy, segmentation supports the development of niche strategies. Thus marketing activities can be targeted at highly attractive market segments in the beginning. Market leadership in selected segments improves the competitive position of the whole organization in its relationship with suppliers, channel partners and customers. It strengthens the brand and ensures profitability. On that basis, organizations have better chances to increase their market shares in the overall market. Summarizing all these advantages, the need for market segmentation is closely related to strategic decisions. Market segmentation is the basis for customer orientation and differentiation.

It is well known that suppliers in mass markets mostly compete on price. Demand for those products that are clearly differentiated from competition and that offer a particular value to customers do has lower price elasticity; hence, only those products can sustain a higher price level and higher margins. The precondition for providing such value addition is detailed knowledge about customers‘ preferences. These preferences will probably be diverse in the total market, but fairly

homogenous within distinguishable segments. Focus on attractive market segments is of special relevance in our fast moving times of Internet economy.

If you‘re trying to segment your market in the traditional way, what you may be looking for would be groups of consumers sorted out in such a way that a certain likeness exists within each group, and a difference exists between them. The variable determining the meaningful likeness or difference between those groups would be the segmentation variable. A trivial segmentation variable, just for the sake of demonstration, would be hair colour. However, after having segmented the customers into groups, it is reasonable to assume that you would expect to do something with it. Let‘s say that you have decided to target a certain segment. You would probably want to do some marketing activities that will appeal to this segment, or else, to communicate some kind of enticing message to it. Sometimes, the segmentation variable could suffice for the purpose, (—listen to me, all you red-heads out there…‖). In most cases (for instance, the segment of those who consume beer only outside their home), you would have to characterize your segment before you could address them. In other words, you would have to define what describes the customers in that segment, beyond your segmentation variable, and also, what makes them different from consumers in other market segments. The characterization of your segment is a task that is not the same as defining your segment. It is a distinct next step. But now, if you can be truly sincere with yourself, it can be that you have already found out that it does not work.

What is meant by that? Well, according to the conventional segmenting method, people in each segment should differ – to some extent – from the people in the other segments. Nevertheless, usually, they do not. Let us say that in a certain segment, there is a majority of women: 60% versus about 50% in the entire population. Even if the difference is statistically significant, is this a ‗feminine segment' then? And what about the other 40% who are men? Let us assume, for example, that in one of the segments, there are a relatively high percentage of religious people. Is this a ‗religious' segment? Absolutely not. The vast majority within this segment is clearly not religious, even though their proportion is higher than it is in the general population. Thus, in each and every segment, with slight variations, you could meet almost every layer of society.

So where did this idea, on which the conventional segmenting method is based, come from? In the distant past, and in traditional societies (sectarian) the people's behavioural patterns were pretty much modelled by their affiliation to a certain gender, a nationality/tribe/race, a certain religion, a social/economic status, a profession, and an age group – much more than today, anyway. There were clear clusters of elements pertaining to appearance, general behaviour and particularly consumption. Then, back in those days, if you knew one element of a particular cluster, you could quite easily guess the others. But all this has changed. As people are becoming gradually more individualistic, and as possibilities have multiplied, people have become less and less definable as types.

Back in the 70‘s it was taught that because of this process, it is worthwhile to conduct psychographic segmentations which include characterizations according to lifestyle groups. A widely accepted approach such as VALS (Values, Attitudes, and Lifestyles) classified people into 10 _lifestyle groups‘, later reduced to eight.

Disappointingly, it was subsequently discovered that the groupings couldn‘t predict, with reasonable credibility or consistency, specific purchase decisions, or preferences in specific product categories. The alternative was to use those groups for purposes of personal-psychological characterization. So, then again, just like in the case of the demographic and socioeconomic classifications, we find in each and every segment of each segmentation project, ?representatives‘ of all lifestyle types, and, the this has the same problem.

This problem has just been worsening with time. The reason why this is so, is that there has been a long lasting consumer trend going on, with numerous implications, because of which the way the market is segmented should be changed.

Firstly, the consumer refuses almost completely to abide by segments that create homogeneous groups (heterogeneous from others) according to demographic, socio-economic variables, or even according to lifestyle. Our customer will not behave and consume under our stereotypical forecasts. He is a ―collector‖, an _eclectic consumer‘. He likes the old (Frank Sinatra), as well as the new (fast internet), the comfortable (frozen foods) as well as the effortful (cooking, DIY), the expensive (BMW) as well as the economical (hardware do-it-yourself stores),

the international (Giorgio Armani) as well as the locally rooted (folk dancing), the sophisticated (Nokia) as well as the simple (family), the epicure (a double Makiato) as well as the crude (football).

Yet, what is it that constitutes the long lasting cultural trend that has been modeling the eclectic consumer? The eclectic consumer is motivated by the Fear of Missing Out. At present, over 70% of the population is motivated by the Fear of Missing Out. About 30% could be considered highly motivated by FoMO. That means a great big number of customers. The eclectic consumer has become frenzied by the abundance of opportunities, and is now addicted to the concept. She does not want to miss anything and so, her life is multiplex, yet laden. She is always accessible by cellphone or email, updated, open to new concepts, and not so afraid of changes. She spends fewer years than she used to spend in the same apartment, same job, and same marriage (that is, if she does not belong to the group of the eternally single). In contrast to the past, she is proud of being flexible and developing, and is not committed to a fixed personality. And she is much less loyal to brands than she used to be. In fact, she embraces new brands so impetuously, that the marketers are convinced as to their own ingenuity.

So, how do you market to the eclectic customer who is afraid of missing out? Well, there are many facets to this art; one of them to focus on, is market segmentation, at an era in which customers refuse to be classified in convenient clusters, considering that this phenomenon is prevalent in many product categories. One central insight for the re-designing of market segmentation is as

follows: the eclectic consumer, who will not miss anything, ‗connects' to different, even contradictory, motivations he has at different times. Because these motivations are not necessarily compatible with one another, the eclectic consumer is constantly in motion from one stereotype to another, from one lifestyle to another. That is why he chooses to drink pro-biotic yogurt when he gets up in the morning, and to grind up a steak & French fries lunch at noon, topping it off with a cigarette. After work, he meets up with friends for a delicate Japanese dinner, which he concludes with a bottle of fine French cognac.

In order to adapt to this consumer reality, the segmentation (and subsequently, the products and services, our advertising and so on) should be formulated not according to groups of people, but according to motivations and uses. ‗Uses' here means, among other things, psychological uses, such as mood control, self-esteem enhancement, and fantasy support, and also social uses, such as signalling others things like group affiliation, specific atmospheres, or impression control.

Note that this constitutes a formation of a real revolution in segmentation thinking. Some might think that this is not so much about segmentation as it is about consumer behavior analysis. Let us recall the original purpose of market segmentation. It is the fraction of the market into smaller units enabling us to focus our marketing/branding/advertising activities, and to achieve differentiation, so that we could win advantages we could not get when working with the entire market. The search for small consumer groups has evidently stopped delivering results. However, the pursuit after groups of ‗purchases/consumptions' rather than of people, could offer new horizons.

According to the old segmentation, each group is characterized by a need/preference/motivation. The new approach preserves this concept. Yet, in the new reality, and according to the new approach, the motivation is no longer common within a defined consumer group. A ‗segment' is now a group of ‗purchases/consumptions' qualified by a certain context of purchasing or consuming a product plus a specific motivation. When we segment according to this approach, we analyze consumer behaviour; we identify the various contexts of product consumption, and the different motivations that characterize consumers who experience those contexts. The new relevant segments could consist of certain moods (such as the ‗I'm going to teach that husband of mine a lesson he's never going to forget' segment), certain social situations (such as the ‗Wow, I haven't seen YOU in a long time' segment), all according to what is relevant to that specific product category. Note that when we meet a given purchasing context (a dinner at a restaurant) there are varied consumer motivations that exist (‗tonight we're going out solo, no kids', versus ‗we're celebrating grandpa's birthday'), and they would be considered different market segments. A specific consumer is likely to participate in one segment, few segments, or no segment. Nevertheless, much like the old segmentation, every segment accounts for a share of our sales, and we can do our profitability calculations accordingly.

For a shampoo manufacturer, the segment of ―I'm going to look fabulous in that party‖ would be responsible for a certain percentage of the income, and so will the segment of ―I do so need a half an hour to myself‖, and the segment of ―my god,

all those expenses are wearing me out, I must cut down a little‖. The same consumer could belong to each of those market segments, at different times.

According to this approach, then, the marketing activities, at all levels, should be aimed towards a context of purchasing/consumption plus a certain motivation, and not towards groups of consumers.

Finally, let us examine one more example, a segment of the big and significant market of fashion accessories, especially the low priced ones. That‘s the segment of ―Hey, that‘s new, where did you get it from?‖ The central benefit motivating this segment is the psychological/social instrumentality: the opportunity to win a little bit of renewed attention from their surroundings. Marketers specializing in this segment tend to launch short-term brands (although many do not go beyond short-term products). Why short? Because novelty is very quick to fade and this need is one of the regenerating needs (that require ever-new sources of satisfaction) aptly catered to by Short-Term Brands. According to the new approach described here, you could launch such a product or a brand; however, it will not be designed to appeal to a particular group of customers. Entirely different customers, with relation to their genders, age groups, socio-economic layers, family orientations, and achievement seeking levels, will purchase that product whenever they experience a longing for getting a little bit of extra attention, again.

2.1.2 A customer's view of segmentation — meet my needs!

Customers segment themselves and take no notice of how companies segment their market(s). When choosing between competing products and services, customers select the proposition that meets their needs better than any other. To win market share, therefore, a company must ensure that their offers meet these needs better than any other, at a price they perceive as providing superior value for money (which does not necessarily mean it has to be the cheapest). As this is how customers operate in a market, then a segmentation project should have these as its segmentation criteria.

On its own this approach to segmentation, while able to provide you with an invaluable insight into how to win a customer's business, still requires you to know how to reach them. The input to this part of a winning proposition, provided by a detailed understanding of who the customers are and where they are to be found, is clearly very important and plays a crucial part in our segmentation process. In addition, by really understanding what underpins a customer's choice, we gain an insight into their motivations, which will lead you to understand what promotional stance to take.

Interestingly, all the reputable marketing books and marketing courses which look at the alternative approaches to segmenting markets include 'needs-based' segmentation (sometimes called 'benefit' segmentation) in their reviews. They also conclude that 'needs-based' segmentation is by far the most successful approach. This is the approach taken by The Market Segmentation Company, for which we

have developed a series of practical steps, tried and tested in numerous markets around the world, and incorporated into our segmentation process.

2.1.3 Key remaining elements of the segmentation process

With the scope of the project clearly defined (project parameters — Step 1 in the process) and the individuals identified for whom it is essential you understand their buying criteria (market mapping — Step 2 in the process), uncovering the segments can now begin.

The key remaining elements of the process are outlined below:

1. Establish a sample of customers which will represent the different decision-makers found in the specified market (referred to as 'micro-segments'), with the difference between them being the key features they use to discriminate between competing offers and the importance of these features (Step 3 in the process).
2. Record personal details about the decision-makers (including their company details if appropriate) which can be used to identify them (also part of Step 3 in the process).
3. Understand the real needs of customers and list the benefits they are seeking, along with the importance of these benefits, for each micro-segment (Step 4 in the process).
4. Bring together those micro-segments that illustrate similar patterns of importance for the benefits in order to form clusters (Step 5 in the process).

5. Verify that the concluding clusters can be regarded as segments (also part of Step 5 in the process). With the segments uncovered, it is time to identify which of them you should be targeting
6. Establish the attractiveness of each segment to your company based on how well each of them meets your requirements (Step 6 in the process).
7. Determine the relative competitive strength of your company for each segment based on how well you, compared with your competitors, meet their requirements (Step 7 in the process).

By combining segment attractiveness and relative company competitiveness you can construct a strategic picture of the specified market. This will help you select the segments that will enable your company to achieve its corporate objectives.

Additional information about each of these key elements appears below. Full details can be found in Chapters 6 through to 10 of *Market Segmentation: How to do it and how to profit from it* (2010 edition published by Good fellow Publishers Ltd, ISBN 978-1-906884-18-5). Techniques that can be applied to each step, worked examples, exercises and worksheets can be found in the relevant chapter.

1. Establish a sample of customers which will represent the different decision-makers found in the specified market (referred to as 'micro-segments'), with the difference between them being the key features they use to discriminate between competing offers and the importance of these features (Step 3 in the process).

This can be achieved as follows: Divide the specified market into identifiable groups of customers and, taking each group in turn, develop it into a micro-segment by carefully listing what the customers in the group regard as their key features for discriminating between competing offers (referred to as ‗Key Discriminating Features' — KDFs). When meaningful differences are known to occur within a group, capture these differences as separate micro-segments.

NB. It is not appropriate to include ‗price' at this stage of the process. Price is better covered in the next step of the process.

Identifying key discriminating features from the customer's point of view will provide the link to understanding the needs that customers are trying to satisfy (looked at in Step 4 and used as the basis for forming segments in Step 5). This is because customers seek out specific features not for their own sake but for the particular benefits that they deliver.

It is all based on the well established principle that customers don't buy features, they buy the benefits delivered by the features. For this sequence to be successful, however, you need to think of features as consisting of both the tangible and intangible components of an offer. Now indicate the relative importance of the key discriminating features to each micro-segment and when further meaningful differences are known to occur within a micro-segment, develop additional micro-segments to accommodate them. An alternative to the above can be to list the customers found in your market (only suitable for markets with a small number of total customers).

A further alternative is to obtain a sample through a market research exercise using a carefully designed sample frame (though this is best deferred until an internal segmentation project has developed a view of how the market splits into segments). In many in-company workshops, participants have been surprised at how much they know about their markets when this stage has been conducted rigorously.

As best as possible, attribute a size to each micro-segment (volume, value or percentage) which reflects how much of the specified market each micro-segment represents. Please note that it is better to be approximately right than precisely wrong.

2. Record personal details about the decision-makers (including their company details if appropriate) which can be used to identify them (also part of Step 3 in the process). For each completed micro-segment, add some details about who it represents using applicable profiling characteristics. As these may not apply to every customer in the micro-segment, indicate the proportion each characteristic represents. Knowing how to identify and reach the members of each concluding segment will be a crucial element to the success of a segmented approach to marketing. The

   standard categories for profiling customers are:

   - demographic/firmographic;
   - geographic;
   - geo-demographic;

- psychographic.

(_Firmographics‘ is often used to describe 'demographics' in business-to-business markets.) Also note the outlet and channel preferences of each micro-segment. This may be how you reach specific segments.

3. Understand the real needs of customers and list the benefits they are seeking, along with the importance of these benefits, for each micro-segment (Step 4 in the process). Help for this step can be obtained by talking with a cross-section of people within the organization who have customer contact and by referring to sales (and lost sales) reports and appropriate past market surveys. Do not forget, however, that the project is looking at the specified market as a whole, not just your customers. Benefits are identified by taking each micro-segment in turn and determining the needs that are being satisfied by its key discriminating features (KDFs) both individually and as a _package‘ (refer to point 1 for key discriminating features). These are the buying criteria that customers regard as being decisive when choosing between alternative offers (and are referred to as _Decisive Buying Criteria‘ — DBCs). Price is included as a decisive buying criteria for every micro-segment. WARNING: 'Price' on its own is rarely the real reason for a particular purchase: what it may be is the final decider between two competing offers that match each other in all other respects (with the lower priced offer providing better value for money). Now indicate the relative importance of the decisive buying criteria to each micro-segment and when meaningful differences are known

to occur within a micro-segment, develop additional micro-segments to accommodate them.

4. Bring together those micro-segments that illustrate similar patterns of importance for the benefits in order to form clusters (Step 5 in the process). Depending on the number of micro-segments you are working with, this may be possible to execute manually. The simplest approach is to represent the importance levels for each segment‘s decisive buying criteria (DBCs) in a way which enables you to look for matching patterns across the micro-segments, such as by using stars ' ★‘ (refer to point 3 for decisive buying criteria). It is also possible to form clusters mathematically which requires importance levels to be indicated numerically. Once the clusters have been formed the information associated with each cluster‘s micro-segments (size, decisive buying criteria importance levels and profiling characteristics) should be consolidated.

5. Verify that the concluding clusters can be regarded as segments (also part of Step 5 in the process). The three most important questions are:

   (a) Is each cluster large enough to justify the development and marketing of a specific offer? , (b) Is the offer required by each cluster sufficiently different from that required by the other clusters? , (c) Can you identify which customers are to be found in each cluster so that you can target them with their appropriate offer? If the answer to all three questions is 'yes', then the clusters can be regarded as segments. _Market segment‘

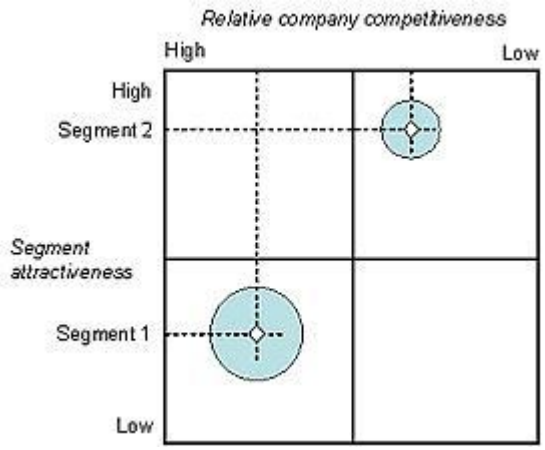
definition, A market segment consists of a group of customers within a market who share a similar level of interest in the same, or comparable, set of needs. It is also important to check your own company‘s ability to focus on the new segments, structurally and in its information and decision-making systems.

6. Establish the attractiveness of each segment to your company based on how well each of them meets your requirements (Step 6 in the process). Here you list the factors that are important to your company when having to decide where it should focus its resources, along with their relative importance to each other. Each segment is then assessed against these factors in terms of how well it can met your requirements and by taking the relative importance of these factors into account an attractiveness score is determined. The results are then transposed onto the vertical axis of a portfolio matrix as this is a useful tool for constructing a strategic picture of your market (an example appears in point 7).

7. Determine the relative competitive strength of your company for each segment based on how well you, compared with your competitors, meet their requirements (Step 7 in the process). The ability of your company to deliver against the buying criteria of each segment is assessed from the segment ‘s perspective and by taking the relative importance of these criteria into account a competitiveness score is determined. This is also determined for each of your main competitors. A relative competitiveness

score for your company is then calculated for each segment by comparing your competitiveness score with the highest score of the competitors. The results are then transposed onto the horizontal axis of the portfolio matrix containing the segment attractiveness scores and the points of intersection for each segment are identified on the matrix. A circle can then be drawn at each intersection to represent the size of the segment it relates to. An example containing two segments appears below.

Figure 2.2 Market Segmentation Attractiveness Matrix



Note: Although ‗high' is positioned on the left of the horizontal axis it can, if you prefer, be placed on the right. It is now possible to redraw the portfolio matrix based on different assumptions about your performance and that of your competitors.

2.1.4 Alternative approaches to market segmentation

Some of the alternative approaches to segmentation found in many marketing textbooks and followed by numerous companies around the world are listed below. A brief comment appears for each of them.

1. Segmentation by products and services
2. Segmentation by demographics/firmographics
3. Segmentation by geography
4. Segmentation by channel
5. Segmentation by psychographics
6. Segmentation by customer needs

Segmentation by products and services

There are three issues to consider here:

1. By solely looking at the products or services bought you may be ignoring other aspects of the purchase which are important to the customer. This often includes, for example, the convenience of sourcing the item (often associated with the preferred distribution channel), or even the reputation of the channel they are buying it from (I buy it because 'Retailer XYZ' stocks it and I trust their choice).
2. By looking at the product or service as a whole there is a danger that you may be overlooking what in particular is attracting the customer to one offer as opposed to another. Understanding the particular features customers use to discriminate between competing offers will help you do this.
3. However, in choosing between competing products or services and their associated features, customers make their decision based on which offer best satisfies their needs. Segmentation cannot, therefore, be left at the product/service level, it has to be taken one step further. This additional step is to understand what it is that the customer is really trying to

achieve from the particular features they choose. The expression, 'don't sell the features; sell the benefits' captures this essential extra step.

The products and services produced (and the distribution channels used) are a company's attempt to ensure they win profitable customer business, both now and in the future. Product/service preferences provide a crucial input into segmentation but in a supportive role rather than a determining role. These preferences are best used to help identify the real buying criteria of customers (see customer needs).

Segmentation by demographics/firmographics

('Firmographics' is sometimes used when referring to 'demographics' in business-to-business markets.) The demographic/firmographic approach assumes that customers differ according to some criteria about either themselves or about the company they work in. So, every 30 – 35 year-old can be targeted with the same offer or everyone who works in a particular industry or in a company of a particular size has exactly the same buying criteria.

Demographic information on its own does not define a marketing proposition, it does not define the product or service required, or the promotional stance to take. Demographics play a role in segmentation, but that role is not to 'define' segments. The role it plays is to help you identify for each segment a profile of the typical customer to be found in each segment. In other words, who is found in each segment. This in turn will help you understand how to reach each segment.

Segmentation by geography

The geographic approach assumes that customers found within a particular geographic area can be targeted with the same offer.

Once again, this approach on its own does not define a marketing proposition; it does not define the product or service required, or the promotional stance to take. It can, however, play a role in segmentation by providing further help in identifying how to reach the customers found in particular segments.

Segmentation by channel

As a distribution channel

There are three issues to consider here:

1. By solely looking at the different channels and specific companies used there is a danger that you may be overlooking what in particular is attracting the customer to one as opposed to another. Understanding the particular features customers use to discriminate between competing channels will help you do this.
2. However, in choosing between competing channels and their associated features, customers make their decision based on which channel best satisfies their needs. Segmentation cannot, therefore, be left at the feature level, it has to be taken one step further. This additional step is to understand what it is that the customer is really trying to achieve from the particular features they choose. The expression, 'don't sell the features; sell the benefits' captures this essential extra step.

3. It also should not be forgotten that the particular channel customers choose to use and the specific company they elect to buy from is only one aspect of a purchase. For whatever product or service line(s) you sell, ensuring there is a match between the specific items you stock and the type of customers your company attracts will be crucial to the success of your business.

The type of distribution channel you elect to be is your attempt to ensure the company wins profitable customer business, both now and in the future. The specific features on which customers focus when selecting between alternative channels provides a crucial input into segmentation but in a supportive role rather than a determining role. These preferences are best used to help identify the real selection criteria of customers. Channels can, however, play an additional role in segmentation when specific segments can be associated with particular routes to market as the channels they use will provide further help in identifying how to reach them.

Through distribution channels

The particular channel customers choose to use, and the specific company they elect to buy from, is only one aspect of a purchase. Even customers for whom channel is the deciding factor will still, in most instances, have to make a choice between the alternative offers available through their preferred channel.

A segmentation project solely focused on distribution channels would therefore overlook the other key buying criteria customers used when deciding between one

offer and another. Strategies developed solely around channels would therefore be unsuccessful.

This does not mean that the issue of channel can be ignored when it plays only a minor role in a segment's buying criteria: there is a limit to how much inconvenience customers will subject themselves to in order to obtain their preferred product or service. Understanding the criteria used by the different channels and the companies within them when they select which specific products or services they will offer is clearly crucial. A separate segmentation project designed to understand their selection criteria is therefore essential.

When distribution channels play a key role in the purchase behaviour of customers and therefore have to be included in your segmentation project, it is important to bear in mind the following two points:

1. By solely looking at the different channels and specific companies used there is a danger that you may be overlooking what in particular is attracting the customer to one as opposed to another. Understanding the particular features customers use to discriminate between competing channels will help you do this.
2. However, in choosing between competing channels and their associated features, customers make their decision based on which channel best satisfies their needs. Segmentation cannot, therefore, be left at the feature level, it has to be taken one step further. This additional step is to understand what it is that the

customer is really trying to achieve from the particular features they choose. The expression, 'don't sell the features; sell the benefits' captures this essential extra step.

The distribution channels used (and the products/services produced) are a company's attempt to ensure they win profitable customer business, both now and in the future. Channel and product/service preferences provide a crucial input into segmentation but in a supportive role rather than a determining role. These preferences are best used to help identify the real buying criteria of customers. Channels can, however, play an additional role in segmentation when specific segments can be associated with particular routes to market as the channels they use will provide further help in identifying how to reach them.

Segmentation by psychographics

Although this approach on its own does not define the product or service required, by identifying the internal drivers of decision-makers it can help define the most appropriate promotional stance to take for different segments. So, like the other alternative approaches to segmentation, it too can contribute to a segmentation project.

2.1.5 Role of Market segmentation

To build competitive advantage both in the short term and in the longer term, companies need to segment their target group. Market segmentation can aid in better achieving strategic and tactical objectives. For strategic purposes, market

segmentation results can be used to prioritize market segment opportunities. Kotler (1989) has called this the four P‘s of strategic marketing (which are additional to the four P‘s of tactical marketing), viz: (1) Probing, (2) Partitioning, (3) Prioritizing, and (4) Positioning. For tactical purposes (the traditional four P‘s, product, price, place, and promotion) market segmentation results can be used to further optimize specific product/service offerings, and can help implementing communication and advertising plans.

only means that segments must react differently to marketing efforts, but also that the required marketing efforts are consistent with the strengths and core competencies of the company. A specific market segmentation solution obtained will be a function of the variables used to segment the market and the methods or procedures used to arrive at a certain classification. The literature on cluster analysis and market segmentation procedures is vast. More recently, the field has seen important developments in the area of finite mixture (latent class) procedures. In this note we would like to summarize developments in the area of cluster analysis and latent class models. References to the best articles on each methodology will be given.

2.1.7 Segmentation Variables

The nature and quality of the identified segments depends in first instance on the variables used. Four classes of segmentation variables can be distinguished, each class having its own specific strenghts:

General observable variables. Examples of segmentation variables that fall into this class include demographic variables, socio-economic variables, etc.

Product-specific observable variables. Examples of this class are usage frequency, and brand loyalty,

General unobservable variables. In this class we encounter variables such as life-style variables, psycho-graphics, etc.

Product-specific unobservable variables. Examples here include benefits, utilities, preferences, intentions, etc.

This classification is well-known. We note that each of these classes of variables may have its advantages. For example, general observable variables are easy to collect, reliable and stable. Although general unobservable bases are weakly related to purchase behaviour, they are in general accessible and are useful for the development of advertising copy. Derived benefits and product importances have been found to score high on identifiability, responsiveness, and actionability. See for a more detailed description Wedel and Kamakura (1998). The use of different types of segmentation variables, making use of the various strengths of these variables is recommended.

2.1.8 Segmentation Methods

After having identified a set of potential segmentation variables, the methods used to identify the segments will be an important aspect in obtaining a successful segmentation. Wedel and Kamakura (1998) describe four classes of segmentation methods:

1. A priori descriptive methods (e.g. cross-tabs),

2. A priori predictive methods (e.g. discriminant analysis)

3. Post hoc descriptive methods(e.g. clustering methods, mixture models)

4. Post hoc predictive methods (e.g. clusterwise regression, CHAID, CART, mixture regression models).

The most important distinction here is between descriptive and predictive methods. Descriptive methods are defined in general as those methods that analyze a set of variables, without a distinction between dependent and independent

variables. Predictive methods are defined in general as methods that analyse a set of variables, whereby one variable is designated as the dependent variable and the other variables are designated as the independent variables.

General Overview of (Descriptive) Cluster Analysis Procedures1

The identification of clusters or market segments in a sample of respondents, described on a number of relevant variables, can be done in many different ways. In general the identification of segments (clusters) is accomplished in seven different steps, viz.: (1) selection of the segmentation variables, (2) deciding whether or not to standardize the involved variables, (3) deciding how to measure the similarity between respondents, (4) selection of a segmentation algorithm or procedure, (5) selecting the most appropriate number of segments, (6) interpretation of the identified segments, and (7) replication and validation of the identified segmentation solution. These steps are used as a framework to review developments and to identify potential issues and points for improvement.

(1). Selection of the segmentation variables

In this step we select the set(s) of variables that is (are) used in the analysis to identify segments. The variables actively used in the analysis to identify segments are called the segmentation bases. Segmentation bases may include demographic variables, socioeconomic variables, product usage and attitude variables, or benefits derived from the product or service. Benefits refer to the values or utilities respondents attach to the various levels of products or services. It is known that including variables on which segments are not distinguished may obscure the segmentation structure. This phenomenon refers to the issue of variable selection.

Milligan (1980), for example, found that the addition of even one irrelevant variable seriously reduced the extent of cluster recovery. Another issue involves the number of variables to be considered in one cluster analysis. Of course, this question cannot be answered in general. Increasing the number of variables increases the probability of including an irrelevant variable. In section 5.1 we discuss this issue further.

(2). Deciding whether or not to standardize the involved variables

If all included variables are measured at interval scale level or higher, it is easy to standardize the variables by dividing them by their standard deviations. Edelbrock (1979) found no substantial differences between the uses of standardized versus non-standardized variables in his resulting classifications. The unstandardized data only performed reasonably in error-free conditions.

In general, we do not recommend standardization (see Fleis and Zubin, 1969) unless there is an explicit reason for doing so. Based on the evidence we recommend to analyze standardized data using the zRANGE-transformation, zRANGE = X/[Max(X)-Min(X)], (see Milligan and Cooper, 1988).

(3). Deciding how to measure the similarity between respondents

After the variables have been selected, the next step is to compute the similarities between the respondents (or more general the research units). Similarities between respondents can be computed using four different measures:

association coefficients (e.g. Simple matching coefficient, Jaccard‘s coefficient, coefficient of Gower),

correlation coefficients,

distance measures (e.g. Euclidean distance, City-Block and Mahalanobis distance), and

probabilistic similarity measures (see Everitt, 1993 for details).

The choice of the most adequate measure depends upon (1) the measurement level(s) of the variables, (2) whether interdependencies have to be taken into account and (3) which aspects of the data are most relevant (a distinction here is made between elevation, shape and scatter). Similarity between profiles can be decomposed in three parts (Cronbach and Gleser, 1953), viz.: (1) shape, the pattern of dips and rises across the variables, (2) scatter, the dispersion of the scores around their average and (3) elevation, the mean score of the case over all the variables (see Skinner, 1978).

In case all the variables are of interval scale measurement level basically all three kinds of similarity measures can be used and the other two factors will determine the choice. If all of the variables are measured at the nominal or ordinal level one can apply any of the association coefficients (e.g. coefficient of Jaccard). Thus, in those instances the problem is ‗just' selecting the most appropriate association coefficient. If we have variables with different measurement level, an option is the use of the association coefficient of Gower (1971). Other possibilities of dealing with mixed measurement levels exist in the finite mixture models. Distance measures can only be used if variables are measured at at least the interval level. To adequately take into account the variance-covariance matrix between the

variables the Mahalanobis distance is preferred over the Euclidean and City-Block distance measures (cf. Art, Gnadadesikan and Kettenring, 1982). Within the class of association coefficients, the coefficient of Gower is preferred (cf. Alderderfer and Blashfield, 1984).

Different similarity measures emphasize different aspects of the data. The correlation coefficient emphasizes shape, and the distance measures confound all three components of similarity (shape, scatter, and elevation). To illustrate, let us look how these three aspects appear in a conjoint context (See Table 2.1). Suppose we have three respondents and for each respondent we have part worth utilities for each of the three levels of one attribute:

Now, define D12 (shape) as the similarity between respondent 1 and 2 only taking shape into account. D12 (scatter) and D12(elevation) are similarly defined. In the example above: D12 (shape) = 0, D12(scatter) > 0, D12(elevation) > 0. What does shape tell us in this context? It tells us that the rank order of utilities is the same for both individuals. Is this information sufficient and could we thus cluster using correlation coefficients? No, because elevation and scatter also tell us something. Elevation takes into account the fact that for respondent 1 attribute 1 is perhaps more important that for respondent 2. Scatter takes into account how important the transitions within attributes are. Compare, for example, respondent 1 and 3. For both respondents the distance would be (near) zero if we would only use shape and elevation, but the distance would increase once the scatter aspect is taken into account. In the context of preference structure measurement results (e.g. part

worth utilities) a distance measure seems a reasonable choice, although the three aspects are mixed in a particular way and it is not clear whether this is an optimal way.

(4). Selection of a cluster analysis algorithm

From a users‘ perspective segmentation procedures can be classified into the following families of cluster analysis methods: (1) hierarchical methods (to derive hierarchical clustering schemes), (2) iterative non-overlapping partitioning methods, (3) overlapping and fuzzy clustering procedures, (4) mixture models or latent class procedures, and (5) factoring methods.

Hierarchical methods

The agglomerative hierarchical approach is well-known and many alternatives within this approach are available. The basic approach of hierarchical clustering is that we start with considering each respondent in the sample as a separate cluster. Subsequently, respondents are classified together in a cluster based on their estimated similarity (using one of the measures discussed in section 4.3). The available alternatives for performing a hierarchical clustering differ with respect to the way in which the similarity between respondents is determined, and subsequently the similarity between clusters is determined. The most well known alternatives to measure the similarity between clusters include _single linkage‘, _complete linkage‘, _average linkage‘, _centroid method‘, _median clustering‘, and _Ward‘s method‘. A few lesser-known approaches were proposed by Gowda and Krishhna (1978), and Hansen and Tukey (1992).

The various agglomerative hierarchical procedures have been compared empirically. Based on the results of several studies we recommended use of the group average method using Euclidean distances, and Ward‘s method (cf.

Edelbrock and McLaughlin, 1980). We note that in the context of conjoint segmentation Vriens, Wedel and Wilms (1996) also found Ward‘s method to outperform the non hierarchical k-means method. Divise methods start out with considering the whole sample as one segment, and then aim to partition the sample tree-like wise. These methods are rarely encountered in the literature, an exception being the CHAID approach which we discuss in section 6. However, some authors have argued that divisive methods can yield better results than agglomerative methods. Divisive methods are not discussed in this report (see for a discussion of a divise method Kaufman and Rouseeuw, 1990).

Iterative non overlapping partitioning methods

Iterative partitioning methods do not proceed according a tree-like approach. Instead, iterative partitioning methods start with a specific classification of the sample in a predetermined number of segments (e.g. either based on a random allocation procedure or based on an hierarchical procedure). Subsequently, respondents are ‗moved around‘, from one segment to another. If the solution improves the transfer is kept, otherwise the respondent goes back to his/her initial segment. Five characteristics distinguish the various partitioning methods: (1) the selection of seed points, (2) type of cluster assignment process or type of pass, (3) the statistical criterion used to assign the points to the cluster, (4) whether a fixed

or variable number of clusters will be formed, and (5) treatment of outliers in the solution.

There are two basic ways to start an iterative method: (i) the defintion of seed points and (ii) the selection of an appropriate starting partition. Seed points are estimates of cluster centroids. When seed points are defined, then the data points are assigned to the nearest cluster centroid on the first pass. We can also use a starting partition to select the seed points (e.g. by using a hierarchical cluster solution).

The type of pass refers to the way in which cases are assigned to clusters. Two approaches exist: (i) k-means passes and (ii) hill climbing passes. K-mean passes refer to the reassigment of cases to the cluster with the nearest centroid (new centroids can be either computed after each reassignment or after entire pass through the data). K-means algorithms have been proposed by Wishart (1975) and Spaeth (1980). Both procedures rely on MacQueens‘ method (MacQueen, 1967).

In the procedure used by Wishart each cluster is represented by its centroid and in the procedure proposed by Spaeth each cluster is represented by its median. Hill climbing passes change the allocation of cases based on whether or not the change improves the value on a particular statistical criterion. Several criteria have been proposed as an optimization criterion in the hill-climbing variant of the iterative partitioning methods3. Examples of such criteria are described in Everitt (1993) and include (i) minimization of trace W (where W refers to the within-cluster dispersion), (ii) minimization of the determinant of W, (iii) minimization of trace

BW-1 (where B refers to the between-segments-sum-of-squares, see Everitt, 1980), etc.

Each of these criteria is associated with particular types of clusters (cf. Aldenderfer and Blashfield, 1984). For example, the trace W criterion is biased in favor of hyperspherical very homogeneous clusters. A disadvantage of this criterion is that it can be affected strongly by transformations of the raw data (e.g. standardization). The det W criterion is not scale dependent, and does not assume that clusters are hyperspherical. To check whether the data can be better represented by non-hyperspherical clusters, we recommend the use in addition to the trace W criterion, the det W criterion. In using the det W criterion it is assumed that the clusters have the same shape and it has a tendency to create clusters of about equal size. To overcome the _similar shape‘ problem other optimization criteria have been proposed by, for example, Scott and Symons (1971) and Maronna and Jacovkis (1974). See also Everitt (1993), and Lukasova (1971). There have been a few Monte Carlo studies investigating the relative performance of the various iterative partitioning methods (Blashfield, 1980; Bayne et al., 1980; Mezzich, 1978; Milligan, 1980; Scheibler and Schneider, 1985). Overall, the k-means procedure and the hill-climbing procedure optimizing trace W were found to perform best, especially when centroids from Ward‘s procedure were used.

Based on such results, some authors have now proposed a sequential procedure (first hierarchical, then, using hierarchical results to set the seed points in the second stage, k-means). In general a good starting solution is very important when using iterative partitioning methods.

Only a few methods, such as ISODATA (Ball and Hall, 1965), allow for a variable number of clusters in the solution or for a residual pool of unassigned points. We will describe later how latent class methods are also able to deal with the latter issue.

Overlapping methods

Potential benefits of overlapping and fuzzy methods arise in situations of multiple uses of brands, multiple usage situations or multiple benefits sought. Overlapping methods allow subjects to have memberships in multiple segments. Fuzzy segmentation methods allow subjects to have partial memberships in various segments. See for additional details Arabie et al. (1981), and Roubens (1982). For a comparison, see Hruschka (1986). I have seen very few applications of this class of methods, and as such will not elaborate upon them.

Finite mixture models or latent class models

In mixture models it is assumed that the observations in a sample arise from two or more groups, of unknown proportions, that are mixed. The purpose is to unmix the sample and to identify the underlying groups or segments within a statistical framework. Mixture models have been especially useful in the field of cluster analysis, and represent a model-based approach to clustering, which connect clustering with classical statistical estimation methods. The approach is discussed further in section 7.

Factoring methods

Ordination techniques attempt to provide some type of dimensional representation, usually based on fewer variables than in the original dataset. In the context of clustering respondents, techniques such as Q-type factor analysis and multi-dimensional scaling can be used cf. Hagerty, (1985). The number of extracted factors need not be an adequate indicator of the number of segments present and is limited by the number of independent observations per subject (Stewart, 1981).

(5). Selecting the most appropriate number of segments

Milligan and Cooper (1985), investigated 30 decision rules for determining the appropriate number of segments. In reviewing the literature on this topic the conclusion is that selecting the most appropriate number of segments in traditional cluster analysis is cumbersome.

In practice, managerial relevance is often used to choose a number of clusters. To be managerially relevant the number of clusters must be small enough to allow complete strategy development. At the same time, each cluster or segment must be large enough to warrant such strategic attention and must be reachable, and defensible against competitors (de Kluyver and Whitlark, 1986, p. 280). To achieve this we recommend a close involvement of the marketing management of the client company. This involvement may come from the marketing manager, the sales people, or people responsible for marketing communications.

(6). Interpretation of the segments

Profiling - Once segments are identified they are usually profiled on both the original segmentation variables and other relevant variables such as demographics and brand usage. This can be done in various ways. One way is to use the segments as a banner in the tabulations. This yields cross tabulations with all the questions in the study. A second way is to use discriminant analysis. In this case the segment memberships are used as the dependent variable, and the profiling questions are use to predict the segment memberships. Those variables that are successful in doing so indicate the meaningful profiling variables.

Simultaneous clustering: improving profiling power. In practice, uncovering these profiling variables can be disappointing because the analysis does not show any statistically significant differences on the profiling variables. There does exist a way to further explore this, however, with a technique called simultaneous clustering. The approach makes use of the fact that in a typical k-means solution, there exist a number of respondents who could be transferred from one segment to another without deteriorating the ‗optimal' k-means solution too much. The method aims to simultaneously optimize a k-means criterion and the between-variance of the profiling matrix.

Once profiled, descriptive segment names are usually developed, almost always with client input. These brief descriptor names must be accurate. Haley's (1968) fruit (cranberry) positioning study (see DeBruikcker and Reibstein, 1983) shows the traditional way to profile segments and give them descriptive names.

Presentation methods

Methods of presentation include:

Presenting the raw means of the segmentation variables across segments (either numerical, through histograms, or spiderplots),

Averages and top box scores of the segmentation variables across segments,

Verbal descriptions of the segments,

Audio-visual _representations‘ of the segments.

In most studies, interpretation and profiling the segments is not sufficient. To develop a marketing plan for the selected or prioritized segments, we sometimes need to go beyond the mere facts (as is usually the case when formulating marketing implications). In Haley‘s cranberry study, for example, once the segments were profiled on their consumers‘ lifestyles and attitudes, Haley developed a chart with plausible buying incentives, promotional actions, copy visual, and recommended media for each of the segments separately.

(7). Replication and validation

Five approaches for validating cluster analysis solutions have been proposed and applied, viz.: (1) the cophenetic correlation coefficient, (2) significance test on variables used to create clusters, (3) replication, (4) significance tests on independent variables and (5) Monte Carlo procedures. The first two validation techniques have serious flaws and should not be used (cf. Aldenderfer and

Blashfield, 1984). The third technique involves the estimation of the degree of replicability of a cluster solution across a series of data sets. One way of doing this is by randomly splitting the sample in two and performing cluster analysis on both sub-samples. The degree to which both solutions are in agreement is a measure for stability. Several variations on this theme exist. The most extensive being consensus or replicated clustering (see Helsen and Green, 1991). A fourth approach for validating a cluster solution entails the use of variables not used to generate the cluster solution. The power of external validation is that it directly tests the cluster solution against relevant criteria. In context of preference structure measurement a possible way to implement this approach is to use holdout observations (for computing predictive accuracy) or a criterion like the *expected mean squared error of prediction* (Hagerty, 1986).

2.1.9 Some Practical Issues in Segmentation Analysis

1. How to deal with large numbers of variables: tandem clustering?

In commercial market research it is rather common to run into datasets that include large number of variables. We have encountered situations in which the client had included several hundred variables all of which were considered potentially important for their segmentation. How do we handle these situations? These large numbers of variables may arise as a result that there are certain topics that need to be covered by a series of variables. For example, the client may want to investigate what the respondent was feeling during a certain consumption experience. To address this one needs to include a set of variables. For example, consider a set of statements that the respondents indicate to agree or disagree with

(e.g. I was happy, I was feeling playful, I was depressed, I was excited, I felt like partying, I felt thirsty, I felt moody, etc. etc.). Practically, it is impossible to include large sets of variables into one cluster analysis. It is also not recommended because one not relevant variable can dramatically affect your segmentation solution. In addition to screening your variables based on their frequency distribution (e.g. eliminating variables that show a skewed distribution), we may use data-reduction techniques such as factor analysis to reduce the number of variables or to reduce the number of variables into a smaller set of underlying dimensions. After this first step we can proceed with our cluster analysis, now based on a manageable number of variables. This two-step procedure is sometimes called tandem clustering. Several authors have argued against this procedure (e.g. Arabie and Hubert, 1994; Wedel and Kamakura, 1998). However, especially in multiple item situations (multiple items to measure the consumer experience, to measure involvement, to measure life styles, etc.) it is not clear whether tandem clustering will negatively affect the segmentation solution (see Green and Krieger, 1995). We recommend therefore the use of tandem clustering in those situations that are characterized by a large number of variables arising from the use multiple item instruments. However, it is important to keep in mind the characteristics of this analysis procedure, basically: you throw away parts of the data that may have been important in the clustering. One way to protect against this is by running several different analyses, with several different procedures.

2. How to deal with missing data and outliers?

Several approaches can be used to deal with missing data. First, we can delete any respondent who has a missing value on any of the segmentation variables. Second, we can replace a particular missing value by the mean of the variable. More advanced procedures are possible, but we consider this to be beyond the scope of this research note. Not much literature is available on how to treat outliers. An observation can be an influential outlier by having a positive or negative effect on the solution. In the latter case it may become a candidate for deletion in the segmentation analyses.

3. How to deal with variables of mixed measurement level?

When using standard segmentation methods, as described in the previous section, it is possible to simultaneously analyze variables with different measurement levels. The problem to solve here is that we need a similarity measure to compute the similarity between subjects based on a set of variables with multiple measurement levels. Gower

(1971) proposed a coefficient for this purpose. Once we have computed the similarities between subjects we can apply standard clustering techniques. Another approach is similar to the tandem approach described earlier. We can start with pre-processing the (partly) categorical data by using correspondence analysis, multiple correspondence analysis, or even use dummy variables for the categorical variables. The goal of this pre-processing is to create a scale that can be used to compute Euclidean distance measures between respondents. This two-step approach may suffer from the same disadvantages as the tandem approach

described earlier. A third, more general, approach to deal with this problem is to use the latent class methodology that is described in the next section. If the nominal variables do not have too many categories we can transform them into dummy variables. We can do the same for ordinal level variables, although some information contained in the variable will be lost. The transformed variables are then directly included in the latent class analysis.

4. How to deal with sampling issues?

The following section is based upon Wedel and Kamakura‘s (1998) treatment of this issue. The basic issue here is that when your sample is not a simple random sample, drawn from the population of interest, the derived segments do not represent the segments in the population. This issue thus arises when using stratified sampling. In this case the respondents found in the identified segments had unequal probabilities of being in the sample. For example, suppose we stratified by company size, and we decided to interview managers from 100 very large, and 100 small companies. Suppose that segment 1 consists for the larger part of respondents from large companies. This is a biased finding because these respondents had a higher probability of being chosen in the first place, and therefore the estimate of the segment-size is biased, as is the estimate of segment means on the relevant variables. Wedel, Hofstede, and Steenkamp (1998) propose a general approach to handle these complex sampling cases; they propose the use of pseudo maximum likelihood methods. From a practical, commercial point-of-view, a lot can be gained here. Wedel, Hofstede and Steenkamp (1998) show that

ignoring the specific features of the sampling design results in an incorrect number of segments and biased estimates of the segment-level parameters.

2.1.10. Market Segmentation with CHAID

Most of the segmentation techniques discussed so far are descriptive methods, i.e. there is not a dependent variable on which the segments are based. There are several segmentation techniques, like discriminant analysis, clusterwise regression, latent class regression models, and the CART and CHAID (CHI-squared Automatic Interaction Detection) approaches that can be used when there is a dependent variable. The other variables are used to explain/predict the dependent variable. The CHAID technique is a widely used method that is easy to apply. A thorough overview of this technique can be found in Magidson (1988, 1994). Below, we outline the main elements of this approach.

The CHAID approach

Chi-square automatic interaction detection (CHAID) is based on a stepwise analysis-of variance procedure (Kass, 1980). In ANOVA, the dependent variable is interval-scaled and the independent variables are categorical. However, the CHAID approach is used when the following components are present:

1. A categorical dependent variable,

2. A set of categorical independent variables (continuous may be used as well but the first step of the analysis is to categorize them),

3. Settings for the various CHAID parameters.

At any point in the CHAID analysis, some subgroup is being analyzed and the ―best‖ predictor is identified. In CHAID categories of the independent variables are merged.

Having a set of several independent variables, each having several categories implies that many different mergers are possible in, say, the first step of the analysis. CHAID will merge those categories of an independent variable that are homogeneous. The question thus becomes ―Which categories can be best merged together?‖ This question is answered in basically two steps:

Within an independent variable. That is which categories can be best merged within a specific independent variable.

Across the independent variables. After we looked within each of the independent variables to select the best mergers within these variables, we have to look which categories should be merged first across the independent variables.

Within an independent variable, we look for the two categories that best can be merged together. For each pair of categories that can be merged together we make a cross table with the dependent variable and compute the corresponding c2 value, and p-value. We do this for all possible mergers in that variable. The ‗best‘ merger here is defined as the one that results in the **highest** *p*–value in a c2 analysis. Recall, that if the c2 statistic of a cross-table is statistically significant this means that we have to reject the hypothesis of independence, and this implies that there may be a relationship between the two variables making up the cross table. Thus, the *p*-value represents the probability that the observed sample relationship

between a predictor and the dependent variable would occur if in fact the two variables were statistically independent. Thus, a high p-value makes that it is very unlikely that the two variables are related: i.e. we are merging categories that are unrelated with the dependent variable.

After this is done, we look at the results across the independent variables. The predictor (i.e. independent variable) with the **lowest** *p*-value is the one that is most likely to be related to the dependent variable. The first split made is the split based on this variable. This is so because if we compare all possible mergers at a given point the one merger resulting in the lowest p-value maximizes the relationship between the dependent variable and the ‗merged' independent variable. More formally, the CHAID algorithm can be described as follows:

Stage 1: Merging (for each independent variable).

1. Form the full two-way cross-tabulations with the dependent variable.

2. For each pair of categories that is eligible to be merged together, compute the c2 statistics to test for independence.

3. For each pair wise c2, compute the corresponding ―pair wise‖ *p*-value. Among the pairs that are found to be non-significant merge the most similar pair (i.e. the pair having the smallest pair wise c2 value) into a single joint category, and go to step 4. If all the remaining pairs are significant, go to step 5.

4. For any joint category containing three or more categories, test to see if any predictor category should be unmerged by testing the significance associated with that category versus the others in that joint category. If a significant c2 is obtained, unmerge that category from the others. Return to step 3.

5. Compute the Bonferroni adjusted *p*-value based on the after-merged table. This adjustment should be used since a series of tests are used that increase the probability of making a so-called type I error. An independent variable having many categories has many possible splits, and thus the probability of finding a low p-value is higher as compared to an independent variable having only two categories.

Stage 2: Splitting

Select as predictor variable the one with the lowest significant adjusted *p*-value. Split the groups in this variable according to the optimally merged categories of that variable. This defines the subdivision of the group in subgroups.

Stage 3: Stopping

Return to step 1 to analyze the next subgroup. Stop when no independent variable has a significant *p*-value or when a subgroup contains too few observations.

The result is a treelike group of segments; it is a so-called tree-building method. The difference with a standard cluster analysis is that the segments are based on a dependent variable, and are thus derived for prediction. Another difference with standard cluster analysis is that the derived segments are not based on the same variable, i.e. it discovers interaction effects among variables. Furthermore, variables that are not discriminating may be incorporated in the analysis in the first stage but will not be included in the final model.

The main advantage of CHAID is that the output is easy to understand and therefore easy to communicate to the management. It is, however, often seen as a useful exploratory data analysis technique of which output may be used as input for other techniques with more predictive power (Shepard, 1995).

Some issues

The CHAID modelling presented above can handle dependent data with more than two categories. In the case these data are measured at the interval level, the procedure can still but used, but there is a substantive loss of statistical power. This is caused by the fact that the c2 statistic does not take the properties of the interval aspect into account. By using a Y-association test (e.g., Magidson 1992) instead of a c2 test, the interval aspect is taken into account. This method is part of the CHAID (*answer tree*) module in SPSS. Since CHAID is a (forward) stepwise approach, i.e. once an independent variable is chosen it cannot be eliminated in a later stage, there is the possibility that a better segmentation solution can be found by using an other variable in an earlier stage. In the last version of CHAID in SPSS there is the option to perform an exhaustive CHAID, which derives all possible orderings and the ―best‖ solution is chosen. This is however very time demanding.

The basic idea of CHAID is that the segments are solely based on statistical criteria. However, this will not always generate the optimal segmentation with respect to the six criteria for successful segmentation. An advantage of CHAID is that the researcher, in cooperation with the marketer, can easily adjust the derived CHAID-tree manually. For example, redefine a split of an independent variable. In this way a balance can be obtained between statistical criteria and managerial arguments.

An important issue is the specification of the stopping rules and other settings. This includes the minimum number of observations in a group that is eligible for splitting, the minimum number of observation in a new created group, the maximum depth of the tree, the significance level at which an independent variable is used and the merging criterion. Unfortunately there are no objective criteria to choose these. It is often a balance between a too small a tree that is not very useful and a too large a tree that is hard to interpret. Depending on the sample size the settings are chosen. If the model is solely used for prediction, e.g. target selection in direct marketing, then the interpretation is less important, we advise to allow for a large tree.

A closely related technique is the so-called classification and regression trees (CART), which was developed by Breiman et al. (1984). The result of this technique is also a treelike group of segments. In contrast with CHAID, the independent variables may be categorical and continuous. The splits are based on some homogeneity measure between the segments formed. Like CHAID, it is a useful method for explorative data analysis and also easy to communicate. A comprehensive discussion on CART can be found in Haughton and Oulabi (1993) and Thrasher (1991).

Applying CHAID in SPSS

Recently SPSS developed a new module for CHAID analysis, the so-called answer tree module. Apart from CHAID, it contains three other methods. The program allows the researcher to define the way variables may be merged (e.g. only

consecutive categories or all categories), the stopping rules etc. In principle, at each step the researcher can manually redefine the tree with respect to the choice of the independent variable, the definition of the split etc. Uninteresting parts of the tree do not have to be developed completely. Often, it is a trial-and-error process to obtain the optimal CHAID-tree.

2.1.11. Finite Mixture Models (Latent Class Segmentation Models)

Finite mixture models refer to a class of procedures that deal with heterogeneity in the data at hand. It is assumed that the data is a mixture of homogenous sub-groups, called components of the mixture (Dillon and Kumar, 1994). Mixture models have been especially useful in the field of cluster analysis, and represent a model-based approach to clustering, which connect clustering with classical statistical estimation methods.

1. Advantages of Latent Class Segmentation Models

Finite mixture models offer several advantages over the more traditional techniques as discussed in the previous sections:

> Mixture models identify market segments, and provide unbiased market segment memberships estimates (i.e. Dillon and Kumar, 1994). For example, when using standard cluster analysis techniques, the sample means of the segments identified from our data may be inconsistent estimates of the population means of these segments. The use of mixture models results in consistent estimates of these parameters,

Mixture models have an opportunity to perform statistical testing to compare different segmentation solutions (e.g. choosing between a two-segment and a three-segment solution), It has been shown repeatedly, that finite mixture or latent class approaches outperform more traditional approaches (see Vriens, Wedel, and Wilms, 1996; Wedel and DeSarbo, 1993),

Mixture models enable us to simultaneously take into account variables measured on different measurement levels. It is also possible to transform nominal and ordinal variables into dummy variables or scale them in to interval level variables and then include them in the analysis,

It allows for both unconditional segmentation (i.e. there is no predictive model underlying the segmentation) and conditional segmentation (i.e. latent class regression models),

Mixture models allow for an explicit treatment of cases where complex samples are involved (Wedel, Ter Hofstede, and Steenkamp, 1998), and

As a last potential advantage we mention that the latent class approach explicit allows us to deal with specific difficulties encountered in identifying international segments (see next section).

2. Basic Concepts

To formulate a mixture model we assume that our respondents arise from a population that is a mixture of S segments in proportions that are unknown to us. Briefly, the specification and estimation of a finite mixture (regression) model proceeds along the following steps:

1. Formulate the statistical distribution of subject i, given (conditional upon) that this distribution comes from segments,

2. Formulate the unconditional distribution (which is a weighted sum of the conditional distributions),

3. Formulate the likelihood (given the assumption that the subjects are independent this is the product of individual distribution functions),

4. Estimate the mixing proportions and model parameters, and

5. Compute the posterior probabilities using Bayes‘ rule.

In more detail this process can described as follows (following Wedel and Kamakura,

1998).

Let:

i = 1,…,I denote the subjects (i.e. respondents, businesses, etc.),

j = 1,…,J denote the set of segmentation variables,

s = 1,…,S denote the segments,

yi = (yij) denote the variable representing the scores of respondents i on variables j,

In order to formulate the finite mixture model, assume that that the objects on which the variables y=(yij) are measured arise from a population which is a mixture of S segments, in proportions a1,...,aS. It is not known, in advance from which segment a particular object arises. The proportions are constrained:

$$\sum_{s=1}^{S} \alpha_s = 1, \alpha_s > 0, s = 1, \ldots., S$$

Given that **yi** comes from segment s, the statistical distribution function of the vector **yi** is represented by the general form fs(yi/θs). θs denotes the vector of all unknown parameters associated with the specific density chosen (e.g. if the **yi** are uni-variate nomal distributed, we would estimate the mean and variance within each segment). The simple idea behind the mixture distributions is that if the distributions conditional upon knowing the segments have been formulated the **unconditional distribution** of **yi** is obtained as :

$$f\left(\frac{y_i}{\Phi}\right) = \sum_{s=1}^{S} \alpha_s f_s\left(\frac{y_i}{\Phi_s}\right)$$

Where Φ =(α, θ). The purpose is to estimate Φ. The likelihood function of Φ is:

$$L\left(y_i, \Phi\right) = \prod_{i=1}^{I} f\left(y_i / \Phi\right)$$

The likelihood of finite mixtures can be maximized basically in two ways: (1) using standard optimization routines, such as Newton-Raphson routine, and (2) by using the Expectation-Maximization algorithm.

Given that an estimate of F has been obtained, estimates of the posterior probability, pis, that observation i comes from mixture component s can be calculated for each observation vector yi by means of Bayes‘ Theorem, where this posterior probability is given by:

$$P_{is} = \frac{\alpha_s f_s(\frac{y_i}{\theta_s})}{\sum_{s=1}^{s} \alpha_s \, f_s \, (\frac{y_i}{\theta_s})}$$

3. Selection of segments

Suppose we want to test whether a _K+1'-segment solution is better than a K-segment solution. The standard likelihood ratio tests cannot be used since this test is not asymptotically chi-squared distributed. A number of alternatives exist for selecting the appropriate number of segments:

The AIC (Akaike‘s Information Criterion) and CAIC measures (Consistent Akaike‘s Information Criterion; see Bozdogan, 1987).

AIC = -2LnL-Pd (where P is the number of parameters to be estimated, and d is an arbitrary constant). For the AIC d = 2, for the CAIC, d = Ln(N+1),

The BIC (Bayesian Information Criterion). For the BIC, d = Ln(N), see Schwartz (1978), and

The entropy statistic (Ramaswamy et al., 1992).

If we use the information measures lower values indicate a better solution. By plotting the values of an information measure against the number of segments, we can look for the breaking point. The breaking point is the number of segments beyond which the reduction in the information statistic is substantially reduced. The above statistics account for over parameterization. In addition, we must assure that the segments are sufficiently separated. To assess the separation of the segments, the entropy statistic can be used to investigate the degree of separation in the estimated posterior probabilities. Values close to 1 indicate that the derived segments are well separated. A value close to zero, indicating that all the posteriors are equal for each observation, is of concern as it implies that the centroids of the segments are not sufficiently well separated. See next section for how we interpret the E value.

2.1.12. Latent class regression models

In the finite mixture model approach, described above, we estimated only the expected values for each of the underlying densities (and dispersion parameters, depending on the distribution assumptions made). It is possible to extent this concept to conditional expectations, or regression models. In this case, the expected values of the underlying densities are estimated, whereby the expectations are formulated as a function of a set of explanatory variables. The purpose here is to search for segments in such a way that the fit of the regression models within each of the segments is maximized. Extensions of the basic, latent class regression, model are the latent class metric conjoint analysis model (DeSarbo, Wedel, Vriens, and Ramaswamy, 1992), and the latent class discrete

choice model (DeSarbo, Ramaswamy, and Cohen, 1993). As an example that is often encountered we describe the latent class metric conjoint analysis model below. Latent class metric conjoint analysis DeSarbo et al. (1992) and Wedel, Vriens and DeSarbo (1991) propose latent class approaches for metric conjoint analysis. These approaches are a generalization of DeSarbo and Cron's latent class methodology (1988). In the latent class metric conjoint analysis approach, the segments and the conjoint model parameters are estimated simultaneously, employing a stochastic framework. The method employs a stochastic framework involving mixtures of multivariate conditional normal distributions. The model is estimated by maximizing the likelihood through an expectation-maximization algorithm (EM-algorithm; Dempster, Laird, and Rubin, 1977). The latent class approach allows for the calculation of posterior probabilities of the memberships of subjects in segments. The model propsed by Wedel, Vriens and DeSarbo (1991) is somewhat simpler and thus we describe this model first.

Let i, j, s, and yij, be defined as before

Let further:

k = 1,…,K denote the kth profile,

Xkj = the value of the jth variable on the kth profile,

bjs = denote the regression parameter for the jth conjoint dummy variable within segment s,

Assume that in a conjoint analysis study, ratings of preferences, **yik** of I respondents with respect to K profiles have been collected. Then, within each segment the IxK vector of profile ratings can be expressed as:

$$Yiks = X\beta s + \varepsilon_{iks}$$

Within segment s subjects are assumed to have identical part-worths for the attribute levels, With

$$\varepsilon = N\ (0,\ \sigma^{2})$$

i.i.d. distributed, and bs the segment-specific part-worth utilities (i.e. regression coefficients). All terms are further defined as before. Let also as be defined as before.

The purpose now is to estimate **$\Phi = (\beta s, \sigma_{\epsilon}, \alpha_{\epsilon})$.** If we assume that the yis are independently normally distributed, the conditional distribution of y i, the (Ix1) vector of ratings for subject i, conditional upon subject i being in segment s is:

$$f_s\left(\frac{y_i}{\beta_s, \sigma_s}\right) = \prod_{j=1}^{J}(2\pi\ \sigma_S^2)^{-1/2}\exp\frac{-(y_{ij} - X_{ij}\beta_s)^2}{2\sigma_s^2}$$

Since yi is drawn randomly from a mixture of densities of the underlying segments in unknown proportions, α_1,..., α_s, the unconditional distribution function can be formulated and the likelihood is obtained as the product of this distribution across subjects. By deriving the partial derivatives with respect to the parameters and setting these to zero we obtain closed expressions for the αs‘s and βs. See for details Wedel, Vriens and DeSarbo (1991). For finding a maximum of the likelihood function we can use the E-M-algorithm. The above model has been

extended to relax the assumption of i.i.d. distributed error terms by DeSarbo, Wedel, Vriens and Ramaswamy (1992).

2.1.13. International Market Segmentation Approaches

In order to be able to identify international market segments, several problems need to be recognized and dealt with. Three important problems may be encountered and should be addressed:

The measurement instruments used in the research are subject to response tendencies that, as is well known, may differ across countries. Therefore, in international research we need to take into account the fact that respondents in different countries may display different response tendencies,

In international marketing research, stratified sampling strategies are often encountered where the sample is stratified by country. The result is often equal sample sizes allocated to each country. Not taking this into account would lead to biased estimates. A valid international segmentation approach would need to take this into account. See Wedel, Ter Hofstede and Steenkamp (1998) for an elaborate discussion of this topic, and

A practical strategy needs to be developed to identify truly international segments

For the latter two issues systematic solutions are available and discussed below. For the first issue, no general solution is known. In specific cases satisfactory solutions are available.

1. Sampling issues

Consider a case where a stratified sampling strategy is used, whereby the sample sizes for the different countries are more or less equal, while the population sizes in the different countries differ substantially. In this case, the probabilities that members from a certain market segment are entered in the sample, within a particular country, are not. We note that straightforward applications of latent class methods, or more traditional types of clustering methods for that matter, are based upon the assumption of simple random sampling. The solution for this situation is to use a pseudo-maximum-likelihood approach (see Wedel, ter Hofstede and Steenkamp, 1998). It has been found that ignoring the sampling characteristics may lead to severely biased estimates of the number, size, and nature of the market segments.

2. International segmentation schemes versus country-specific Segmentation schemes: a practical strategy

There are two hypotheses possible:

1. Each (or some) of the involved countries have the same number and type of segments in common (or have some of the segments in common), although the size of the international segments may differ. We refer to this as an international segmentation scheme.

2. The involved countries have no common segments. This means that the segments are not comparable in terms of nature of the segments: i.e. the identified segments differ structurally on the variables on which the segments are identified. We refer to this as a country-specific segmentation scheme.

If clients are interested to investigate the possible existence of international segments we propose to proceed in the following way.

1. Start with the application of the concomitant variable latent class segmentation models (see Kamakura, Wedel, and Agrawal, 1994). In this approach the first hypothesis described above is used. In the concomitant variable, latent class segmentation models the prior and posterior segment membership probabilities are modelled as function of a set of external variables. In international segmentation these variables include country membership variables (defined as a series of dummy variables). This approach allows us to test whether the segment-sizes and posterior segment-membership probabilities vary across the levels of the concomitant variables, for a particular segment. For example, if all the concomitant variables representing country-memberships are zero, this indicates an international segmentation, spanning all of the involved countries, where each segment is equally big. If certain concomitant variables are statistically larger than/smaller than zero, this means that segment sizes differ across countries. If the parameters of the concomitant variable models are statistically different from zero, we know that the identified segments differ in sizes across countries, and that the respondents in the various countries have different probabilities of belonging to a particular segment. We note that this first step also can be done without concomitant variable models, by first starting with a standard latent class model,

followed by a discriminant analysis in which the country dummies are used to predict segment membership.

2. To assess whether the respondents all fit in the international segmentation scheme we look at the posterior probabilities. The information contained in the posterior probabilities can be summarized in a statistic: the entropy statistic, E. If E approaches 1 this means that respondents have either probability 0 or 1 that they belong to a certain segment. Hence, this points to a confident classification. If E approaches 0 this means that subjects are difficult to classify and that they do not agree with the derived international segmentation. If the latter situation arises we need to investigate the second hypothesis stated above: the country specific segmentation. We note that this is a very strong advantage of the latent class approach because the posterior probabilities allow us to do this.

2.1.14. Application Recommendations

Although it is difficult to give recommendations that are generally applicable, we believe that it is very useful from practical point of view to have some guidelines.

Descriptive/predictive

If there is a predictor variable, it is advised to use a method that is based on this. It will give segments that discriminate optimally with respect to the variable of interest. Hence, use CHAID or mixture regression models.

Use of latent class methods

Latent class methods outperform the traditional segmentation methods. This holds for descriptive as well as for predictive methods. The disadvantages are that latent

class models are harder to estimate and harder to understand for researchers that are not so familiar with statistical techniques. At this moment, the strongest recommendation for the application of latent class model is in the areas: (1) conjoint segmentation, (2) international market segmentation, and (3) segmentation under complex sampling conditions.

Hierarchical/non-hierarchical

In practice one often only uses the non-hierarchical clustering method. This is risky since the method is sensitive for the initial choice of the cluster centers. On the other hand, if one only uses hierarchical methods, the problem is that observations cannot swap between different clusters. Any misclassification at a particular stage cannot be corrected later on. Therefore, we recommend starting with a hierarchical clustering to obtain a reasonable estimate of initial cluster solutions and to obtain an idea about the number of segments. Then, subsequently, use a non-hierarchical procedure to fine-tune the solution. In the hierarchical model it is advised to use Ward‘s method.

Standardization

If the range of values of particular variables differs heavily, it is actually necessary to standardize the data. Otherwise the variables with the largest range will dominate the analysis. It is recommended to standardize them by dividing the score by the range of that variable (zRANGE = X/[Max(X)-Min(X)]. Standardizing the variables when the scale of the variables is (nearly) identical is not recommended.

Reducing the number of variables

Often there are a large number of variables that one wants to include in the segmentation.

On one hand this can sometimes give problems. For example:

The algorithm cannot handle the size of the dataset,

Some types of variables may dominate the analysis. For example, when there are ten price related variables and 2 quality related variables, the first group will have a larger impact on the segmentation. In this case it can be useful to use factor analysis, although it is not recommend from a theoretical point of view. The results of factor analysis can be used in two ways. First, use the factor scores, as the new variables on the segmentation will be performed. Second, choose per factor the variables with the highest factor loading and use these for the analysis. The advantage of the first approach is that the information of all variables is used. The disadvantage is that it is often quite hard to define the factors and by that the segments. It depends on the situation which approach is preferred.

By increasing the number of variables we have a higher probability of including an irrelevant variable that can deteriorate the solution.

2.1.15. Market Segmentation and Marketing Management

Market segmentation results need to lead to a market segmentation strategy. This encompasses more than the numerical and statistical results of using the techniques described in the previous sections. To ensure the relevance of the identified market segments, and to ensure the identified market segmentation fits with the organizational strengths and weaknesses (SWOT-analyses), and can be

fully implemented in the marketing planning of the clients‘ company. We recommend to work with multiple segmentation bases since this increases the extent to which we are satisfying our segmentation criteria. For each set of segmentation variables we perform separate initial segmentation analyses. After these initial analyses we recommend to discuss these preliminary findings extensively with the several groups within the clients‘ company, i.e. the market researchers, their internal client, sales people, marketing communication people, engineering, and product development staff. After receiving their feedback, attempts can be made to synthesize the segmentation analyses into one market segmentation, that addresses all of the segmentation criteria.

2.1.16. The summary of above literature

What is Market Segmentation?

Market segmentation describes the division of a market into homogeneous groups which will respond differently to promotions, communications, advertising and other marketing mix variables. Each group, or ―segment,‖ can be targeted by a different marketing mix because the segments are created to minimize inherent differences between respondents within each segment and maximize differences between each segment. Market segmentation was first described in the 1950‘s, when product differentiation was the primary marketing strategy used. In the

1970‘s and 1980‘s, market segmentation began to take off as a means of expanding sales and obtaining competitive advantages. In the 1990‘s, target or direct marketers use many sophisticated techniques, including market

segmentation, to reach potential buyers with the most customized offering possible.

Why Use Market Segmentation?

There are many good reasons for dividing a market into smaller segments. The primary reasons: Easier marketing. It is easier to address the needs of smaller groups of customers, particularly if they have many characteristics in common (e.g. seek the same benefits, same age, gender, etc.). Find niches. Identify under-served or un-served markets. Using ―niche marketing,‖ segmentation can allow a new company or new product to target less contested buyers and help a mature product seek new buyers. More efficient use of marketing resources by focusing on the best segments for your offering— product, price, promotion, and place (distribution). Segmentation can help you avoid sending the wrong message or sending your message to the wrong people.

When Do You Use Market Segmentation?

Any time you suspect there are significant, measurable differences in your market, you should consider market segmentation. Identified segments must be big enough. Market must be large enough to warrant segmenting. Don‘t try to split a market that is already very small. Differences must exist between members of the market and these differences must be measurable through traditional data collection approaches (i.e., surveys). Once the market is segmented, you must be able to design marketing communications that address the needs of the desired segments. If you can‘t develop promotions and advertising that speak to each

segment, there is little value in knowing that those segments exist. Each segment must be reachable through one or more media. You must be able to get your message in front of the right market segments for it to be effective. If one-eyed, green aliens are your best marketing opportunity, make certain there is a magazine, cable program or some other medium that targets these people (or be prepared to create one). Segments must not only differ on demographic and psychographic characteristics, they must also differ on the benefits sought from the product. If everyone ultimately wants the same things from your product, there is no reason to segment buyers. However, this is seldom the case. Even commodities like sugar and paper plates can benefit from segmentation. The expected profits from expanding your markets and more effectively reaching buyer segments must exceed the costs of developing multiple marketing programs, re-designing existing products and/or creating new products to reach those segments.

How Do You Segment a Market?

There are two basic ways to segment a market:

A priori. A priori segmentation involves dividing a market into segments without the benefit of primary market research. Manager intuition, analysis of secondary data sources, analysis of internal customer databases or other methods are used to group people into various segments. Previous ―post hoc‖ segmentation studies are considered to be ―a priori‖ when applied to the same markets at some point in the future. Some examples of a priori segmentation schemes are: Heavy versus moderate and light users, Men versus women , Younger versus older users, North

versus south regions, VALS or PRIZM clusters, Buyers versus non-buyers, Re-application of previous post hoc segmentation schemes, Post hoc. Primary market research is used to collect classification and descriptor variables for members of the target market. Segments are not defined until after collection and analysis of all relevant information. Multivariate analytical techniques are used to define each segment and develop a scoring algorithm for placing all members of the target market into segments.

What Information Do You Use to Segment a Market?

There are two types of information used in market segmentation: Classification variables. Classification variables are used to classify survey respondents into market segments. Almost any demographic, geographic, psychographic or behavioral variable can be used to classify people into segments. Demographic variables — Age, gender, income, ethnicity, marital status, education, occupation, household size, length of residence, type of residence, etc. Geographic variables

— City, state, zip code, census tract, county, region, metropolitan or rural location, population density, climate, etc. Psychographic variables — Attitudes, lifestyle, hobbies, risk aversion, personality traits, leadership traits, magazines read, television programs watched, PRIZM clusters, etc. Behavioral variables — Brand loyalty, usage level, benefits sought, distribution channels used, reaction to marketing factors, etc.

Descriptor variables. Descriptors are used to describe each segment and distinguish one group from the others. Descriptor variables must be easily obtainable measures or linkable to easily obtainable measures that exist in or can

be appended to customer files. Many of the classification variables can be considered descriptor variables. However, only a small portion of those classification/descriptor variables are readily available from secondary sources. The trick is to identify descriptor variables that effectively segment the market in the primary research effort which are also available or can be appended to individual customer records in customer databases. This permits us to operationalize the market segmentation scheme developed in the primary research effort by applying it to existing customer and market information. DSS utilizes a number of proprietary procedures to achieve this important linkage. Here are some examples of descriptor variables: Census bureau demographic characteristics. Third-party classification variables like PRIZM or VALS

Geographic characteristics or regions, Panel data and scanner data on buying habits and usage levels, Customer data collected by companies for internal use

What Analytical Techniques Are Used to Segment a Market?

Most multivariate analytical techniques can be used and probably have been used in some way to create post hoc market segments. There is no ideal methodology that works with every segmentation study. Each methodology has advantages and disadvantages. Segmentation studies generally require the use of two or more methodologies to produce the best results. In nearly every case, multiple techniques should be tested before selecting the ―best‖ solution. There are 3 categories of analytical techniques applied to market segmentation: data

preparation, data analysis, and classification. The most common techniques for each category are:

Data Preparation

Factor analysis

Correspondence analysis

Conjoint analysis

Data Analysis

Cluster analysis

Chi-square Automatic Interaction Detection (CHAID) or Classification and Regression

Trees (CART)

Artificial neural networks

Latent class structure

models Classification

Discriminant analysis

Multiple regression

Multivariate logit

Multidimensional scaling (MDS)

Each of these analytical techniques, as well as other techniques not listed, can be applied to survey data to produce market segments. Below, we briefly describe how they are often used in segmentation studies.

Data Preparation

Numerous techniques can be used to aid the segmentation process. Factor analysis can reduce the number of variables to a more manageable size while also removing correlations between each variable. The coordinates produced by correspondence analysis, when calculated at the individual or group level, can be clustered to produce market segments. Correspondence analysis can also be used to convert nominal data (like yes/no answers) to metric scales. Utilities from conjoint analyses can be used in segmentation because they represent the relative value individuals‘ place on all key attributes that define a product or service. In fact, conjoint utilities represent the most effective basis variables because they are derived from respondent preferences between product options or from actual choices of preferred products.

Data Analysis — Cluster Analysis

Cluster analysis is the most frequently used method of segmenting a market. The underlying definition of cluster analysis procedures mimics the goals of market segmentation: to identify groups of respondents in a manner that minimizes differences between members of each group while maximizing differences between members of a group and those in all other groups. However, there is one key difference between clustering and segmenting respondents — clusters produce groups of respondents who have similar responses on key variables while segmentation finds groups of respondents who have similar behaviors when purchasing and seeking products in the market.

Both hierarchical and iterative cluster analysis procedures can be used, but hierarchical procedures are difficult to evaluate once you exceed 100 or 200 survey respondents. Among the various iterative cluster analysis procedures, the K-Means method is most often used. K-Means cluster analysis can be found in all of the most popular statistical programs (SAS, SPSS, BMDP, Statistica, SYSTAT).

Data Analysis — CHAID and CART

CHAID and CART are known as ―Classification Tree Methods.‖ These methods divide respondents into groups and then further divide each group into subgroups based on relationships between segmenting basis variables and some dependent variable. The dependent variable is usually a key indicator such as usage level, purchase intent, etc. These procedures create tree diagrams, starting at the top with all respondents combined and then branching into 2 or more groups at each new level of the tree. Subdivisions are determined by finding the survey variable that produces the greatest difference in the dependent variable among individual response categories or groups of response categories on that survey variable.

CHAID is the most commonly used classification tree method, but it cannot handle continuous dependent variables so a combination of CHAID and CART is sometimes used. Both CHAID and CART have the ability to process non-metric and non-ordinal data.

Unlike cluster analysis, classification tree methods create true segments when they divide respondents. However, these segments are only based on one dependent variable. Other methods, including cluster analysis, divide respondents based on 10‘s or even 100‘s of data elements.

Data Analysis — Artificial Neural Networks

Artificial Neural Networks or ANNs offer another means to segment respondents. The Kohonnen architecture is one self-organizing ANN that can be used for segmentation. It is called self-organizing because, like cluster analysis, there is no dependent variable specified in the model. The ANN attempts to group respondents based on their similarities. It differs from cluster analysis in its ability to ignore noisy data. Atypical individuals have less impact on the segmenting calculations and each successive iteration makes ever smaller adjustments to the network weights so the calculations quickly stabilize, ignoring infrequent respondent characteristics. The greater the variation or uncertainty in respondents‘ answers, the better ANNs perform compared to cluster analysis.

Data Analysis — Latent Class Structures

Latent class analysis is often described as ―factor analysis for categorical variables.‖ It is used to find underlying constructs within sets of variables. However, latent class analysis can also be used to cluster categorical variables into segments based on responses across a broad range of categorical variables. Latent classes attempt to find the underlying constructs which motivate people to buy a particular product or to desire certain features in that product.

Classification Algorithms

There are a number of classification algorithms or analytical methods which can be applied to market segmentation. Discriminant analysis can be used to classify respondents into predefined segments based on descriptor variables like census data. The segmentation scheme determines which respondents belong in each market segment. The classification or scoring program then creates the means of identifying potential members of each segment based on limited information (usually data which can be obtained from secondary sources). When a limited set of information can be used to accurately predict which market segment each individual belongs, you have a successful classification algorithm. Multiple regression and multinomial logic can be used in the same manner to create classification schemes for your market segments.

How Many Segments Should I Have?

Unfortunately, there is no definitive answer. Experience, intuition, statistical results and common sense all must be applied to decide on the number of segments to retain. If you have several very small segments, you may need to change the criteria for segmentation or remove some of these respondents as outliers. Too many segments can lead to developing many different marketing programs for small, very similar, markets. Here a few rules of thumb for segmentation:

Large enough: Majority of segments must be large enough to be economically feasible to target marketing and product design efforts.

Relevant: The segments must be relevant to your company‘s products/services.

Reachable: Segments must be reachable through one or more marketing mix variables (price, promotion, features or distribution).

Different: There must be clearly defined differences between market segments to make some segments more desirable than others. If many of the segments want essentially the same features and intend to buy at the same frequency or volume level, then these segments do not exhibit meaningful differences.

How Much Does Market Segmentation Cost?

Because of the amount of information that must be collected and the detailed analyses that must be conducted to identify segments, market segmentation is one of the more expensive research projects you will consider. Telephone surveys of more than 30 minutes are common and multi-phase projects using combinations of telephone and mail surveys are the norm when collecting data for market segmentation.

Although expensive, very few research projects can have the long term impact that market segmentation can produce. This research methodology provides the information necessary to identify new markets, redesign marketing programs and increase profitability. Few research projects can achieve more than one of those goals.

2.2 Retail in India

The Indian retail industry is divided into organised and unorganised sectors. Organised retailing refers to trading activities undertaken by licensed retailers, that is, those who are registered for sales tax, income tax, etc. These include the corporate-backed hypermarkets and retail chains, and also the privately owned large retail businesses. Unorganised retailing, on the other hand, refers to the traditional formats of low-cost retailing, for example, the local *kirana* shops, owner manned general stores, *paan/beedi* shops, convenience stores, hand cart and pavement vendors, etc.

India‘s retail sector is wearing new clothes and with a three-year compounded annual growth rate of 46.64 per cent, retail is the fastest growing sector in the Indian economy. Traditional markets are making way for new formats such as departmental stores, hypermarkets, supermarkets and specialty stores. Western-style malls have begun appearing in metros and second-rung cities alike, introducing the Indian consumer to an unparalleled shopping experience. The Indian retail sector is highly fragmented with 97 per cent of its business being run by the unorganized retailers like the traditional family run stores and corner stores. The organized retail however is at a very nascent stage though attempts are being made to increase its proportion by 9-10 per cent by the year 2010 bringing in a huge opportunity for prospective new players. The sector is the largest source of employment after agriculture, and has deep penetration into rural India generating more than 10 per cent of India‘s GDP.

Liberalization of the Indian economy and rationalisation of business procedures have already ensured a high economic growth with a rapidly expanding base for the manufacturing and high-end services sectors. Fresh avenues for gainful employment to a predominantly young and talented population have created high disposable incomes that translate in to higher consumption and thus better opportunities for all verticals of Retail to flourish. India is the 4th largest economy as regards GDP (in PPP terms) and is expected to rank 3rd by 2010 just behind US and China. On one hand where markets in Asian giants like China are getting saturated, the AT Kearney's 2006 Global Retail Development Index (GRDI), for the second consecutive year Placed India the top retail investment destination among the 30 emerging markets across the world. Over the past few years, the retail sales in India are hovering around 33-35 per cent of GDP as compared to around 20 per cent in the US.

The last few years witnessed immense growth by this sector, the key drivers being changing consumer profile and demographics, increase in the number of international brands available in the Indian market, economic implications of the Government increasing urbanization, credit availability, improvement in the infrastructure, increasing investments in technology and real estate building a world class shopping environment for the consumers. In order to keep pace with the increasing demand, there has been a hectic activity in terms of entry of international labels, expansion plans, and focus on technology, operations and processes. This has lead to more complex relationships involving suppliers, third

party distributors and retailers, which can be dealt with the help of an efficient supply chain. A proper supply chain will help meet the competition head-on, manage stock availability; supplier relations, new value-added services, cost cutting and most importantly reduce the wastage levels in fresh produce. Large Indian players like Reliance, Ambanis, K Rahejas, Bharti AirTel, ITC and many others are making significant investments in this sector leading to emergence of big retailers who can bargain with suppliers to reap economies of scale. Hence, discounting is becoming an accepted practice. Proper infrastructure is a pre-requisite in retailing, which would help to modernize India and facilitate rapid economic growth. This would help in efficient delivery of goods and value-added services to the consumer making a higher contribution to the GDP.

International retailers see India as the last retailing frontier left as the China‘s retail sector is becoming saturated. However, the Indian Government restrictions on the FDI are creating ripples among the international players like Walmart, Tesco and many other retail giants struggling to enter Indian markets. As of now the Government has allowed only 51 per cent FDI in the sector to _one-brand‘ shops like Nike, Reebok etc. However, other international players are taking alternative routes to enter the Indian retail market indirectly via strategic licensing agreement, franchisee agreement and cash and carry wholesale trading (since 100 per cent FDI is allowed in wholesale trading).

The Current Status

India‘s retail industry accounts for 10 percent of its GDP and 8 percent of the employment to reach $17 billion by 2010. The Indian retail market is estimated at US$ 350 billion. But organised retail is estimated at only US$ 8 billion. However, the opportunity is huge-by 2010, organised retail is expected to grow at 6 per cent by 2010 and touch a retail business of $ 17 billion as against its current growth level of 3 per cent which at present is estimated to be $ 6 billion, according to the Study undertaken by The Associated Chambers of Commerce and Industry of India (ASSOCHAM). Indian retailing is clearly at a tipping point. India is currently the ninth largest retail market in the world. And it is names of small towns like Dehradun, Vijayawada, Lucknow and Nasik that will power India up the rankings soon.

Organised retail in India has the potential to add over Rs. 2,000 billion (US$45 billion) business by the Year 2010 generating employment for some 2.5 million people in various retail operations and over 10 million additional work-forces in retail support activities including contract production & processing, supply chain & logistics, retail real estate development & management etc. It is estimated that it will cross the $650-billion mark by 2011, with an already estimated investment of around $421 billion slated for the next four years.

Organized Retail Penetration (ORP) is the highest in footwear with 22 per cent followed by clothing. Though food and grocery account for largest share of retail spend by the consumer at about 76 per cent, only 1 per cent of this market is in the organized sector. However, it has been estimated that this segment would multiply

five times taking the share of the organized market to 30 percent in the coming years.

2.2.1. Slowdown in Indian Economy – myth or reality for retail players?

The current slowdown in the Indian economy notwithstanding, the retail segment in the country seems to be in for a big time expansion led by most major Indian business majors and global players. Even though the CB Richard Ellis report released in April 2008, placed India at a dismal number 44 in the list of preferred destinations for global retailers looking to expand, fresh announcements in the media belie this fact. However, going through these years of learning, nearly all stake holders in the industry are re-considering their retail plans. A need for consolidation in retail business is evident and to give it effect many have hit the drawing boards again – not necessarily means that there is any down turn in the industry. In spite of the fast track growth of the retail industry, India is still undergoing through the initial development phase of modern retail.

2.2.2. Private Consumption & Retail

The country‘s dynamic retail landscape presents a grand opportunity to investors from across the globe, to use India as a strategic business hub. With the changing face of retail, the Indian consumer is in for a rapid transformation. With retail spending growing at double digit, Private Final Consumption Expenditure (PFCE)

at current prices was estimated at Rs. 26,07,584 crores in 2007-08 as against Rs. 23,12,105 crore in 2006-07.

As per the Images F&R Research estimates for India Retail Report the Indian Retail market stood at Rs.1,330,000 crores in 2007 with annual growth of about 10.8 per cent. Of this, the share of organised Retail in 2007 was estimated to be only 5.9 per cent, which was Rs.78,300 crores. But this modern retail segment grew at the rate of 42.4 per cent in 2007, and is expected to maintain a faster growth rate over the next three years, especially in view of the fact that major global players and Indian corporate houses are seen entering the fray in a big way. Even at the going rate, organised retail is expected to touch Rs.2,30,000 crores (at constant prices) by 2010, constituting roughly 13 per cent of the total retail market. The consumer spending is ultimately pushing the economy into a growth-and-liberalisation mode. The Indian market is becoming bolder by the day, and with the economy now expected to maintain its growth at over 8-9 per cent and average salaries being hiked by about 15 per cent, there will be lot more consumption. A short recount of the factors that have earned India the top spot among the favoured retail destinations:

Healthy investment climate

A ‗Vibrant Economy': India topped A T Kearney's list of emerging markets for retail investments for three consecutive years and stood 2nd only behind Vietnam this year. The 2nd fastest growing economy in the world, the 3rd largest economy in terms of GDP in the next 5 years and the 4th largest economy in PPP terms after

USA, China & Japan, India is rated among the top 10 FDI destinations. Barring recent political disturbances, India has been sailing smooth with 2nd stage reforms in place, India can be reasonably proud of having put in place some of the most widely accepted Corporate Ethics (Labour Laws, Child Labour Regulations, Environmental Protection Lobby, Intellectual Property Rights, and Social Responsibility) and major tax reforms including implementation of VAT, all of which make India a perfect destination for business expansion.

In terms of international tourist spending, India is the fastest-growing market in Asia Pacific, according to the Visa Asia Pacific release. The economy has been growing at about 9 per cent a year, which shows that India‘s growth rate can actually exceed that of China by 2015. The Indian economy is expected to grow larger than Britain‘s by 2022 and Japan‘s by 2032, to become the third-largest economy in the world after China and US, and finally become the second largest economy after China by 2050, so the global economic forecasts say.

A report by investment banker Goldman Sachs, credits India with the potential to deliver the fastest growth over the next 50 years with an average rate of more than five per cent a year for the entire period. All these are clear portends in terms of investments and returns. Total FDI (foreign direct investment) inflow in 2007-08, was to the tune of USD25 billion – up 56% over previous year - with investments in infrastructure development and capital market continuing to flow in at a rapid pace.

To sustain an ambitious GDP growth target of nine per cent per annum, India needs to invest around USD 500 billion in development of Infrastructure over the next five years. Of this, about USD150 billion is expected to come from foreign investment. Indian retail industry itself has attracted total investment of over Rs.20,000 crores in creating infrastructure, systems & shop-fit.

At the heart of the Indian growth story is its population, the generators of wealth, both as producers and consumers. With the largest young population in the world - over 890 million people below 45 years of age, India indeed makes a resplendent market. The country has more English speaking people than in the whole of Europe taken together. Its 300 million odd middle class, the ―Real‖ consumers, has attracted the attention of the world, and as the economy grows so does India's middle class. It is estimated that 70 million Indians earn a salary of over USD 19,500 a year, a figure that is set to rise to 140 million by 2011. The number of effective consumers is expected to swell to over 600 million by this time – sufficient to establish India as one of the largest consumer markets of the world.

2.2.3. The retail revolution

In this land of 15 million retailers, most of them owning small mom and pop outlets, we also have a modern retail flourishing like never before. There is little room for conflict as evidenced from the fact that India presents a unique case of consumption-driven economy: while the US reels under recession, where supply clearly outstrips demand, India confronts inflation, where the industry and retailers are as yet unable to provide what the consumer demands.

Over the last few years Indian retail has witnessed rapid transformation in many areas of the business by setting scalable and profitable retail models across categories. Indian consumers are rapidly evolving and accepting modern retail formats. New and indigenised formats such as departmental stores, hypermarkets, supermarkets, specialty and convenience stores, and malls, multiplexes and fun zones are fast dotting the retail landscape.

2.2.4. The Indian retail market segments

The Indian retail market has been gaining strength, riding on the sound vibes generated by a robust economy that has given more disposable incomes in the hand of the consumer who will keep demanding better products and services, and a better shopping environment. In the overall Retail pie, Food and Grocery was the dominant category with 59.5 per cent share, valued at Rs.792,000 crores, followed by Clothing and Accessories with a 9.9 per cent share at Rs.131,300 crores. Interestingly, out-of-home food (catering) services (Rs.71,300 crores) has overtaken Jewellery (Rs.69,400 crores) to become the third largest retail category, with a 5.4 per cent market share – this largely reflects the massive employment opportunities to youngsters in the services sector and accompanying changes in consumer lifestyles. Consumer durables (Rs.57,500 crores) is the fifth largest retail category followed by Health & Pharmaceuticals (Rs.48,800 crores), Entertainment (Rs.45,600 crores), Furniture, Furnishings & Kitchenware (Rs.45,500 crores), Mobiles & Accessories (Rs.27,200 crores), Leisure retail

(Rs.16,400 crores), Footwear (Rs.16,000 crores), Health & Beauty Care services (Rs.4,600 crores) and Watches & Eyewear (Rs.4,400 crores) in the order.

In the Organised retail segment, the picture is different altogether. Clothing & Fashion Accessories is the largest category with 38.1 per cent of the market share, valued at Rs.29,800 crores, followed by Food & Grocery accounting for 11.5 per cent of the organised retail market at Rs.9,000 crores, Footwear with 9.9 per cent of the organised retail market share at Rs.7,750 crores, Consumer Durables with 9.1 per cent market share at the fourth place (Rs.7,100 crores), and Out-of-home food (catering) services and Furniture, Furnishings & Kitchenware retail in the order. The mobile & accessories retail market has shown fastest growth in 2007 (25.6%) over the previous year, the other two prominent categories being out-of-home food (catering) services where growth was 25.1 per cent and books, music & gifts leisure category which achieved 23.3 per cent growth. India‘s biggest USP and asset base is its youthful population, whose appetite for leisure and entertainment is galloping at 14 per cent p.a. With the rapid addition of malls with multiplexes there is a coming together of leisure retail, cinema and gaming. It is indeed difficult to analyze each of these components in isolation. All players are after all trying to get to capture a share of consumer‘s mind – his time and money.

As the consumer‘s spend on leisure and entertainment increases, the mix of his spends is going through a churn like never before. Leisure and entertainment are recession proof. The affluence across the country has touched a large part of the population and there is no looking back. Multiplexes, leisure retailers across books, music, gaming all form a shared existence and whilst the shares of the pie

keep shifting the overall leisure and entertainment business is well on its way to a Rs. 60,000 Crores mark by 2010-11. In the organised retail segment, however, the fastest growth was recorded in the tiny health & beauty care services category (Rs.660 crores), which grew at the rate of 65 per cent in 2007 over the previous year – again a reflection of rise in services sector employment that demands proper grooming. The second fastest growing organised retail category is that of Entertainment (53.8%), followed by the mobile phones & accessories and the food & grocery retail categories, both of which achieved 55.2 per cent growth in 2007. Much of the stupendous growth opportunity in Catering services (25.1%) and leisure retail (23.3%) categories was utilised by the unorganised retailers because organised players could not keep up to the desired growth momentum. A closer study of the retail growth story at constant prices shows that in both these categories growth of organised retail was higher in 2006 (41.7% and 26.1% respectively) as compared to 2007 (37% and 25%). At constant prices, growth in the fashion & accessories retail category, both in the overall market and the organised retail segment, have been consistently positive since 2004: while the overall market grew 12.8 per cent in 2007, the organised segment grew 35.5 per cent. In jewellery retail, the overall market growth was higher in 2007 (9.6%) as compared to the previous year (9.2%) but growth in organised retail was slightly at a slower pace in 2007 (31%) as compared to the previous year. The overall market growth in the time-wear category has declined from 10.7 per cent in 2005 to 9.7 per cent in 2006 and further down to 8.9 per cent in 2007. However, growth in organised retail was higher in 2007 (16.6%) as compared to 2006 (14.8%). Popularity of mobile phones is to a large extent responsible for the dampening of

the overall market growth in this category while the renewed enthusiasm in the organised segment is on account of the fillip from luxury brands and offerings that are positioned more as a high-end lifestyle statement than on the functionality aspect of the product. Footwear retail, the overall market as well as its organised segment, has grown faster year after year but growth in 2007 was especially remarkable: the overall market grew 12 per cent in 2007 as against a 9.2 per cent growth in 2006 while the organised segment grew 42.3 per cent and 36.4 per cent respectively for the two years. The global brands have actually turned the heat on, and the domestic brands too appear to have accepted the challenge in the true spirit.

Growth in the health and beauty care category has been remarkable in 2007, though the organised segment growth in 2007 (57.5%) was slightly lower as compared to 2006 (59.1%). The demand is stupendous but organised players have hardly much to boast of in terms of innovative concepts and global standards when it comes to providing the customers with an experience that is superior and radically different from what the unorganised segment offers. This category needs to be positioned as a ―wellness‖ category that provides individualised services to customers with synergies of health & beauty care, pharmaceuticals and specialised clinical services – all at one place. Another category that merits special mention is Furnishings and Furniture retail, where the overall market grew at seven per cent in 2007 as compared to just 3.2 per cent in 2006 – thanks to the housing sector boom. The organised segment also grew faster at 29.7 per cent in 2007 as compared to 23.1 per cent the previous year, but this Rs.45,500 crores category

calls for better attention from organised players**.** Is India ready for ready-to-assemble furniture? May be not, but surely the market will change in next couple of years. Global players need to understand that Indian homes are different and so are the Indian environments, maintenance standards. At present most large players entering this segment are busy experimenting and in the process have lost monies too. Consumer durables and the mobile phone & accessories categories have both grown faster in 2007 as compared to 2006. At constant prices, the overall food & grocery retail market grew slightly higher at 2.3 per cent in 2007 as compared to a 2.2 per cent annual growth in the previous two years. But the organised retail segment in this category is simmering in the true sense – a 50 per cent growth in 2007 as compared to 42.9 per cent in 2006, and lot more fireworks can be expected this year and the years ahead. Valued at Rs.9,000 crores, this organised market constitutes barely 1.1 per cent of the total food & grocery retail market. Time-wear (48.9%) and Footwear (48.4%) are the most organised of all retail categories. Clothing & fashion accessories retail comes next with the organised segment controlling 22.7 per cent of the market.

Segment analysis

The structure of Indian retail is developing rapidly with shopping malls becoming increasingly common in the large cities and development plans being projected at 150 new shopping malls by 2008. However, the traditional formats like hawkers, grocers and tobacconist shops continue to co-exist with the modern formats of retailing. Modern retailing has helped the companies to increase the consumption of their products, for example: Indian consumers would normally consume the rice

sold at the nearby kiranas viz. Kolam for daily use. With the introduction of organized retail, it has been noticed that the sale of Basmati rice has gone up by four times than it was a few years back; as a superior quality rice (Basmati) is now available at almost the same price as the normal rice at a local kirana. Thus, the way a product is displayed and promoted influences its sales. If the consumption continues to grow this way it can be said that the local market would go through a metamorphoses of a change and the local stores would soon become the things of the past or restricted to last minute unplanned buying.

Food and grocery retail

The food business in India is largely unorganized adding up to barely Rs.400 billion, with other large players adding another 50 per cent to that. The All India food consumption is close to Rs.9,000 billion, with the total urban consumption being around Rs.3,300 billion. This means that aggregate revenues of large food players is currently only 5 per cent of the total Indian market, and around 15-20 per cent of total urban food consumption. Most food is sold in the local ‗wet' market, vendors, roadside push cart sellers or tiny kirana stores. According to McKinsey report, the share of an Indian household's spending on food is one of the highest in the world, with 48 per cent of income being spent on food and beverages.

Apparel retail

The ready-mades and western outfits are growing at 40-45 per cent annually, as the market teams up with international brands and new entrants entering this

segment creating an Rs.5 billion market for the premium grooming segment. The past few years has seen the sector aligning itself with global trends with retailing companies like Shoppers‘ stop and Crossroads entering the fray to entice the middle class. However, it is estimated that this segment would grow to Rs. 3 billion in the next three years.

Gems and Jewellery retail

The gems and jewellery market is the key emerging area, accounting for a high proportion of retail spends. India is the largest consumer of gold in the world with an estimated annual consumption of 1000 tonnes, considering actual imports and recycled gold. The market for jewellery is estimated as upwards of Rs. 650 billion.

Pharmaceutical retail

The pharma retailing is estimated at about Rs. 300 billion, with 15 per cent of the 51 lakh retail stores in India being chemists. Pharma retailing will follow the trend of becoming more organised and corporatized as is seen in other retailing formats (food, apparel etc). A few corporates who have already forayed into this segment include Dr Morepen (with Lifespring and soon to be launched Tango), Medicine Shoppe, Apollo pharmacies, 98.4 from Global Healthline Pvt Ltd, and the recently launched CRS Health from SAK Industries. In the south, RPG group‘s Health &

Glow is already in this category, though it is not a pure play pharma retailer but more in the health and beauty care business.

Music Retail

The size of the Indian music industry, as per this Images-KSA Study, is estimated at Rs.11 billion of which about 36 percent is consumed by the pirated market and organized music retailing constitutes about 14 percent, equivalent to Rs.1.5 billion.

Book retail

The book industry is estimated at over Rs. 30 billion out of which, organized retail accounts for only 7 per cent (at Rs.2.10 billion). This segment is seen to be emerging with text and curriculum books accounting to about 50 per cent of the total sales. The gifting habit in India is catching on fast with books enjoying a significant share, thus expecting this sector to grow by 15 per cent annually.

Consumer durables retail

The consumer durables market can be stratified into consumer electronics comprising of TV sets, audio systems, VCD players and others; and appliances like washing machines, microwave ovens, air conditioners (A/Cs). The existing size of this sector stands at an estimated US$ 4.5 Billion with organized retailing being at 5 per cent**.**

2.2.5. Space – and the freedom to grow

Each time one takes stock of the country‘s retail estate scenario, one invariably comes across statistics to show that every city in the country is bursting at its seams with shopping centre activity. If mall space were to be taken as an indication of the level of activity, we find that the country has witnessed nearly

12-fold growth in the last five years, with total mall space having increased from just about 3.7 million square feet in 2002 to over 47 million square feet in 2007. Also, the opening up of the real estate sector to FDI has brought India in the international investment spotlight. FDI inflow in to the sector has propelled the realty sector growth at over 30 per cent per annum. There is yet a lot more to unfold on India‘s retail landscape in the years ahead.

2.2.6 Shopping Centers

Currently, there are about 280 operational shopping centres in various formats and sizes (including some partly operational), and this number is expected to rise to almost 500 by end-2010. Of the new malls coming up, 40 per cent are concentrated in the smaller cities. Shopping centre business alone is estimated to become a Rs 40,000 crores business by 2010-11.

The emergence of Shopping Centers is already beginning to define a new lifestyle for India. There is no doubt a huge demand for clean, contemporary shopping and entertainment complexes that will house India‘s brands and retail formats and offer New India an exciting and rewarding shopping experience for the whole family. A number of winning solutions will doubtless emerge over the next few years but the dominant centers for the long term will be those that are designed around the Indian consumer and cater to the long term specific needs of a particular location. A shopping center doesn‘t serve all India. It serves consumers living largely within a five to fifteen kilometres radius of that center. So a successful shopping center in Trivandrum will be designed differently to one in

Ludhiana. The tenant mix will be different. The food court will have a different menu offer and local services such as transport and logistics will be tailored to the needs of the local community.

Organised retailing in small-town India is already growing at over 50-60 per cent a year, compared to 35-40 per cent growth in the large cities. About 200 tier-III cities with population of less than 2 million and another 500 rural towns have the potential to be the hub for rural markets, where organised retailing can effectively set base – each of these 700 centres will on an average be catering to about 1,000 villages.

2.2.7. Supply Chain, Logistics & Infrastructure

Organised retail is a function of strong supply chain and robust physical infrastructure. Basic supply chain framework takes care of operational performance at each nodal point – from order to delivery. In view of this, major retailers will have to continuously upgrade their back-end, front-end and supply chain dynamics in order to provide a standard of value and services to their customers.

Corporate bigwigs such as Reliance, AV Birla, Tata, Godrej, Bharti, Mahindra, ITC, RPG, Pantaloon, Raheja and Wadia Group are expected to invest close to Rs.1 trillion in the business of retail over the next five years. Reliance Retail is investing Rs.30,000 crores in setting up multiple retail formats backed by a 68-strong distribution network, with expected sales of over Rs.100,000 crore by 2010.

The Future Group‘s Pantaloon Retail and RPG‘s Spencer‘s are also going all out to maintain their dominant position on India‘s retail horizon. Subhiksha has earned global accolades for its fast-track growth. The Lifestyle India, Indiabulls, Wadhawan Group, Vishal Retail, petroleum majors IOCL, BPCL & HPCL, and others are firming up more and more ambitious retail expansion plans by the day. While global retailers Metro AG and Shoprite Holdings increase their presence on the Indian retail landscape, the Bharti – Wal-mart combine is scouting locations for their joint retail venture. The recent tie-up between Tata and Tesco further adds to the action in retail.

2.2.8. Regional Retailers

Realising the big picture of retail regional retailers too are waking up to position themselves strong against MNCs and Indian Corporate big wigs jumping into the ring. But the match has only just begun and promises to be a show stealer. Some of the larger regional players are looking to tie up with international retailers. Although multinational and large Indian retailers clearly have the advantage of size and power, local companies have survived by adapting. As the big guns introduce new branding strategies, improved merchandising and management techniques, local retailers are fast to catch up and emulate to co-exist.

The big players in what has traditionally been a fragmented market makes existing food retailers look really tiny, even in the case of large regional chains that have

been around for years. The South particularly has a large concentration of such businesses which are giving the MNC / Corporate Retailers a run for their money.

Even as multinational retailers are firming up their India strategies, franchising is emerging as the preferred option. Franchisee activity is expected to pick up in tier-II cities as well. According to a Frost & Sullivan research, the overall Indian third-party logistics (3PL) market, estimated at about USD 890.3 billion in 2005, is expected to grow at a compound annual growth rate of 21.9 per cent to reach USD 3,556.7 million in 2012. Shop-fit and Technology are the other sunrise retail support sectors that offer immense opportunity for investment and growth.

2.2.9. Retail Formats

There is strong emergence of India specific retail formats irrespective of the size. For example, hypermarkets, supermarkets or convenience stores that are emerging in India today are specifically designed for the Indian consumer. A store in India will have non-vegetarian section sealed-off from the rest of the store out of respect for a group of consumers. Spices, vegetables and grains are seen stacked high in a special section. Retailers in India have realised he relevance of designing and executing world-class hypermarket formats in India, specifically catering to the Indian consumer yet offering world-class products at prices that India can afford.

Indian department stores stock brands that Indian consumers want and design their store layouts based on their few years of experience. But this experience is unique to them alone. A new entrant has all those lessons to learn. The existing players are developing a sophisticated knowledge base on consumption patterns and

preferences that are a critical tool in defining competitive positioning going forward.

Just the right time to think retail

Developments indicate that this is just the right time to think retail. Fuel and passenger vehicles are two of the mega businesses that can tremendously gain in the evolving scenario. The auto sector needs to explore innovative collaborative opportunities with the retail sector to add value to the shopping experience of passenger vehicles. So far operated through dealership network with showrooms mostly in not-so-happening premises, auto showrooms are now beginning to move to retail centres to grab attention of new generation upwardly mobile customers. With increasing income, easy credit facilities and 'change every year' new found attitude (initially started with mobile handsets) Indian consumers are likely to make spontaneous decisions on automobile buys as well. Oil companies - unable to raise prices of transportation fuels in line with rising global crude oil prices - are now looking at alternate revenue streams; a major reason why added emphasis is being placed on forecourt retailing.

Hypermarkets that have ventured into the retail petroleum business have met with considerable success due to competitive fuel pricing, discounted prices linked to loyalty programmes and cross-merchandising. India's oil majors can certainly take the lead to fuel the retail growth collaborating with real estate developers, auto companies, consumer brands, retailers and service providers. This will also facilitate travel & tourism in no small measure. This has happened across the globe and is now happening in India.

2.3 Shopping Tourism

Travel and tourism are two other sectors that will immensely benefit connecting with retail.

Shopping is becoming an increasingly relevant component of the tourism value chain. Shopping has converted into a determinant factor affecting destination choice, an important component of the overall travel experience and, in some cases the prime travel motivation. Destinations have thus an immense opportunity to leverage this new market trend by developing authentic and unique shopping experiences that add value to their touristic offer while reinforcing, and even, defining their tourism brand and positioning. More importantly, shopping is one of the major categories of tourists' expenditure, representing a significant source of income for national economies both directly and through the many linkages to other sectors in the economy.

Few publications have been written specifically on the contribution of shopping tourism to tourism overall and its impact on destinations. Over the past six decades tourism has experienced continued expansion and diversification, becoming one of the largest and fastest-growing economic sectors in the world. Many new destinations have emerged, challenging the traditional ones of Europe and North America. Despite occasional shocks, international tourist arrivals have shown virtually uninterrupted growth – from 277 million in 1980 to 528 million in 1995, and passing the 1 billion mark in December 2012. As global economic recovery took hold and departures from the emerging economies continued to show strong growth, this figure rose to 1.087 billion international arrivals in 2013.

UNWTO's long-term outlook and assessment of future tourism trends is positive. The number of international tourist arrivals worldwide is expected to increase by 3.3% a year on average from 2010 to 2030. This represents some 43 million more international tourist arrivals every year, reaching a total of 1.8 billion arrivals by 2030.

As an internationally traded service, inbound tourism has become one of the world's major trade categories. The overall export income generated by inbound tourism, including passenger transport, exceeded US$ 1.3 trillion in 2012, accounting for as much as 9% of global GDP when we factor in its direct, indirect and induced impact. Globally, the sector provides 1 in every 11 jobs and for many developing countries it is one of the main sources of foreign exchange income, creating much needed employment and opportunities for development.

Defining Shopping Tourism

Given the relatively recent evolution of shopping tourism as a primary motivation for travel, few solid definitions of the concept exist. As early as 1991 Jansen-Verbeke[1] questioned how, where and when a shopping environment can function as a tourist attraction. She dismissed the notion that a shopping destination must be a 'shopping paradise', some kind of border area with a specific fiscal regime to stimulate the attractiveness of visitors for shopping. Instead, she proposed the concept of a shopping experience as the cornerstone of its own branch of tourism, including the traditional shopping areas of towns and cities, shopping centres located in peri-urban areas and centres that have been transformed from their previous uses as ports, industrial/agricultural hubs, or theme parks. In 2004, Moscardo[2] suggested that until 2004, '[i]f, traditionally, the consumption of tourists focused on specific goods and services (hotels, restaurants, cultural or entertainment offers), modern tourists, who often enjoy a high purchasing power, nowadays are consumers of wider goods, such as fashion, crafts or design'.[3] When attempting to define shopping tourism, researchers have offered various answers, focusing on different elements of the experience. Stansfield[4] for example, has observed that an individual's shopping behavior is different when on holiday. Travelling away from home impacts tourists' shopping habits. Spending increases, more non-essential items are bought, and purchases are made on unusual days

[1] In: Tourism Management: Research - Policies - Practice. Vol. 12 (1991), No. 1. – pp. 9-14.

[2] - 3 Moscardo, G., 2004, Shopping as a destination attraction: an empirical examination of the role of shopping in tourists' destination choice and experience". Journal of Vacation Marketing, 10 (4), pp.294-307.

[4] In: Butler, R.W., 1991, West Edmonton Mall as a tourist attraction, Canadian Geographer, 35, pp.287–295

(e.g., Sundays, evenings and/or holidays). Various authors have observed that when a domestic and foreign customer are offered the same retail environment, it can be anticipated that the foreign customer will purchase more items, while also spending more per item, and at a different time or day than the domestic customer. By straying from the conventional norms of consumption, tourists demonstrate shopping as more of a leisurely activity than their average purchases at home. On the basis of these observations, one approach may be to define shopping tourism as a contemporary form of tourism fostered by individuals for whom purchasing goods outside of their usual environment is a determining factor in their decision to travel. Leisure has always served as one of the primary motivations for travelling, as tourists seek enjoyable activities in places outside of their homes. However, it was not until recently that shopping was considered a leisurely activity, not to mention one that could be a driving force behind tourism. As Dallen Timothy has observed, 'consumption is not just about products. It is about consuming places, spaces and time'[5]. Hence the growing space that shopping occupies in the destination management agenda.

Opportunity in Indian Retail

Favourable demographic and psychographic changes relating to India's consumer class, international exposure, availability of quality retail space, wider availability of products and brand communication are some of the factors that are driving the retail in India. Development of India as a sourcing hub shall further make India as an attractive retail opportunity for the global retailers. Retailers like Wal-Mart, GAP, Tesco, JC Penney, H&M, Karstadt-Quelle, Sears (Kmart), etc stepping up their sourcing requirements from India and moving from third-party buying offices to establishing their own wholly owned / wholly managed sourcing & buying offices shall further make India an attractive retail opportunity for the global players.

[5] Timothy, D., 2004, Shopping Tourism, Retailing and Leisure, Aspects of Tourism vol.23 p.11

Though lucrative opportunities exist across product categories, food and grocery, nevertheless, presents the most significant potential in the Indian context as consumer spending is highest on food. Further, ‗wet groceries' i.e. fresh fruits and vegetables is the most promising segment within food and grocery though initially all retailers foraying in to this segment had to face had wide spread protest from traders, small shop keepers.

The next level of opportunities in terms of product retail expansion lies in categories such as apparel, jewellery and accessories, consumer durables, catering services and home improvement. These sectors have already witnessed the emergence of organized formats though more players are expected to join the bandwagon. Some of the niche categories, like Leisure and entertainment (Books, Music and Gifts in particular), offer interesting opportunities for the retail players. Currently the fashion sector in India commands a lion's share in the organised retail pie. This is in line with the retail evolution in other parts of the world, where fashion led the retail development in the early stages of evolution and was followed by other categories like Food & Grocery, Durables etc. Fashion across lifestyle categories makes up for over 50 per cent of organised retail and with the

kind of retail space growth that India is witnessing we can certainly foresee a very healthy prospect for the fashion industry.

Investment Opportunities in the Retail Sector

AT Kearney‘s study on global retailing trends found that India is the least competitive as well as least saturated of all major global markets. This implies that there are significantly low entry barriers for players trying to setup base in India, in terms of the competitive landscape. The report further stated that global retailers such as Walmart, Carrefour, Tesco and Casino would take advantage of the more favourable FDI rules that are likely in India and enter the country through partnerships with local retailers. Other retailers such as Marks & Spencer and the Benetton Group, who operate through a franchisee model, would most likely switch to a hybrid ownership structure.

A good talent pool, unlimited opportunities, huge markets and availability of quality raw materials at cheaper costs is expected to make India overtake the world's best retail economies by 2042, according to industry players. The retail industry in India, according to experts, will be a major employment generator in the future. Currently, the market share of organised modern retail is just over 4 per cent of the total retail industry, thereby leaving a huge untapped opportunity.

The Potential of the Indian Retail Sector

The high growth projected in domestic retail demand will be fuelled by:

- The migration of population to higher income segments with increasing per capita incomes

- An increase in urbanisation
- Changing consumer attitudes especially the increasing use of credit cards
- The growth of the population in the 20 to 49 years age band
- There is retail opportunity in most product categories and for all types of formats
- Food and Grocery: The largest category; largely unorganised today
- Home Improvement and Consumer Durables: Over 20 per cent p.a. CAGR estimated in the next 10 years
- Apparel and Eating Out: 13 per cent p.a. CAGR projected over 10 years

Opportunities for investment in supply chain infrastructure: Cold chain and logistics. India also has significant potential to emerge as a sourcing base for a wide variety of goods for international retail companies

- Many international retailers including Wal-Mart, GAP, JC Penney etc. are already procuring from India.

The sector is expected to see an investment of over $30 billion within the next 4-5 years, catapulting modern retail in the country to $175-200 billion by 2016, according to Technopak estimates.

Of the total organised retail market of Rs 550 billion, the business of fashion accounts for Rs 300.80 billion, which translates into nearly 55 per cent of the organised retail segment in the country. Total fashion sector was estimated at Rs 1,914 billion and forms about 15 per cent of the country's retail market of Rs

12,000 billion. Commanding such a large chunk of the organised retail business in India, fashion retailing has indeed been responsible for single-handedly driving the business of retail in India.

Challenges in Retailing

The industry is facing a severe shortage of talented professionals, especially at the middle-management level. Most Indian retail players are under serious pressure to make their supply chains more efficient in order to deliver the levels of quality and service that consumers are demanding. Long intermediation chains would increase the costs by 15 per cent. Lack of adequate infrastructure with respect to roads, electricity, cold chains and ports has further led to the impediment of a pan-India network of suppliers. Due to these constraints, retail chains have to resort to multiple vendors for their requirements, thereby, raising costs and prices. The available talent pool does not back retail sector as the sector has only recently emerged from its nascent phase. Further, retailing is yet to become a preferred career option for most of India‘s educated class that has chosen sectors like IT, BPO and financial services.

Even though the Government is attempting to implement a uniform value-added tax across states, the system is currently plagued with differential tax rates for various states leading to increased costs and complexities in establishing an effective distribution network. Stringent labor laws govern the number of hours worked and minimum wages to be paid leading to limited flexibility of operations and employment of part-time employees. Further, multiple clearances are required

by the same company for opening new outlets adding to the costs incurred and time taken to expand presence in the country. The retail sector does not have _industry' status yet making it difficult for retailers to raise finance from banks to fund their expansion plans. Government restrictions on the FDI are leading to an absence of foreign players resulting into limited exposure to best practices. Non-availability of Government land and zonal restrictions has made it difficult to find a good real estate in terms of location and size. Also lack of clear ownership titles and high stamp duty has resulted in disorganized nature of transactions.

Emerging trends in Retailing in India

Sourcing **-** The CPG industry is following the IT outsourcing trend. Indian subsidiaries of global CPG players have proved themselves in terms of quality and production capabilities. This has led several international companies to source from India. For example, Hindustan Lever the subsidiary of Unilever exports a wide range of products like soaps, detergents, oral care and skin care products to other Unilever subsidiaries. A new Export Orientated Unit is being set up by Hindustan Lever in Pune, to cater to the needs of this business. Unilever is also setting up a global sourcing office in India to buy products and raw materials from low-cost locations for its subsidiaries across the world. India is moving towards embracing global patent and trademark standards that will facilitate outsourcing by CPG companies. With quotas in textiles sector also being relaxed, global retailers are increasingly focusing their sourcing efforts (both apparel and non-apparel) from India. Wal-Mart has announced it will increase its annual sourcing from

India from current USD0.5 billion to USD5 billion. This trend is likely to be followed by most big retailers.

Forward integration / Alternate channels - In a bid to close the distance between the company and the end consumer by cutting down the distribution channel, some manufacturers of consumer goods are establishing company-owned stores and service formats and exploring the store-in-store concept to enhance margins and increase customer value.

Rapid expansion and format migration - After making years of investments in customer acquisition, setting up of systems / processes and consequent operational losses, many leading retailers have passed their ―learning‖ phase and are getting into the consolidation/aggressive rollout phase. Today, few of them are making modest money out of the business. This provides confidence to the investors to infuse much needed capital in the business, and will lead to further expansion, innovation and format migration. For example the department store chain, Shoppers Stop, will soon have its Initial Public Offering (IPO) and use all the funds raised, for rapid expansion in existing format and roll-out of grocery stores. Other leading retailers such as Trent (Westside) and Landmark Group (LifeStyle) are also considering new formats in home improvement and hypermarkets.

Region specific formats - With organized retail penetrating into Class II towns, retailers have started differentiating and experimenting with store sizes and formats. For example, in the departmental store format, while most Class A cities

and metros have large stores of 50,000+ square foot in size, stores in Class II towns have stabilized in the 25,000-35,000 square foot range.

Development of discount formats **-** Large discount formats, or hypermarkets, aiming at retail consolidation by providing a single point of contact between brand-owners and customers, are now emerging as major competitors to both unorganized and organized retailers. Penetration of organized retailing into the lower income brackets and consumer demand for increased value-for-money has improved the prospects of these formats. Big Bazaar, promoted by Pantaloon, and Giant, promoted by the RPG group, provide two examples of this trend.

Convenience stores at gas stations **-** India is now showing signs of aligning with global trends in petro-retailing with increasing sales coming from non-fuel related products. With deregulation, private players entering this sector force existing petro-retailers to review their business models. Dealer and company owned convenience stores at service stations are on the rise. State run oil giants have entered into joint ventures with FMCG companies and food retailers to sell food and groceries at select markets.

Malls potential **-** Over 300 large malls will come up by the end of 2016. The urban Indian consumer prefers _All-under-one-roof‟ shopping destination together with eating-out and entertainment venues. This creates excellent development potential for malls. With the shortage of retail space, this makes an excellent potential for development of malls. As there is a shortage of retail space in central locations, most of these malls are being set up in the suburbs of large metropolitan

cities. In addition, foreign retailers are restricted from investing in property and thus have to depend on malls for their outlets.

2.4 Retail formats in India

This section begins by analyzing the retail formats in the present Indian scenario and proceeds to outline the key strategic factors in retailing. In the last part the paper shows the challenges facing retail and our recommendations for making organized retailing a success. Each of the retail stars has identified and settled into a feasible and sustainable business model of its own.

Modern retailing has entered India in form of sprawling malls and huge complexes offering shopping, entertainment, leisure to the consumer as the retailers experiment with a variety of formats, from discount stores to supermarkets to hypermarkets to specialty chains. However, *kiranas* still continue to score over modern formats primarily due to the convenience factor. The organized segment typically comprises of a large number of retailers, greater enforcement of taxation mechanisms and better labour law monitoring system. It's no longer about just stocking and selling but about efficient supply chain management, developing vendor relationship quality customer service, efficient merchandising and timely promotional campaigns. The modern retail formats are encouraging development of well-established and efficient supply chains in each segment ensuring efficient movement of goods from farms to kitchens, which will result in huge savings for the farmers as well as for the nation. The Government also stands to gain through more efficient collection of tax revenues. Along with the modern retail formats,

the non-store retailing channels are also witnessing action with HLL initiating Sangam Direct, a direct to home service. Network marketing has been growing quite fast and has a few large players today. Gas stations are seeing action in the form of convenience stores, ATMs, food courts and pharmacies appearing in many outlets.

In the coming years it can be said that the hypermarket route will emerge as the most preferred format for international retailers stepping into the country. At present, there are 50 hypermarkets operated by four to five large retailers spread across 67 cities catering to a population of half-a-million or more. Estimates indicate that this sector will have the potential to absorb many more hypermarkets in the next four to five years.

Traditionally, the small store (kirana) retailing has been one of the easiest ways to generate self-employment, as it requires minimum investments in terms of land, labour and capital. These stores are not affected by the modern retailing as it is still considered very convenient to shop. In order to keep pace with the modern formats, kiranas have now started providing more value-added services like stocking ready to cook vegetables and other fresh produce. They also provide services like credit, phone service, home delivery etc. The organized retailing has helped in promoting several niche categories such as packaged fruit juices, hair creams, fabric bleaches, shower gels, depilatory products and convenience and health foods, which are generally not found in the local kirana stores. Looking at the vast opportunity in this sector, big players like Reliance and K Rahejas has

announced its plans to become the country's largest modern retainers by establishing a chain of stores across all major cities.

Apart from metro cities, several small towns like Nagpur, Nasik, Ahmedabad, Aurangabad, Sholapur, Kolhapur and Amravati as witnessing the expansion of modern retails. Small towns in Maharashtra are emerging as retail hubs for large chain stores like Pantaloon Retail because many small cities like Nagpur have a student population, lower real estate costs, fewer power cuts and lower levels of attrition. However, retailers need to adjust their product mix for smaller cities, as they tend to be more conservative than the metros. In order for the market to grow in modern retail, it is necessary that steps are taken for rewriting laws, restructuring the tax regime, accessing and developing new skills and investing significantly in India.

Conventional Formats

- **Kiranas -** These are food and non-food neighborhood counter stores, also called _mom and pop stores‘ in western countries. These are big chunks forming the segregated and unorganised retail segment. These are family-owned and- run retail-outlets picking the goods from wholesalers totalling to around 12 million stores across India.

- **Mandis -** These are the largest chunk of unorganised retail catering to urban and rural masses. Mandis are physically located at different regions to enhance convenient shopping. The sellers bring across various products like eatables, vegetables and fruits, pulses, cereals, spices etc. The most

prominent of them are sabzi mandis found in most of the localities across India.

- **Village Haats -** This form is operating in rural areas where buyers and sellers gather once in a week or month from nearby villages and small towns to cater their livelihood and leisure needs. These haats are a source of entertainment and socialization among rural masses.

- **Push Cart Vendors -** The are categories of vendors roaming from door to door in various localities selling fruits, vegetables, and other eatables, from which mostly housewives makes purchases that too on credit.

Table 2.1 Table of Conjoint Context

	Attribute		
	Level 1	Level 2	Level 3
R1	-0.50	0.20	0.30
R2	-0.20	0.05	0.15
R3	-0.50	0.23	0.27

Table 2.2 Upcoming Retail Formats

Modern Formats	**Area (sq. ft)**	**Points of Differentiation**
Shopping Malls	60,000-7,00,000	Multi-format, multiproduct, multi-brands, catering lifestyle needs
Hypermarts	50,000-70,000	Multi-verticals
Supermarts	5,000-10,000	Single vertical
Departmental Stores	20,000-50,000	Single Vertical
Apparel Stores	20,000-25,000	Multi-branded, single vertical on specific needs of customers
Exclusive formats	500-5,000	Owned/Franchised single product

Table 2.3 Advantages of Conventional and Modern Organised Retail Formats

Conventional	Modern Organised
Low operating-cost and overheads	Large bargaining power
Proximity to consumers	Range and variety of goods
Long operating-hours	Quality assurance (brand related, durability)
Strong relations with customers	Convenience and hygiene

In the long run, both traditional and modern retail formats will co-exist providing the benefits of both the formats. Organised retail will dominate the traditional formats, which will fall under the new evolving hub-and-spoke, and cash-&-carry models.

Cash-&-carry Wholesale Model

Cash-&-carry is a form of retail trade in which goods are sold from a wholesale warehouse operated either on a self-service basis where customers settle the invoice on-the-spot or pay cash and carry the goods away themselves. The cash-&-carry player also performs many value-added functions, including selling and promoting, buying and assortment building, bulk-breaking, warehousing, transporting, financing, risk-bearing, supplying market information, and providing management services.

Hub-and-spoke Model

Retail Chains are entering residential areas with the hub-and-spoke model, whereby one large store supports various smaller stores in the nearby residential areas. This is win-win model is well-suited to the Indian business scene where large stores obtain supplies from the warehouse and supplies to the consumers, involving both large payers acting as wholesalers and local kiranas as retail

outlets. With efficient supply chain management, availability of space and proper technology in place, this will not take much time. The Piramyd Retail‘s Trumart Stores (food and grocery) in Mumbai and Pune are based on a similar model.

SWOT Analysis

Strength

- Retailing is a ―technology-intensive" industry. It is technology that will help the organised retailers to score over the unorganised retailers.

  Successful organised retailers today work closely with their vendors to predict consumer demand, shorten lead times, reduce inventory holding and ultimately save cost. Example: Wal-Mart pioneered the concept of building competitive advantage through distribution & information systems in the retailing industry. They introduced two innovative logistics techniques – cross-docking and EDI (electronic data interchange).

- On an average a super market stocks up to 5000 SKU's (Stock keeping Units) against a few hundreds stocked with an average unorganised-retailer

Weaknesses

- **Less Conversion level**: Despite high footfalls, the conversion ratio has been very low in the retail outlets in a mall as compared to the standalone

  counter parts. It is seen that actual conversions of footfall into sales for a mall outlet is approximately 20-25 percent. On the other hand, a high street store of retail chain has an average conversion of about 50-60 percent. As a result, a stand-alone store has a ROI (return on investment) of 25-30

percent; in contrast the retail majors are experiencing a ROI of 8-10 percent.

- **Customer Loyalty**: Retail chains are yet to settle down with the proper merchandise mix for the mall outlets. Since the stand-alone outlets were established long time back, so they have stabilized in terms of footfalls & merchandise mix and thus have a higher customer loyalty base.

Opportunity

- The Indian middle class is already 30 crores & is projected to grow to over 60 crores by 2010, making India one of the largest consumer markets of the world. The **IMAGES-KSA** projections indicate that by 2015 India will have over **55 crores people under the age of 20** – reflecting the enormous opportunities possible in the **kids and teens retailing segment.**

- Organised retail is only 4 percent of the total retailing market in India. It is estimated to grow at the rate of 27 percent p.a. and reach Rs. 1,37,000 crores by 2010**.**

- **Percolating down**: In India it has been found out that the top 6 cities contribute 66 percent of the total organised retailing. While the metros have already been exploited, the focus has now been shifted towards the tier-II cities. The 'retail boom' of which 85 percent has so far been concentrated in the metros, is beginning to percolate down to these smaller cities and towns. The contribution of these tier-II cities to total organised retailing sales is expected to grow to 20-25 percent**.**

- **Rural Retailing:** India's huge rural population has caught the eye of the retailers looking for new areas of growth. ITC launched India's first rural mall "Chaupal Saga" offering a diverse range of products from FMCG to electronic goods to automobiles, attempting to provide farmers a one-stop destination for all their needs. "Hariyali Bazar", started by DCM Sriram group, provides farm-related inputs & services. The Godrej group has launched the concept of 'agri-stores' named "Aadhaar" which offers agricultural products such as fertilizers & animal feed along with the required knowledge for effective use of the same to the farmers. Pepsi on the other hand is experimenting with the farmers of Punjab for growing the right quality of tomato for its tomato purees, pastes.

Threats

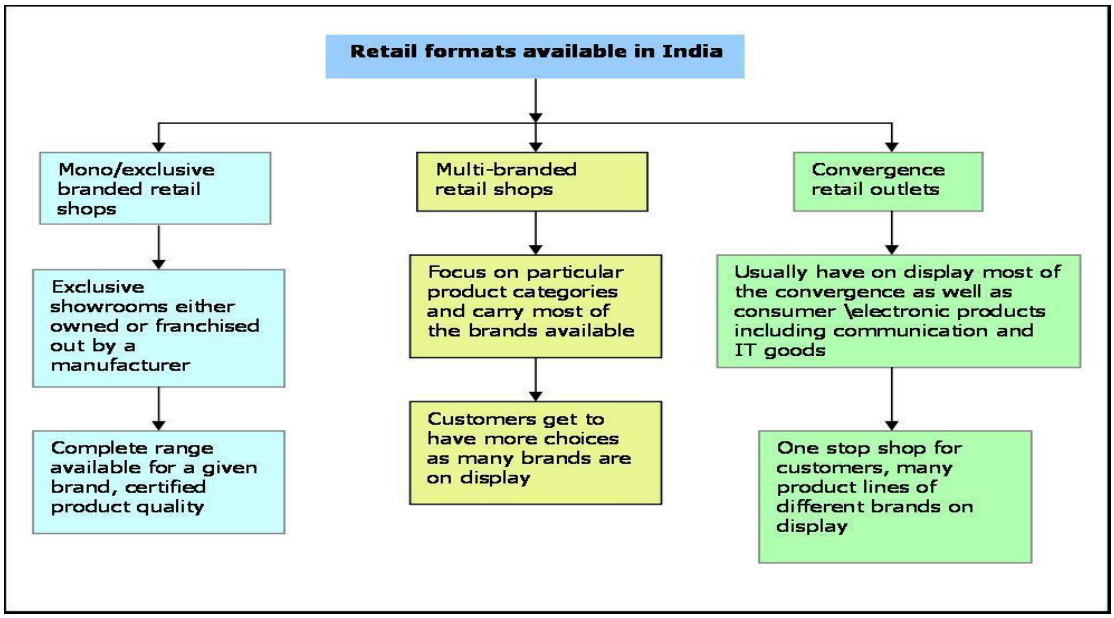
- If the unorganised retailers are put together, they are parallel to a large supermarket with little or no over-heads, a high degree of flexibility in merchandise, display, prices and turnover.

- **Shopping Culture** has not developed in India as yet. Even now malls are just a place to hang around, largely confined to window-shopping.

Retail Formats available in India – Based on Brand and Convenience classification

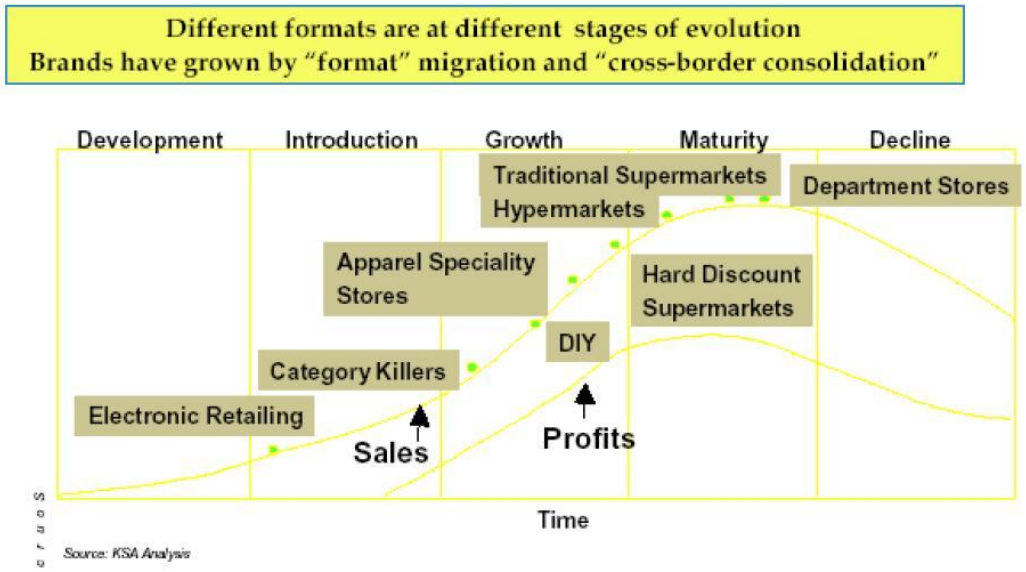
The diagram shows the retail formats classification bases on Brand and Convenience

Figure 2.3 Retail format



Growth Pattern of retail format in India

Figure 2.4 Retail life cycle



Structure of Retail Formats

Structure of the retailing industry according to **ownership patterns**:

• An unaffiliated or independent retailer

• A chain retailer or corporate retail chain

• A franchise system

• A Leased Department (LD)

• Vertical Marketing System (VMS)

• Consumer Co-operatives

A new entrant in the retail environment is the **'discounter' format**. It is also is known as **cash-and-carry or hypermarket**. These formats usually work on bulk buying and bulk selling. Shopping experience in terms of ambience or the service

is not the mainstay here. RPG group has set up the first 'discounter' in Hyderabad called the Giant. Now Pantaloon is following suit. Two categories of customers visit these retail outlets. 1. The small retailer. For example, a customer of Giant could be a dhabawala who needs to buy edible oil in bulk. 2. The regular consumer who spends on big volumes (large pack sizes) because of a price advantage per unit.

2.4.1 Key Strategic Factors in Retailing

The key to success is identifying a superior value-promise and who is in a better position to do it than retailers? Retailers are the closest to the point of purchase and have access to a wealth of information on consumer shopping behaviour. Retailers have some unique advantages for managing brands such as continuous and actionable dialogue with consumers, control over brand presentation at point-of-sale, control over shopping environment, display location/adjacencies, and signage. And they have used this advantage with tremendous success.

As seen, the role of the intermediary is being diminished gradually, which has obvious implication of backlash of the trade channel upwards towards the suppliers. This is more severe in countries such as India, where the channel economics in favour of the middlemen is still strong enough given the fragmentation of the retail sector. Therefore when FoodWorld, the largest grocer in India has a ―direct supply‖ contract with over 20% of its key suppliers, it gives rise to conflict of interest with the distribution infrastructure that suppliers have painstakingly built over the years. Thus companies like HLL have evolved a

distinct distribution channel altogether (called ―Modern Trade‖) to service the needs of such large grocers. Even the mom and pop stores (known as kirana shops) are affected due to this ―unfair‖ back-end advantage extended by the supplier to its leading accounts (the emerging supermarket chains).

The strategies adopted by the retailer to compete with branded goods are illustrated by the following diagram. Branding the store and following a private label strategy is the key strategy which helps the retailer to compete with branded products.

2.4.2. Challenges Ahead For Retailing

The unorganised nature of retailing has stunted its growth over several years. "Lack of industry status affects financing prospects and stunts growth of the industry", says Kishore Biyani, managing director, Pantaloon Retail India. In the current scenario, only players with deep pockets have been able to make it big. In addition to the advent of Internet, there are many other challenges which retailers have to address.

• ***Human Resources***

Availability of trained personnel and retaining the human resources is a major challenge for these big retailers. The bigwigs like Crossroads offer high compensation and create a cohesive environment that makes an employee proud to be a part of such big retail chains.

• ***Space and Infrastructure***

To establish a retail shop / mall, the real estate and the infrastructure are very vital. The expenditure and availability on both the accounts do hinder the growth of the retail chain. The lack of secondary infrastructure also affects the logistics and supply chain management for retail companies.

• ***Absence of retailer friendly laws***

India still does not have retail-friendly laws especially relating to the movement of goods from one state to another. Retailers need to put in a whole lot of products from different parts of the country - at times from outside the country - on the shelf. But question of multiple tax levels is an issue. Then there are laws like shops cannot be open for all seven days, shops have to be open after or close before a certain time which affects operations.

• ***Lack of technical know-how***

The Indian government does not encourage any foreign direct investment (FDI) in the retail industry. FDI is normally one of the ways of getting technical inputs. And because of this dearth of FDI in this sector, development in terms of people, skills etc is happening the hard way.

2.3.3. Future perspective

We should see fundamental shifts in the way Indians shop in the very near future. The Year 2003 could well be a landmark year for organised Indian retailing. According to a recent study done by ETIG the organized retail industry is expected to grow by 30 per cent in the next five years and is expected to touch Rs. 45,000 crore. Thus, the growth potential for the organised retailer is enormous. In

the next 2-3 years, India will finally see operations of a number of very serious international players - notwithstanding the current restrictions on FDI in retail. Metro from

Germany is a very successful and resourceful retailer and their cash & carry format should offer a good run for money to others. Some others will also find perfectly legitimate ways to operate in India, for example, Marks & Spencer, Mango and Shoprite.

Change Accelerators

The following factors will be significant in driving growth in the retail sector:

• Consumer factors

• Increase in income

• Working women

• Changes in lifestyle – demand for ―global‖ trends

• Supply side factors

• Growing importance of retailing in political and economic agenda

• Real estate reforms to be undertaken in the next 24 months

• Major restructuring of the manufacturing sector easing product

supply constraints for efficient retailing

• Reduction in import duties-offering more global sourcing options

Which categories will grow?

The single biggest opportunity in India in organised retailing is bound to be food and groceries; it is in this sector that the largest amount of consumer spends is

concentrated. This sector has maximum opportunity for investments and entrepreneurs to come in and try to make the supply chain a little more efficient. Consumer durables is another promising sector because, with increasing purchasing power, consumers tend to spend the most on this category. Also, there is nothing to prevent a company from putting up shops outside the city limits, because consumer durables are a premeditated purchase. Furthermore, availability of finance options has increased spending in this sector. Third are home products - with increasing private ownership of homes by relatively young couples, across most major cities in India, national retail chains offering home furniture (and accessories) have great potential.

Finally, personal care products, pharmaceutical products, and healthcare services have tremendous growth potential. Recently, we have seen some interest from organised healthcare players like Max, Fortis, Birlas and the Reliance group

Where is this growth going to happen?

The top 15 cities in India cater to 33 per cent of total urban population, but as high as 38 per cent of Sec A and B (the top two socio-economic consumer strata) urban population. The next 15 cities only add to another 7 per cent of Sec A and B population. So logically, the focus will be restricted only to the top 15 cities. Research conducted by KSA Technopak, shows that today 96 per cent of total organised retail is in the top 10 cities, of which the top six cater to 82 per cent. However, the rate of growth will be higher in the bottom four of the top 10, which will have a 20

per cent share by 2005 against the present share of 15 per cent.

Which formats will grow?

KSA Technopak's research suggests the top four formats to emerge in the next five years are:

* Shopping Malls

* Specialty Stores (in new categories such as office products, specialty food, optical and travel)

* Departmental Stores

* Supermarkets

Recipe for Success

• **Focus on the consumer**: It is clear that consumers have changed and they are looking for something different. Understanding their evolving needs, aspirations and lifestyles is the underlying key to success for any retailer. The primary emphasis should be on access, experience and service and the secondary emphasis on product and price. There should be an effort to improve service by having better trained sales staff, better availability of products, and minor but important conveniences, e.g. delivery of goods either to the car or even home. Collaborative advertising and promotion can then round off this effort

• **Brand the store**: branding the store will increase volume and enhance customer loyalty. Branding is critical to maintaining competitive differentiation in an increasingly challenging retail environment. However, the brand needs to be clearly communicated to the customer.

• **Develop private label brand:** Private labels act as margin generators, increasing sales volume by positioning the label as providing higher perceived value to consumers. In the long run, they also increase the retailers‘ bargaining power with national brand suppliers. Private labels generate customer loyalty by providing exclusive products, which works towards differentiation strategy, much sought after by the retailers.

In terms of geography some entrepreneurs should put efforts in creating custom-developed solutions for tapping the rural and semi-urban spending potential. Even in non-metro urban centres, there are very good opportunities in looking at starting or expanding operations. Some cities that should see greater organised retail action in the future would be Ludhiana, Chandigarh, Lucknow, Nagpur, Ahmedabad, Surat, Pune, Kochi, Thiruvananthapuram, Guwahati and Bhubaneshwar.

In terms of format malls have a sustainable competitive advantage over other formats. Consumer preferences are shifting towards malls from traditional markets. As a result of consumer shifts, retailers also prefer to be located in malls in anticipation of higher footfall. KSA Consumer Outlook 2000SM shows that increasingly consumers prefer "All Under One Roof‖ destination for shopping as well as eating out and entertainment. These findings together indicate an excellent potential for a mall with the following features:

• a superior well-managed leisure experience

• targeted at all members of the household

• comprising of shopping, dining and entertainment, all under one roof

• a wide range of products and services

• proximity to homes

Conclusion - How can it be done?

For a start, these retailers need to invest much more in capturing more specific market intelligence as well as almost real-time customer purchase behaviour information. The retailers also need to make substantial investments in understanding/acquiring some advanced expertise in developing more accurate and scientific demand forecasting models. Re-engineering of product-sourcing philosophies - aligned more towards collaborative planning and replenishment should then be next on their agenda. The message, therefore, for the existing small and medium independent retailers is to closely examine what changes are taking place in their immediate vicinity, and analyse whether their current market offers a potential redevelopment of the area into a more modern multi-option destination. If it does, and most commercial areas in India do have this potential, it would be very useful to form a consortium of other such small retailers in that vicinity and take a pro-active approach to pool in resources and improve the overall infrastructure. The next effort should be to encourage retailers to make some investment in improving the interiors of their respective establishments to make shopping an enjoyable experience for the customer.

2.5 Choice of a Retail Store and Retail Store Format

Introduction

Lately, retail has been one of the growth areas in the global economy. It has witnessed a high growth rate in the developed countries and is poised for an exponential growth, in the emerging economies. Along with the rapid growth, retailing scenario has also been characterized by increasing competition and emergence of increasingly new retailing formats (Popkowski Leszczyc, Sinha, and Timmermans, 2000). With an overlap of merchandise being offered across different formats, the competition has become intense and unpredictable in terms of the direction where it is coming from. In the light of these, the study of how consumers choose retail stores, and what drives the store choice, cannot be overemphasized. The emergence of a variety of retail formats, offering a diverse mix of offerings to the consumers, adds further confusion to the domain of store choice. One way to look at the problem of store choice is then to acknowledge the emergence of various store formats and incorporate them into the models for store choice.

Outline

The one of the objective of this research is to integrate store choice and format choice in a single framework. We first do an extensive literature of the store choice and the format choice, literature and identify gaps. We then propose an integrated model of store and format choice to address the issue of format and store choices.

2.5.1 Literature Survey – Store Choice

Store choice has been a subject of wide research and has been studied from various perspectives. The store choice behaviour of shoppers has been found to be quite similar to the brand choice behaviour of the consumers, with a difference being the incorporation of the spatial dimension in store choice (Sinha and Banerjee, 2004). Therefore, while brand choice is independent of the location aspect, and is not affected by it, the store choice is very much influenced by location (Fotheringham, 1988; Meyer and Eagle, 1982). One view, in the store choice literature gives primacy to the store location and believes that the consumers are influenced by the travel costs of shopping (Brown 1989; Craig, Ghosh, and McLafferty 1984; Huff 1964) and store location therefore plays an important role in the store choice. A number of studies, have considered, and pointed out the primacy of store location (Arnold, Oum and Tigert, 1983; Freymann, 2002) in store choice.

Another view in store choice literature focuses on the store attributes. Price is one of the easily noticeable attributes and considerable work exists (Bell, Ho and Tang, 2001; Freymann, 2002; Arnold, Oum and Tigert, 1983), on how, the price of store offerings, affects the store choice. The role of store atmospherics, store ambience and store environment has also been studied as a part of store attributes. A number of studies (Kotler, 1973; Baker, Grewal and Levy, 1992) have studied these and found important relation with consumer store choice. Then there are studies which look at how store environment cues influence consumers' store choice decision criteria, such as perceived merchandise value and shopping

experience (Baker, Parasuraman, Grewal, and Voss, 2002). Store choice, has also been studied, taking the store image (Martineau, 1958) into account and has been found to affect store choice.

Yet another view of store choice, gives more importance to the consumer side, and has looked at the consumer attributes, as well as the situational and tasks associated with shopping. So the store choice has been seen in the context of the risk reduction strategies of the shoppers (Mitchel and McGoldrick, 1996; Mitchell and Harris, 2005). In addition work on store choice has also been done on the role of situational factors (Wu, Petroshius, and Newell, 2004; Mattson, 1982) and the task-store attribute relationship (Kenhove, Wule, and Waterschoot, 1999). It has also been found to be dependent on the timing of shopping trips, with consumers visiting smaller local store for short ―fill-in‖ trips and larger store for regular shopping trips (Kahn and Schmittlein, 1989). It has also been shown by Bell and Lattin (1998) that there exists a logical relationship between a household's shopping behavior and store preference. A narrower segment of the store choice research has been devoted to studying individual difference variables, such as demographic, socio-economic, or psychological variables, as the key predictors of store choice (Bellenger, Robertson and Hirschman 1976; Douglas, 1976; Monroe and Guiltinan, 1975; Winn and Childers, 1976). One drawback of the research in this field has been that though the studies identified relationships, the strength of relationship with the store choice was found to be weak (Mattson, 1982).

2.5.2 Literature Survey – Format Choice

The literature on format choice is limited in nature and is of more recent origin. The recent interest in store formats is mainly attributed to innovations in the mix that the retailers are coming up with, owing to the competition. A store format has been defined as the mix of variables that retailers use to develop their business strategies and constitute the mix as assortment, price, and transactional convenience and experience (Messinger and Narsimhan, 1997). It has also been defined as a type of retail mix used by a set of retailers (Levy and Weitz, 2002). Different store formats are derived from various combinations of price and service output (Solgaard and Hansen, 2003).

The format literature can be traced back to the discussion on cross shopping, which was first discussed in the trade literature in the late 1970s (Cort and Dominguez, 1977). This is recognised now as the ―incidence of consumers shopping at different types of retailer formats for products also commonly referred to intra-type competition (i.e. two different retail formats that sell substitutable products or services)‖ (Carpenter and Moore, 2006, p4). It has evolved further with studies based on grocery as well as other sectors and has dealt with issues dealing with, within chain choice (Cort and Dominguez, 1977), within product sector choice (Cassill and Williamson, 1994), choices based upon marketing and store attributes (Gehrt and Yan, 2004; Hansen and Deutscher, 1977) and multi-channel choices (Schoenbachler and Gordon, 2002). With the new formats being introduced in the evolving markets, the retail offering of these store formats in the

evolving markets has also been studied across different product categories (Sinha and Banerjee, 2004).

The choice of retail formats is richer in studies with consumer attributes as explanatory variables, and a lot of work has specially been devoted to the consumer demographics. The study of Crask and Reynolds (1978) dealt with frequent and non-frequent shoppers to the departmental stores, and found frequent patrons tended to be younger, more educated, and had higher incomes. In another study, Sampson and Tigert (1992) found that warehouse club members were more upscale as compared to the general population, were more educated and had higher incomes. Similarly, Arnold (1997) found significant differences between the demographic profiles (e.g. age, education, household size) of large-format department store shoppers and non-shoppers. Similarly, the work of Carpenter and Moore (2006) found that certain demographic groups were associated with certain store formats. In addition their study also examined store attributes (e.g. price competitiveness, product selection, and atmosphere) as drivers of format choice. Bhatnagar and Ratchford (2004) developed a general model of retail format choice for non durables, and they demonstrated that the retail format choice depended on a number of factors such as travel costs, consumption rates, perishability of products, inventory holding costs of consumers etc.

Studies have also been conducted on shopper behaviour and format choice. In a study of store choice behaviour among audio equipment shoppers, Dash et al. (1976) found shoppers having higher levels of pre-purchase information shopped

at specialty store, while those with low pre-purchase information purchased at departmental stores. In another study, Bell and Lattin (1998) demonstrated that large basket shoppers preferred EDLP formats, while, small basket shoppers, preferred HiLo stores, similar results were arrived at by Bell, Ho and Tang (2001).

2.5.3 Summary of the literature

A review of the literature on store choice and format choice reveal the following:

The research in store choice literature has a beginning far earlier, than format choice literature.

The literature in store choice had more width as well as depth, as compared to the store choice literature.

The literature looks at both as separately, and few overlapping studies, which incorporate store as well as format choice are available.

The store choice literature, is heavily loaded towards studies examining, store choice based on store attributes. Even where, the consumer attributes (demographics etc.) have been used, the relationship is weak. In format choice, the stress is more on consumer attributes (mainly demographic), leading to particular types of formats.

The store choice literature, has studied, stores within the same format, and fails to identify that competition exists across formats. In addition, store attributes, identified in the literature are not unique to the store, but are rather shared by a number of stores operating within the same format (and some even across formats).

2.5.4 The question of shopper attributes in store choice

That the store choice, can be completely captured based only on store attributes and its images, and ignoring the shopper attributes especially when there is considerable consumer diversity, seems to be preposterous. The store choice will be governed both by the store attributes as well as the consumer attributes. Looking at the store attributes in isolation is like looking at the supply side (what the store offers), and not on the demand side (what the shopper is looking for). To capture store choice, it is essential that, both the store attributes as well as the shopper attributes are captured. The problem comes from the fact that in the existing studies on store choice, the relationship between, store choice and the shopper attributes are very weak. This could be a consequence of, low store loyalty and significant store switching which is significant for grocery store purchases (Kau and Ehrenberg, 1984; Uncles and Hammond, 1995; Popkowski Leszczyc and Timmermans, 1997}. It is fairly, well established in the literature that store switching, is quite widespread and unstable even in short term. Actually it has been seen that many consumers regularly visit two or more stores simply because they undertake shopping trips from different places as home, office etc (Solgaard and Hansen, 2003), different preferences for stores, based on the composition of the basket of goods, to profit from the lowest prices at the various stores or by engaging in multi-stop, multipurpose trip behaviour (Popkowski Leszczyc and Timmermans, 1997). Another reason why the relations are weak might be, that these did not take into account the store formats.

One way of overcoming, these problems, is in integrating store choice, format choice, and the consumer attributes within the same framework. In the literature pertaining to store choice the consumers evaluate a group of stores on a set of attributes and then, depending upon their individual preferences, patronize the best store. It has generally been seen that all the stores in the choice set are in the same formats (Bhatnagar and Ratchford, 2004). This indicates that the first choice for the shopper is that of the format and store is the subsequent choice. An analysis of the store switching behaviour by (Galata, et.al, 1999), revealed modest levels of inter-format switching, but a large extent of intra-format switching, their study further found that when shoppers switch they choose a store of the same format. This again indicates that, the choice is at two levels, the format and then the store.

Accordingly, this research attempts to model the store choice of shoppers based on the demographic variables and other socio-economic factors. The demographic variables, are stable within a short range, and the literature on the format choice (Galata, Randolph, Bucklin, and Hanssens, 1999) has also assumed, the format choice to be relatively stable in short term. Therefore, incorporating the format choice, through the stable demographic variables, it is possible, to model store by incorporating consumer variables at a higher level and store attributes at a lower level.

Plausible Contribution

This research attempts to make a number of contributions to the marketing research. This research attempts to model the store choice, which is is depending on their demographic attributes (are almost stable in nature), and their preference of importance on the store images (dynamic).

2.6 Cluster Validity Function

While clustering and segmentation algorithms are unsupervised learning processes, users are usually required to set some parameters for these algorithms. These parameters vary from one algorithm to another, but most clustering/segmentation algorithms require a parameter that either directly or indirectly specifies the number of clusters/segments. This parameter is typically either k, the number of clusters/segments to return, or some other parameter that indirectly controls the number of clusters to return, such as an error threshold. Setting these parameters requires either detailed pre-existing knowledge of the data, or time-consuming trial and error. The latter case requires repeated runs with different parameters, and still requires that the user has sufficient domain knowledge to know what a good clustering ―looks‖ like. However, if the data set is very large or multidimensional, human verification can become difficult.

2.6.1 Literature Review on determining number of clusters

Past literature indicates on determining the numbers of clusters are given below. In most real life clustering situations, an applied researcher is faced with the dilemma

of selecting the number of clusters or partitions in the final solution (Everitt, 1979; Sneath & Sokal, 1973). Virtually all clustering procedures provide little if any information as to the number of clusters present in the data. Non-hierarchical procedures usually require the user to specify this parameter before any clustering is accomplished and hierarchical methods routinely produce a series of solutions ranging from n clusters to a solution with only one cluster present (assume n objects in the data set). As such, numerous procedures for determining the number of clusters in a data set have been proposed (Dubes & Jain, 1979; Milligan, 1981c; Perruchet, 1983). When applied to the results of hierarchical clustering methods, these techniques are sometimes referred to as stopping rules. Often, such rules can be extended for use with non-hierarchical procedures as well.

The application of a stopping rule in a cluster analytic situation can result in a correct decision or in a decision error. Basically, two different types of decision errors can result. The first kind of error occurs when the stopping rule indicates k clusters are present when, in fact, there were less than k clusters in the data. That is, a solution containing too many clusters was obtained. The second kind of error occurs when the stopping rule indicates fewer clusters in the data than are actually present. Hence, a solution with too few clusters was obtained. Although the severity of the two types of errors would change depending on the context of the problem, the second type of error might be considered more serious in most applied analyses because information is lost by merging distinct clusters.

Determining the number of clusters in a data set, a quantity often labelled k as in the k-means algorithm, is a frequent problem in data clustering, and is a distinct issue from the process of actually solving the clustering problem. For a certain class of clustering algorithms (in particular k-means, k-medoids and Expectation-maximization algorithm), there is a parameter commonly referred to as k that specifies the number of clusters to detect. Other algorithms such as DBSCAN and OPTICS algorithm do not require the specification of this parameter. The correct choice of k is often ambiguous, with interpretations depending on the shape and scale of the distribution of points in a data set and the desired clustering resolution of the user. In addition, increasing k without penalty will always reduce the amount of error in the resulting clustering, to the extreme case of zero error if each data point is considered its own cluster (i.e., when k equals the number of data points, n). Intuitively then, the optimal choice of k will strike a balance between maximum compression of the data using a single cluster, and maximum accuracy by assigning each data point to its own cluster. If an appropriate value of k is not apparent from prior knowledge of the properties of the data set, it must be chosen somehow. There are several categories of methods for making this decision.

Rule of thumb - One simple rule of thumb (Kanti Mardia et al., 1979) sets the number to $K \approx \sqrt{n/2}$ with n as the number of objects (data points).

The Elbow Method - Another method looks at the percentage of variance explained as a function of the number of clusters: You should choose a number of clusters so that adding another cluster doesn't give much better modelling of the

data. More precisely, if you graph the percentage of variance explained by the clusters against the number of clusters, the first clusters will add much information (explain a lot of variance), but at some point the marginal gain will drop, giving an angle in the graph. The numbers of clusters are chosen at this point, hence the "elbow criterion". This "elbow" cannot always be unambiguously identified (David J. Ketchen, Jr & Christopher L. Shook, 1996). Percentage of variance explained is the ratio of the between-group variance to the total variance. A slight variation of this method plots the curvature of the within group variance (Cyril Goutte, Peter Toft, Egill Rostrup, Finn Årup Nielsen, Lars Kai Hansen, March 1999). The method can be traced to speculation by Robert L. Thorndike, 1953.

Information Criterion Approach - Another set of methods for determining the number of clusters are information criteria, such as the Akaike information criterion (AIC) (Akaike, Hirotugu, 1974), Bayesian information criterion (BIC)

(Schwarz, Gideon E, 1978), or the Deviance information criterion (DIC) : if it is possible to make a likelihood function for the clustering model. For example: The k-means model is "almost" a Gaussian mixture model and one can construct a likelihood for the Gaussian mixture model and thus also determine information criterion values (Cyril Goutte, Lars Kai Hansen, Matthew G. Liptrot & Egill Rostrup , 2001).

An Information Theoretic Approach (Catherine A. Sugar and Gareth M. James , 2003). Rate distortion theory has been applied to choosing k called the "jump" method, which determines the number of clusters that maximizes efficiency while minimizing error by information theoretic standards. The strategy of the algorithm

is to generate a distortion curve for the input data by running a standard clustering algorithm such as k-means for all values of k between 1 and n, and computing the distortion (described below) of the resulting clustering. The distortion curve is then transformed by a negative power chosen based on the dimensionality of the data. Jumps in the resulting values then signify reasonable choices for k, with the largest jump representing the best choice.

The distortion of a clustering of some input data is formally defined as follows: Let the data set be modelled as a p-dimensional random variable, X, consisting of a mixture distribution of G components with common covariance, ?. If we let $c_1 ... c_K$ be a set of K cluster centers, with c_X the closest center to a given sample of X, then the minimum average distortion per dimension when fitting the K centers to the data is:

$$d_K = \frac{1}{p} \min_{c1 \ldots c_K} E[(X - c_X)^T \Gamma^{-1} (X - c_X)]$$

This is also the average Mahalanobis distance per dimension between X and the set of cluster centers C. Because the minimization over all possible sets of cluster centers is prohibitively complex, the distortion is computed in practice by generating a set of cluster centers using a standard clustering algorithm and computing the distortion using the result. The pseudo-code for the jump method with an input set of p-dimensional data points of X.

The choice of the transform power Y = (p / 2) is motivated by asymptotic reasoning using results from rate distortion theory. Let the data X have a single,

arbitrarily p-dimensional Gaussian distribution, and let fixed K = floor(α^p), for some α greater than zero. Then the distortion of a clustering of K clusters in the limit as p goes to infinity is α^{-2}. It can be seen that asymptotically, the distortion of a clustering to the power $(-p/2)$ is proportional to α^p, which by definition is approximately the number of clusters K. In other words, for a single Gaussian distribution, increasing K beyond the true number of clusters, which should be one, causes a linear growth in distortion. This behavior is important in the general case of a mixture of multiple distribution components.

Let X be a mixture of G p-dimensional Gaussian distributions with common covariance. Then for any fixed K less than G, the distortion of a clustering as p goes to infinity is infinite. Intuitively, this means that a clustering of less than the correct number of clusters is unable to describe asymptotically high-dimensional data, causing the distortion to increase without limit. If, as described above, K is made an increasing function of p, namely, K = floor(α^p), the same result as above is achieved, with the value of the distortion in the limit as p goes to infinity being equal to α^{-2}. Correspondingly, there is the same proportional relationship between the transformed distortion and the number of clusters, K.

Putting the results above together, it can be seen that for sufficiently high values of p, the transformed distortion $d_K^{-p/2}$ is approximately zero for K < G, then jumps suddenly and begins increasing linearly for K >= G. The jump algorithm for choosing K makes use of these behaviors to identify the most likely value for the true number of clusters.

Although the mathematical support for the method is given in terms of asymptotic results, the algorithm has been empirically verified to work well in a variety of data sets with reasonable dimensionality. In addition to the localized jump method described above, there exists a second algorithm for choosing K using the same transformed distortion values known as the broken line method. The broken line method identifies the jump point in the graph of the transformed distortion by doing a simple least squares error line fit of two line segments, which in theory will fall along the x-axis for K < G, and along the linearly increasing phase of the transformed distortion plot for K >= G. The broken line method is more robust than the jump method in that its decision is global rather than local, but it also relies on the assumption of Gaussian mixture components, whereas the jump method is fully non-parametric and has been shown to be viable for general mixture distributions.

Choosing k using the Silhouette - The average silhouette of the data is another useful criterion for assessing the natural number of clusters. The silhouette of a datum is a measure of how closely it is matched to data within its cluster and how loosely it is matched to data of the neighbouring cluster, i.e. the cluster whose average distance from the datum is lowest (Peter J. Rousseuw , 1987). A silhouette close to 1 implies the datum is in an appropriate cluster, whilst a silhouette close to

− 1 implies the datum is in the wrong cluster. Optimization techniques such as genetic algorithms are useful in determining the number of clusters that gives rise to the largest silhouette (R. Lleti, M.C. Ortiz, L.A. Sarabia, M.S. Sánchez, 2004).

2.6.2 Present Contribution

In this research it has been desired to utilize the following (Xie and Beni‘s function, the compactness and separation validity_s, function, the Partition Index SC,Dunn's Index (DI) and Alternative Dunn Index (ADI)) algorithms that can efficiently determine a reasonable number of clusters/segments to return from any non- hierarchical clustering/segmentation algorithm. In order to identify the correct number of clusters to return from a non-hierarchical clustering/segmentation algorithm, this research utilizes the above mentioned cluster validity function.

Most of the research often uses clustering analysis as a tool for market segmentation. In this research, the k-means partitioning method is used to segment the customers based on store image attributes. Such non-hierarchical method of clustering algorithm always tries to find the best fit for a fixed number of clusters and the parameterized cluster shapes. However this does not mean that even the best fit is meaningful at all. Either the number of clusters might be wrong or the cluster shapes might not correspond to the groups in the data. To overcome such problems, this research utilises the (Xie and Beni‘s 1991) cluster validity function which is to solve the cluster membership,c‘, and to measure the effectiveness of cluster. Xie and Beni‘s validity index aims to quantify the ratio of the total variation within clusters and separation of clusters. The Xie and Beni‘s function is

$$XB(c) = \frac{\sum_{i=1}^{c}\sum_{j=1}^{N}(\mu_{ij})^{m}\left\|x_j - v_i\right\|^2}{N \min_{i,k}\left\|x_j - v_i\right\|^2}$$

and $XB(c)$ is the optimal number of cluster which should minimize the value of the index. In continuation to further validate, the compactness and separation validity _s, function is used. The validity function is

$$s = \frac{\sum_{i=1}^{c}\sum_{k=1}^{n}(\mu_{ik}{}^2)d^2\,(v_i - x_i)}{n \min_{ij} d^2(v_i - x_i)}$$

The smaller the value of s', the better the compactness and separation between the clustering groups of in-cluster samples. The goal should therefore be to minimize the value of s. At the same time, this approach also allows us to determine the minimal objective function of clustering algorithm. Till further to authenticate, the Partition Index SC is used, and this is the ratio of the sum of compactness and separation of the clusters.

$$SC = \sum_{i=1}^{c} \frac{\sum_{j=1}^{N}(\mu_{ij})^{m}\left\|x_j - v_i\right\|^2}{N\sum_{k=1}^{c}\left\|x_j - v_i\right\|^2}$$

SC is useful when comparing different partitions having equal number clusters. A lower value of SC indicates a better partition. Utilisation of mentioned validity function is well explored in soft computational problems and rarely used in business and social science studies. The results of validity index are will shown in

the figure form. The analysis found the optimum number of cluster is four and thus the value obtained from such validity function can be used to initialize iterative optimisation-based clustering methods and model identification methods. Further to substantiate, the data set was validated with two more Indexes. They are Dunn's Index (DI) and Alternative Dunn Index (ADI). Dunn's Index (DI) is originally proposed to use at the identification of "compact and well separated clusters". So the result of the clustering has to be recalculated as it was a hard partition algorithm.

$$DI\ (C) = \min_{i\varepsilon c}\{\min_{j\varepsilon c, i\neq j}\left\{\frac{\min_{x\varepsilon}c_i, y\varepsilon c_j d(x,y)}{\max_{k\varepsilon c}\{\max_{x,y\varepsilon c} d(x.y)\}}\right\}\}$$

The main drawback of Dunn's index is computational since calculating becomes computationally very expansive as c and N increase. In the alternative Dunn Index (ADI), where the aim is modifying the original Dunn's index, the calculation becomes simpler, when the dissimilarity function between two clusters $(\min_{i\neq j, y\epsilon c_j} d(x,y)$ is rated in value from beneath by the triangle-nonequality:

$$d(x,y) \geq |\ d(y,v_j) - d(x,v_j)\ |$$

Where v_j is the cluster center of the j-th cluster.

$$ADI\ (C) = \min_{\ i\epsilon c}\{\min_{j\epsilon c, i\neq ji}\left\{\frac{\min_{xi\epsilon Ci, xj\epsilon Cj}\ |\ d(y,v_j) - d(x_i,v_j)\ |}{\max_{k\epsilon c}\{\max_{x,y\epsilon c} d(x.y)\}}\right\}\}$$

Note, that the only difference of SC, S and XB is the approach of the separation of clusters. In the case of overlapped clusters the value of DI and ADI are not really reliable because of re-partitioning the results with the hard partition method.

Plausible Contribution

This research attempts to make use of the above mentioned cluster validity function for this research. The above mentioned cluster validity functions are usually used for image processing analysis where the data is images in nature. But such validity functions are rarely used in social science research. So, utilizing such validity function in social science research and finding its merit and value in social science research is reasonable value addition to the literature.

2.7 Stability of segmentation

2.7.1 Literature Background

The segmentation methods that are discussed so far are based on the assumption that the composition and profile of the uncovered segments are stable over time, or at the very least, during the sampling period. Aside from the stability and substantiality requirements for effective segmentation, from a managerial perspective, violation of the assumption of stability or stationarity may invalidate model estimation when data collection spans a long time period, such as in tracking studies. For example, Kamakura, Kim and Lee (1996) considered the possibility that the decision process and the brand preferences could change within the two year period of their scanner data.

Calantone and Sawyer stated in 1978 that the extent to which segments are stable overtime is a neglected aspect of segmentation studies. If a segment identified at one point in time changes in terms of desired benefits or size, the marketing effort targeted at that segment cannot be maintained. Current segment sizes may provide insufficient justification to warrant a target marketing strategy for specific segments if those segments appear to be decreasing in size or changing in terms of demographic and socioeconomic characteristics. Market studies are needed that assessing the stability of segments over relatively long periods of time. (Wind,1978). On the basis of the work by (Calantone and Sawyer, 1978) and (Bockenholt and Langeheine, 1996) two major sources of segmentation stability was formulated. They are

1. Manifest change: The segment membership is stable, but change may occur in the preference or choice structure of customers (or firms) in a segment over time.

2. Latent change: The preference structure of segments is stable, but change may occur in segment size and / or the segment membership of consumers (or firms) overtime. In this research, the early one is referred as internal stability and the later one is referred as external stability of segmentation.

Accordingly, we can distinguish two types of model that address the stability criterion for effective segmentation. One is the mixture model approach for manifest change. It is an extension og the GLIMMIX approach and includes structural changes in the regression part of the model (i.e., time series, Markov,

Hazard and other approaches that address time-dependence within segments). The other is mixture approach for latent change which addresses changing segment membership and size in basically two ways : through a simple extension of the concomitant variable approach and/or through Markov – type transition probabilities among segments overtime.

The above said method comprises all of the applications of dynamic segmentation that have appeared in the marketing literature to date. For completeness, here it is summarized them in the table form (see Table 2.4).

Table 2.4 Marketing application of dynamic stability

Authors	Year	Application	Change Model	Distribution	Mixture Model
Poulsen	1990	Analyses of new brand tries	MC	B	MIX
Poulsen	1990	Analyses of new brand tries	LC	B	MIX
Wedel et al	1995	Hazard model of brand switching	MC	P	GLIMMIX
Ramaswamy	1997	Brand switching after a new product introduction	LC	M	MIX
Kamakura et al.	1996	Nested logit model of brand choice	MC	NM	GLIMMIX
Bockenholt and Langeheine	1996	Purchase quality and brand choice	LC	P-M	GLIMMIX

MC = Manifest change, LC = Latent change
B = Binomial, P = Poisson, M = Multinominal, NM = Nested multinominal

Abratt (1993) and Lockshin, Spawton and Macintosh (1997) stated that segments' stability is a relevant criterion for judging market segmentation. Wright.M (1996)

stated that a segment structure that is unstable does not enable proper marketing action, targeting and positioning. The issues related to measuring the segments' stability is an important concern in segmentation research, which, however, is also a neglected research area. Wedel.M (1998) and Malhotra.N.K (1996) have classified two main ways of understanding segment stability. They are internal and dynamic stability. Calantone.R.J. and Sawyer.A.G (1978) has mentioned that the internal stability relates to whether the analysis of independent samples from a given time period yield consistent segment solutions regarding the nature and number of identified segment. Lockshin, Spawton and Macintosh (1997) and Souter and Sweeney (2003) stated that the dynamic stability refers to whether identified segments at a given point remain unchanged over time in terms of segment number, size and profile.

2.7.2 Contribution to the existing literature

Dynamic segmentation models have been developed only recently, relatively little is known about their performance. For example, which factor determines the choice between models of manifest and latent change? The only satisfactory answer at this point in time is that primarily substantive arguments should be used to guide such choices. If one can assume that segment sizes and memberships are stable, the manifest-change may be more appropriate. If not, obviously, a latent change model is needed. In addition, empirical arguments, such as the relative fir of the two types of the model, may be used. However limited experience in this field (Poulsen, 1990) shows that the two types of approach may yield comparable fit measures. The third solution may be to include both latent and manifest

changes in the same model, and test for the statistical significance of the two components. Several researchers have attempted that approach, with little success because of problems of collinearity, identification and instability of the estimates.

However, Measuring the internal stability of market segmentation based on fuzzy clustering approach is not been explored in the literature and it is simple and one-pass method to compare with hard segmentation results. Thereby, measurement becomes easy.