10 Differential Patterns in Consumer Purchase Preferences using Self-Organizing Maps: A Case Study of China

Bernd Schmitt and Guido Deboeck

Abstract

Market research and consumer segmentation is still in its infancy in the People's Republic of China. Yet, marketers desperately need information about differential response patterns and consumer segments to formulate viable strategies for the drastically growing Chinese consumer market. This chapter presents the results of consumer surveys with hundreds of respondents conducted in Beijing and Shanghai in the spring of 1997. The data were analyzed using self-organizing maps, a technique that offers easy and convenient visualization of the data obtained without imposing stringent linear constraints on the data. Results indicated important differences in the number and type of consumer subsegments in the Beijing and Shanghai markets with respect to influences on purchase decisions (and, to some degree, lifestyles) but large similarities in media behavior, importance of product attributes, and brand attitudes. The study illustrates the usefulness of the methodology of self-organizing maps. Marketers may use the results of the study as a model of how to collect and analyze consumer surveys to formulate strategies and tactics for the Chinese consumer market.

10.1 Introduction

The current Chinese Consumer Revolution represents a transformation of unprecedented historical proportions, and offers Western companies an unprecedented opportunity to capitalize on events. In order to do so, marketers of consumer goods need to have a strong understanding of China's complex and rapidly-changing consumer market. Until now, our understanding of this market has been limited by the lack of reliable data and methods for analyzing perception and consumption patterns in the marketplace. This chapter presents the data of two large-scale consumer surveys that were conducted in Beijing and Shanghai in the spring of 1997. We analyze the data using self-organizing maps, a technique that offers easy and convenient visualization of the data obtained without imposing stringent linear constraints on the data.
10.2 What do we Know about Chinese Consumers?

Before 1979, when Deng Xiaoping declared his open-door policy, China's consumers had little choice in the marketplace. They had to consume products of dismal quality produced by state-owned enterprises in a regulated market. Today, China's consumers have a vast selection in a large variety of categories from low-end/budget items to luxury goods. They have access to new media and information. They are exposed to branded goods and their logos and advertisements. New modern retail stores are rising everywhere. So, how do consumers make purchase decisions in this new environment? How often do they watch the news? How important are different product attributes (e.g., product features, price, quality, the brand image, and the retail environment) to them? What are their attitudes toward brands? Most importantly, are there meaningful differences between different groups of consumers that may be treated as market segments with different wants and needs?

Unfortunately, compared with most other markets, little is known about today's consumers in the People's Republic of China. A major reason for this lack of knowledge is the lack of an infrastructure for market research. Although market research is increasingly being conducted to provide information and intelligence for market decisions, the quality of most market research remains questionable. Samples are frequently biased (e.g., mail and telephone surveys are biased toward those consumers that can easily receive mail and have phones); responses can be unreliable (due to untrained interviewers who in the worst case fake data); and consumers have little experience with Likert scales and semantic differentials and show strong inclinations to give socially desirable responses (e.g., in focus-group research). Moreover, most of these data consist of focus group or surveys that focus on broad-scale geographic and demographic segments. Finally, results are often presented in a poor manner and without consideration of patterns in the data and differences among consumer groups. Being able to provide differential responses and relating them to relevant descriptors - in short, consumer segmentation - is of critical importance.

In this chapter, we first review the few segmentation studies that have been published or made publicly available about the Chinese consumer market. Most of this research has been conducted by commercial research firms rather than academic researchers. We then present the methodology and results of a survey conducted in Beijing and Shanghai in the spring of 1997. In contrast to prior research, the data of the survey describe consumer segments on a more finely tuned scale, focusing on demographic characteristics as well as attitudes and lifestyles. Using the self-organizing map methodology, we identify consumer segments for two key markets in China - the Beijing and Shanghai markets. Aside from providing substantive results that are beneficial for marketers who intend to tackle the Chinese market, we show the usefulness of self-organizing maps for exploring survey data based on consumer attitudes and cognitive responses to marketing variables.

The survey is part of a broader research project that focuses on the behavior of today's Chinese consumers in terms of their perception and response to products, brand campaigns, advertising, retailing and other consumption-relevant behavior. This project is conducted at the China-Europe International Business School (CEIBS) in Shanghai. We will therefore refer to it throughout this article as the CEIBS Survey.

10.3 A Selective Review of the Prior Segmentation Research

Most segmentation research done in the People's Republic of China today focuses on geographic segmentation, i.e., the identification of differences between rural and urban consumers, consumers in the coastal provinces and inner provinces, and between the North and the South. This type of research was pioneered by Gallup China in 1994 with a sample of 3400 people, a roughly representative sample of the adult Chinese population. Respondents were interviewed in their homes. The survey revealed sharp differences between rural and urban consumers. For example, results indicated that the urban population was significantly more likely than the rural population to study advertisements before the purchase of a "da jian" - a big ticket item or durable. Moreover, 52% of urban respondents (vs. 38% of all respondents) would pay higher prices for products of high quality. 41% of urban consumers (vs. 30% of all respondents) stated that they would buy a leading brand regardless of price. Brand-name recognition of foreign brands was also the highest in the cities. For example, overall, Coca Cola had 62% recognition, but 94% in the selected cities. The corresponding numbers for Pepsi Cola were 42% for the rural sample and 85% for the selected cities. Finally, the Chinese spent one-third of their income on food; the figure rises to 37% among urbanites and even to 41% in Shanghai.

Dynamics Decision, a research firm, has investigated broad differences among cities. Each month, a sample of 9000 families is surveyed in cities across China on their income level and other demographics as well as consumer-goods expenditure. The average per capita income as well as disposable income is roughly the same for Beijing, Tianjin and Shanghai but more than twice as much for Shenzhen.

Cooper's and Lybrand has conducted focus groups in Shanghai and distinguishes three segments based on age and sex: women aged 30-45 appreciate "value and convenience." Men aged 30-45 are "utility shoppers," and consumers aged 30 and under are "highly aspirational and interested in ownership and leisure." Moreover, the focus groups indicate that young women among young consumers are the least concerned about price. Many of them, even if their income is low, spend all of their income on cosmetics and fashion. They favor foreign-invested department stores and boutiques for their atmosphere and service.

Louis Harris, another major research firm, conducted a survey of 2500 consumers in Beijing, Shanghai, Guangzhou, Tianjin and Chengdu. The methodology was based on studies by Yankelovich and Shelly, who are considered pioneers in the measurement of social attitudes in the US. The survey indicated that young, more-affluent and better-educated consumers were much more likely to try new products than older, less-affluent and less-educated consumers. Based on their attitudes and behavior toward new product categories or new brands within a category, they were categorized as the innovators. The survey also found that entrepreneurs and others working in the private sector are more likely to
experiment with new products than government employees. Moreover, the young, affluent and educated are more likely to try products admired by others. They have a strong desire to conform to the norms of the reference group.

10.4 The CEIBS Survey

The purpose of the survey conducted at CEIBS was threefold: (1) to provide a more comprehensive survey of the Chinese consumer market in terms of response variables than prior research; (2) to use a reliable methodology and large samples in two key cities, Beijing and Shanghai; and (3) to identify differences between Beijing and Shanghai consumers as well as subsegments within each market.

The surveys were conducted in the spring of 1997 in Beijing and Shanghai by research assistants trained in survey methodology. Participants were given a small present for participation and recruited based on a stratified-sample methodology based on age, gender and income with 30 respondents per cell. Great efforts were made to produce reliable results. To this end, respondents were recruited randomly (following the stratification scheme) but had to complete the survey in rooms that were rented in downtown Beijing and Shanghai in the presence of the research assistants. This procedure provided a safeguard against the practice of faking data by either respondents or field assistants who may have asked somebody else to complete the survey or complete the surveys themselves.

The survey took about 45 minutes to complete. Respondents were asked a variety of questions, mostly on seven-point rating scales, regarding

- attitudes toward advertisements and media;
- influences on purchase decisions;
- media behavior;
- brand attitudes;
- brand loyalty and perceptions of brand parity;
- attitudes toward products of different firms (e.g., state-owned enterprises, joint ventures, imported goods);
- shopping behavior;
- consumer lifestyles.

Moreover, a variety of demographics were assessed, such as sex, education, type of occupation, sector of occupation and income. Not all responses will be discussed here. The variables included will be identified in more detail in Section 10.6.

10.5 Methodology

Simple tabulations of the demographic descriptors and consumer responses are the most common ways used to analyze survey results. Other traditional approaches to find patterns in the data include clustering techniques, e.g., K-means, principal component analysis (PCA), and factor analysis. Multi-dimensional scaling is often used for representing perceived similarities among products or brands.

A common shortcoming of all these methods is the lack of convenient visualization for the decision-maker (e.g., the manager who may be interested in formulating marketing strategies and tactics for consumer segments). In the case of the K-means clustering the set of model vectors that represents the clusters is still a set of points in a high-dimensional space; additional dimensionality-reduction methods are therefore needed for visualization.

Moreover, with traditional clustering methods, the interpretation of clusters is often difficult. Most clustering techniques tend to assign the data to clusters of a particular shape even if there are no actual clusters in the data. If the goal is not just to compress the data but also to make inferences about the cluster structure, it is essential to analyze whether the data set exhibits a clustering tendency.

PCA and factor analysis are standard methods for linear projections of the data onto a much smaller subspace such that the variance of the original data is preserved as much as possible. Indeed, the eigenvalue criterion frequently used is a measure of the variance explained by the model. Yet, as a linear model, factor analysis imposes stringent assumptions on the data and, as the other methods discussed, has severe limitations in visualizing the structure of (nonlinear) data.

Multi-dimensional scaling (MDS) techniques are methods for creating a space in which similarities are represented as distances, given a certain metric. There exists a multitude of MDS approaches: MDS for metric data (where the exact values of the distances between data items are meaningful) and MDS for non-metric data (where the order of the distances between the data items is important). Most MDS methods are computationally intensive and do not construct an explicit mapping function that could be used for mapping new data items. Instead, the projections of all samples are computed in a simultaneous optimization process.

In contrast to the standard methods of representing survey data for segmentation, self-organizing maps, which are widely used in engineering and becoming of increasing value in finance, economics and marketing applications as well, provide easy visualization, impose few assumptions and restrictions, and, like other data-mining techniques, are able to handle large volumes of data in order to isolate patterns and structures in survey data.

SOM is a non-parametric regression technique which is most often used to form a two-dimensional topological representation of the input data. SOM belongs to a class of unsupervised neural network techniques; it is a data-driven approach that extracts relationships from data. The neurons of a SOM represent the general form of the data. Training of neurons through successive presentation of sample data vectors produces an "elastic net" that is stretched to cover the input space of the data. Instead of looking at vast quantities of tabular data or statistics, a SOM provides a map that allows us to visualize an abstraction of the original data. Furthermore, a component plane representation of a SOM provides information about the correlations between data, the division of data in the input space, and the relative distributions of the components. For more details on the SOM approach see Chapter 11.

The SOMs in this paper were created using Viscovery SOMine from Budapics Inc. (a software product that is commercially available and is discussed in detail in Chapters 13 and 15). All responses to the CEIBS questionnaire were treated.
equally, i.e., given the same weight (or priority). The responses were preprocessed by normalization of the columns to variance one so as to obtain a uniform scale for all attributes.

10.6 Major Results

The CEIBS survey of consumers in Beijing and Shanghai, conducted in early 1997, yielded 930 responses of which 436 were from people interviewed in Beijing and 494 from people interviewed in Shanghai. Several records had missing data. When using traditional methods those records with missing data would pose severe problems. SOMs, however, can handle records with missing values.

Since prior research indicated significant differences between consumers in different cities in China, data were analyzed separately for the Beijing and Shanghai samples to avoid bias toward unjustified generalizations. As shown below, the results justify the separate treatment of Beijing and Shanghai consumers in many respects.

As described earlier, the CEIBS survey gathered the following basic demographic data on consumers in Beijing and Shanghai: gender (male = 1, female = 2); age (student = 1, young = 2, middle-aged = 3, old = 4, senior = 5); income level (none = 0, less than 800 RMB per month = 1, between 800 and 3000 RMB per month = 2, more than 3000 RMB per month = 3); education (primary = 1, secondary = 2, college = 3, more than college = 4); occupation (manual = 1, mental = 2, student = 3, teacher = 4, other = 5); sector of employment (state-owned-enterprise = 1, collective = 2, government = 3, foreign enterprise = 4, school = 5, private enterprise = 6, other = 7).

The first four of these variables (gender, age, income level, and education) are progressively ordered. The last two (occupation and sector of employment) do on first sight not appear to be metrically ordered.

In principle one cannot use metric values for inputs to a SOM if the attributes do not represent metric differences. To be used in a SOM, these would have to be separated into binary subattributes, e.g., "occupation is manual" becomes zero or one, "occupation is mental" becomes zero or one, etc.

Alternatively, if there is a background scale for each of these two attributes then the coding can be used as presented above. For example, the scoring of "occupation" could be based on a progressive scale of skill and knowledge level: from manual or not skilled, to teacher, the highest skill level. In this case, "other" would need to stand for a skill level higher than "teacher". Note that in Chinese culture to be called a "teacher" has a very special meaning.

Likewise, "employment" could be based on a progressive scale from public to private employment, or from work in less profitable to more profitable environments. Implicitly, there may be a correlation with income since the salaries and bonuses in state-owned enterprises may be a lot less than in private enterprises. In both cases "other" would have to be considered a higher level than its precedent. In consequence, it is important in the design and coding of a questionnaire to keep in mind the metric ordering of the responses to various attributes. Note that it is very usual to scale items as was done here. This problem is common to all numeric methods of data analysis - and is not peculiar to SOM.

However, other methods often tend to hide this. In view of these limitations we will give only marginal importance to "occupation" and "sector of employment" in the analysis that follows.

Table 10.1 shows the results of using SOM for a basic clustering of this demographic data from the CEIBS survey. The number of clusters was not predetermined. SOM applied to this data mapped into four clusters; the data self-organized itself by minimizing the distance between the input vectors and the original model vectors. Furthermore, the original model vectors adapted during the training process so as to reflect over time the topology of the input data.

The four main clusters in this SOM included: a cluster that matched 73% of the input vectors, a second cluster that matched 14% of the inputs, a third that attracted 8%, and a fourth cluster covering 4% of the input vectors. About 1.4% or 13 responses out of 930 did not fall into any of these four clusters.

The first of these clusters represents mainly a middle-aged group with income between 800 and 3000 RMB per month. Individuals in this cluster have a secondary education and are either involved in collective work units or the government. The second cluster represents mainly female students with no income. They already have a secondary education and are in college. The third cluster represents mostly middle-aged and older people who have higher income than the ones in the first group, work for private enterprises, and have a secondary education. Finally, the fourth cluster represents older and more senior people, mostly females, who work for state-owned enterprises or collective work units and have a secondary education.

The last column in Table 10.1 shows the weighted average profile of the people.

| Table 10.1. Basic demographic clusters in the CEIBS survey |
|-------------|----|----|----|----|----|----|
|             | C1 | C2 | C3 | C4 | C9 | Total |
| Matching records | 677 | 130 | 72 | 36 | 13 | 928 |
| Matching records (%) | 72.95 | 14.01 | 7.76 | 3.88 | 1.4 | 100 |
| Age Group | Mean | Std. deviation | Mean | Std. deviation | Mean | Std. deviation |
| Age Group | Mean | Std. deviation | Mean | Std. deviation | Mean | Std. deviation |
| Income | Mean | Std. deviation | Mean | Std. deviation | Mean | Std. deviation |
| Gender | Mean | Std. deviation | Mean | Std. deviation | Mean | Std. deviation |
| Education | Mean | Std. deviation | Mean | Std. deviation | Mean | Std. deviation |
| Occupation | Mean | Std. deviation | Mean | Std. deviation | Mean | Std. deviation |
| Work Unit | Mean | Std. deviation | Mean | Std. deviation | Mean | Std. deviation |
interviewed: they are middle-aged with incomes between 800 and 3000 RMB per month, with secondary education, working mostly for the government.

Our analysis of the preferences, attitudes and behavior of these groups of people is divided into several parts: part 1 examines consumer purchase decisions; part 2 examines media influences; part 3 investigates the importance of product attributes; and part 4 analyzes consumer attitudes regarding brands. Finally, part 5 summarizes difference in lifestyle and dining habits of people in Beijing and Shanghai.

10.6.1 Influences on Purchase Decisions

Figures 10.1 and 10.2 show the SOM results regarding factors influencing consumer purchase decisions in Beijing and Shanghai, respectively. Each of these figures contains both an overall SOM map and the component planes showing the individual factors influencing the purchase decisions. The overall SOM map indicates several distinct clusters - some large (e.g., clusters 1 to 3 in Figure 10.1) and some very small clusters (e.g., clusters 4 to 6 in Figure 10.2). The main difference between the maps for the Beijing and Shanghai samples is that the former shows six clusters and the latter only three (two of which are very small). To determine the optimal cluster size we tuned each of these maps by selecting a cluster significance that gave the best trade-off between number of clusters and their overall significance. In other words, we selected the maximum cluster number before a significant drop in cluster significance occurred.

Figure 10.1. Factors influencing purchase decisions by consumers in Beijing. CEIBS survey 1997. The main map shows six clusters; the component planes on the right show the influences of various variables, e.g., advice of friends, TV ads, advice of colleagues, sales person, the product features, price, quality, brand name, the opinion of others. Blue areas in these planes indicate lower values; the red areas indicate the higher values. (This figure can be seen in color in the Color Plate Section.)

Figure 10.2. Factors influencing purchase decisions by consumers in Shanghai. CEIBS survey 1997. The main map shows three clusters; the component planes on the right show the influences of various variables, e.g., advice of friends, colleagues, sales person, TV ads, price, features, quality, brand name and the opinion of others. Blue areas in these planes indicate lower values; the red areas indicate the higher values. (This figure can be seen in color in the Color Plate Section.)

The component planes in Figures 10.1 and 10.2 show the respective contribution of each variable to the SOM map. The scale at the bottom of each plane goes from the lowest values to the highest values for each variable. For example, in the case of income, the lowest values correspond with the income class that had no income, while the highest values correspond to the highest income class (3000 RMB or more per month).

The component planes thus can be read as follows: in cluster 1 (Figure 10.1), for consumers in Beijing the quality and features of products have very high importance; on the other hand the opinion of friends, colleagues, and sales persons play very little part in their purchase decisions. Consumers in Shanghai are likewise highly influenced by quality and features of products, but appear to attach more importance to the opinion of friends, colleagues and/or the sales person. TV ads play a more important role in Shanghai than in Beijing, as evidenced by the small area of very high values in the TV ads plane in Figure 10.1 compared with Figure 10.2.

While the visual interpretation of the differences between both consumer groups is quite straightforward from these displays, a more detailed analysis of the shape and difference among clusters can be obtained from a "statistical summary." A sample of such an analysis is provide in Table 10.2. The first part of the table shows the purchase decision influences of consumers in Beijing. Note that the SOM produced six clusters (cluster 0 in the last column represents input vectors that did not match closely to the six main clusters identified by the map).

The significance of these clusters is shown by the number of matching input vectors (out of a total of 463 responses) as well as the percentage of records in each cluster. Note that in the Beijing sample there are three main clusters with
Table 10.2. Summary of clusters of influences on purchase decisions

<table>
<thead>
<tr>
<th></th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
<th>C4</th>
<th>C5</th>
<th>C6</th>
<th>C0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beijing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Matching records</td>
<td>155</td>
<td>152</td>
<td>68</td>
<td>54</td>
<td>4</td>
<td>5</td>
<td>45</td>
</tr>
<tr>
<td>Matching records (%)</td>
<td>35.63</td>
<td>34.94</td>
<td>15.63</td>
<td>1.38</td>
<td>0.92</td>
<td>1.15</td>
<td>10.34</td>
</tr>
</tbody>
</table>

- Age level
- Income
- Gender
- Education
- Occupation
- Unit
- Friends
- TV ads
- Colleagues
- Sales person
- Features
- Price
- Quality
- Brand
- Opinion

Shanghai

<table>
<thead>
<tr>
<th></th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
<th>C0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Matching records</td>
<td>443</td>
<td>11</td>
<td>11</td>
<td>28</td>
</tr>
<tr>
<td>Matching records (%)</td>
<td>89.86</td>
<td>2.23</td>
<td>2.23</td>
<td>5.68</td>
</tr>
</tbody>
</table>

- Age level
- Income
- Gender
- Education
- Occupation
- Unit
- Friends
- TV ads
- Colleagues
- Sales person
- Features
- Price
- Quality
- Brand
- Opinion

between 15% and 36% of matching records and three small clusters with less than 2% of matching records.

The main differences between the three large clusters in Beijing are as follows:

- Cluster 1 represents mainly middle-aged male consumers (35 to 45 years of age), with 800 to 3000 RMB income per month, who have a college degree or higher and are working for the government. The purchase decisions of this group are mainly influenced by product quality, product features, and price. Friends, colleagues, and the opinion of sales persons have a slightly higher influence on this group than on the previous one.

- Cluster 3 represents mainly older people (40 to 49 years of age) who have incomes between 800 and 3000 RMB per month and work in collectivist settings. The purchase decisions of this group are predominately determined by the quality of products and the product features, to a lesser extent by price, brand image, or the opinion of others. A similar group – cluster 6 – whose members work in foreign enterprises, are more influenced by colleagues and sales persons.

Among the consumers surveyed in Shanghai, there is only one dominant cluster, representing 89% of the input records. This group represents mainly middle-aged people, aged 30 to 39 years old, who have a college degree. Their purchase decisions are dominated by product features and quality, by brand names and price, and to a lesser extent by friends and colleagues (but more so than in Beijing). The opinion of the sales person is again the least important influence.

A graphic representation of these various influences is shown in Figure 10.3. The x-axis in this figure shows the clusters or groupings of input vectors; the y-axis shows the sample attributes and the x-axis (vertical) displays the relative importance of each attribute in each cluster.

10.6.2 Media Influences

Several questions of the CEIBS survey pertained to the influence of the media. Consumers in Beijing and Shanghai were asked how often they watch TV every week, when they watch TV, how often they listen to the radio and when. In addition they were asked to rate on scales ranging from 1 to 7 whether they saw ads as deceptive, entertaining, fun to watch, or informative.

Figure 10.3. Three-dimensional plot of factors influencing purchase decisions of consumers in Beijing.
Table 10.3. TV and radio advertising potential

<table>
<thead>
<tr>
<th></th>
<th>TV watching hours/week</th>
<th>TV time*</th>
<th>Radio listening hours/week</th>
<th>Radio time*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beijing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>2.62</td>
<td>4.9</td>
<td>1.68</td>
<td>3.97</td>
</tr>
<tr>
<td>St. dev.</td>
<td>1.16</td>
<td>0.74</td>
<td>1.01</td>
<td>2.30</td>
</tr>
<tr>
<td>Shanghai</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>2.64</td>
<td>4.94</td>
<td>1.79</td>
<td>3.20</td>
</tr>
<tr>
<td>St. dev.</td>
<td>1.08</td>
<td>0.69</td>
<td>0.95</td>
<td>2.28</td>
</tr>
</tbody>
</table>

*A value of 3 corresponds to 13:00 to 17:00 hours; a value of 4 corresponds to 17:00 to 21:00 hours and a value closer to 5 corresponds to 21:00 hours or later.

As Table 10.3 shows, there were few difference between consumers in Beijing and Shanghai: both groups watch about 2.6 hours of TV per week, mostly around 9 pm in the evening. On average consumers listen for about 1.7 hours per week to the radio and do this mostly between 5 and 9 pm in the evenings.

Moreover, consumers in both cities think that current advertisements on TV are not very informative (5.39 to 5.58 on a scale of 1 to 7, where 1 stands for very informative and 7 for do not think that ads are informative); they consider some ads as fun to watch (4.12 to 4.32) or entertaining (3.82 to 4.07), yet marginally deceptive (3.1 to 3.25).

The CEIBS questionnaire also asked about "Internet use": 32 out of 436 (or 7.3%) of the consumers interviewed in Beijing currently use the Internet; in Shanghai there were 69 out of 494 or 14% of the respondents using the Internet.

10.6.3 Importance of Product Attributes

Product features, quality and to a lesser extent price and brands were important factors in the purchase decisions of consumers. In this section we analyze these product-related influences in detail. The people interviewed were asked to rate on a scale of 1 to 7, where 7 indicated "very much", whether features, price, functionality/quality and brand image exercised an influence on their decisions. They were also asked about the importance packaging, the retail environment and advertising played in their purchase decisions. Figure 10.4 shows the result of a SOM on the responses to these questions. In particular, we found that

- senior people attach more importance to quality and price than features, packaging, and retail environment. The least important influences on their part were advertising and promotion.
- middle-aged people attributed more importance to product features and quality than to all the other factors.

10.6.4 Consumer Attitudes on Brands

We compared SOMs derived from consumer attitudes regarding brand products. Table 10.4 highlights some of the differences.
10.6.5 Chinese Lifestyle and Dining Habits

While there are several lifestyle questions in the CEIBS survey that could be analyzed, the remainder of this paper will be restricted to the analysis of one aspect of lifestyle in the survey. The CEIBS survey included several questions on dining habits and food preferences. The specific questions analyzed here include: How many times per week do you go out for business and private dinner? How much money per person do you spend? What restaurants do you prefer (Western, Chinese, no difference)?

The SOM shown in Figure 10.5 shows that an overwhelming majority of the people interviewed go about four to six times per month to a restaurant for private dinner. They also go about two to three times per month to a restaurant for a business dinner. On average they spend about Yuan 41 per person (US$ 4.9) for a private dinner and Yuan 36 per person (US$ 4.3) for a business dinner. They have a small bias towards Chinese restaurants (although the average score between 1.7 and 1.4 would indicate that the preference between Western and Chinese restaurants is almost even).
Table 10.3 also shows there are smaller groups in Beijing (cluster 2) who take less business dinners and spend substantially less on dining for both private as well as business purposes. On the other hand, a very small group in Shanghai has six business dinners plus two private dinners per week and spends on average Yuan 176 per person (US $21) for business dinners (five times as much as the majority group). This group also spends three times as much for private dinners (Yuan 112 or US$ 13.5 per person). This small group in Shanghai has a strong preference for Chinese restaurants (average score 2.1).

10.7 Conclusions

In this chapter, we analyzed preferences, attitudes and behavior of Chinese consumers in Beijing and Shanghai based on data from a survey conducted at CEIBS in Shanghai during the spring of 1997.

The key finding of the study is the presence of Beijing and Shanghai segments as well as distinct Beijing and Shanghai subsegments regarding purchase decisions. Aside from these differences, we found that

- The dominant influences on purchase decisions in general are product quality and features; price and brands to a lesser extent; and the opinion of friends, colleagues or sales persons to the least extent.
- TV watching and radio listening occupy a small amount of time; ads on TV are considered somewhat deceptive, but occasionally fun to watch; ads are considered not very informative.
- Product quality and features are more important than other product related influences.
- Most people interviewed buy brand products because they think those products are of superior quality, they like to change brands, particularly of clothes, but they do not think that there are many differences between brands of appliances.
- The dining habits of the majority of people interviewed indicates that they go to restaurants about six times a month for private dinners and about four times a month for business dinners, except for a small group in Shanghai who go more frequently to restaurants and also spend substantially more on dining.

The study illustrates the usefulness of self-organizing maps as a methodology for representing survey data. We used self-organizing maps to represent the data because they are a nonlinear, non-parametric regression technique that yields topological, two-dimensional representations of the data that allow for easy visualization of patterns and structures in the data. Furthermore, SOM analysis is a data-driven approach that has no problems in dealing with incomplete data vectors or in handling non-metric data. SOMs are thus an ideal data analysis instrument for identifying patterns as well as similarities and differences in the marketplace, which can be used to formulate targeted marketing strategies and tactics.