ISSN: 2688-8653

Research on the Consumption Structure of Chinese Urban Residents in Recent Years

Yumin Zhang*

College of Sciences, Jinan University, Guangzhou 510632, China

Abstract

Understanding the changes in residents' consumption level and consumption structure can provide a decision-making basis for the government to formulate policies and adjust the industrial structure. In this paper, the per capita expenditure data of urban residents in 31 provinces, municipalities, and autonomous regions in China's four five-year plans since the 21st century are analyzed and evaluated. To study the distribution law of living consumption of urban residents in China, cluster analysis and then factor analysis were carried out. Combined with the scores on public factors and comprehensive scores, the comprehensive per capita consumption level of each province, city, and autonomous region was evaluated. Finally, it analyzes the changes in consumption level and consumption structure of urban residents and puts forward reasonable suggestions.

Keywords

Urban Residents; Consumption Structure; Consumption Level; Cluster Analysis; Factor Analysis.

1. Introduction

Over the past 40 years of reform and opening-up, China's economy has achieved rapid development. From 1978 to 2017, China's GDP grew by 33.5 times in real terms, with an average annual growth rate of 9.5 percent, doubling every eight years on average, making China one of the fastest major economies in the world. China's per capita GDP has kept rising, and it has successfully moved from low-income countries to upper-middle-income countries. However, China's economic growth for a long time mainly relies on investment and export, household consumption is one of the shortboards. In the long run, if consumption does not increase correspondingly, the problem of overproduction is inevitable. Therefore, at present, it is very important to increase consumer demand for China's economic growth. Consumption is the most stable and lasting driving force for economic growth. The 13th Five-Year Plan clearly states that "consumption should play a fundamental role in economic growth and focus on expanding household consumption". Studying the characteristics and changing rules of urban household consumption is of great significance for how to effectively expand domestic demand and realize the sound operation of the national economy. Study of urban residents in food, clothing, shelter, the composition of the four basic aspects, an in-depth understanding of consumer behavior and consumption characteristics of urban residents, grasp the hot spot and direction of the urban resident's consumption demand, to guide the town residents reasonably expanding demand, promote consumption structure upgrade, it will cause the upgrading of industrial structure So that the economy in a healthy and reasonable range of development. Xue [1], combined with the five development philosophies, analyzes the Chinese provincial panel data. Qiu [2], combined with the high Keynes consumption function model and econometric methods of Mianyang city residents' consumption level and structure were analyzed. Both the results show that the consumption structure upgrade can effectively promote the quality of economic development, from the perspective of data to validate the importance of the consumption structure upgrades.

ISSN: 2688-8653

As common statistical analysis methods in multivariate statistics, cluster analysis and factor analysis can aggregate data for the overall analysis, so they are often used in macroeconomic data analysis [3-7]. Yang [8] studied and demonstrated the consumption structure data of Chinese residents in 2018 based on factor analysis and cluster analysis, while Shen [9] analyzed the changes in consumption structure of rural residents in Chongqing based on the data from 2004 to 2018 based on factor analysis. This paper intends to conduct cluster analysis and factor analysis based on the data of 31 provinces and autonomous regions in 2003, 2008, 2013, and 2018 in China's four Five-Year plans (the tenth, eleventh, twelfth, and thirteenth Five-Year plans) since the 21st century, and make a comparative analysis and comprehensive evaluation of the above results. Finally, some policy suggestions are given to promote rational consumption and optimize consumption structure.

In this paper, the rest of the structure is arranged as follows, Section 2 to the consumption structure of the related theory knowledge, Section 3 introduces the theoretical basis of factor analysis, Section 4 introduced for clustering analysis of theoretical basis, Section 5 in real data based on factor analysis and cluster analysis, simplify the structure of urban residents consumption index, and it is concluded that the factor comprehensive score. Finally, according to the analysis results, the conclusion and corresponding policy suggestions are put forward.

2. Consumption Structure Related Theory

2.1. Basic Concept of Consumption Structure

The basic connotation of consumption structure mainly includes the following:

(1) The proportion of different types of consumption materials consumed by people in the process of consumption.

(2) In the process of consumption, the respective percentage of each consumer goods and services in terms of quantity and their coordination, substitution, and proportional relations among each other.

(3) The proportion of all kinds of consumption materials formed in the movement of contradiction between consumption and supply in total consumption expenditure and their relationship.

(4) The sum of the correlation and quantitative proportion of various social factors and natural factors as well as between social factors and natural factors in people's living consumption.

2.2. Classification of Consumption Structure

From the specific content of residents' consumption, the consumption structure is divided into four basic aspects: clothing, food, housing, and transportation. It is convenient to compare the changes in resident consumption structure in different periods and regions. The consumption of Chinese residents is mainly divided into eight aspects: food, clothing, housing, household equipment and supplies, medical care, transportation and communication, culture, education and entertainment, and other goods or services. The basic method to study the consumption structure of urban residents in China is to reflect the consumption structure of urban residents by studying the proportion of the consumption expenditure of urban residents in the eight consumption materials and their relationship with each other.

2.3. The Changing Trend of Consumption Structure

At different income stages, people have different demands for the quantity and quality of consumption materials. With the improvement of income level, the demand for consumption materials also develops towards the enjoyment and development of information. According to Maslow's hierarchy of needs theory, human needs always develop from the lower level to the

ISSN: 2688-8653

higher level. The development and change of consumption structure can be roughly divided into three stages:

(1) Meet the basic survival stage. The income level is very low, can only afford the most basic consumer demand, among which food consumption accounts for a large proportion.

(2) Meet basic survival and pursue higher consumption quality. The income level has been improved, the basic survival needs have been met, and the proportion of the consumption of survival materials has been reduced, and the quality of life has been improved to some extent.

(3) Consumption diversification stage. The income level is higher, the life is richer, the demand of each aspect has been better satisfied.

3. Factor Analysis

The concept of factor analysis originated from the statistical analysis of intelligence tests by Karl Pearson and Charles Spearman in the early 20th century. In recent years, with the development of computers, factor analysis has been widely used in medicine, psychology, meteorology, economics, and other fields. Factor analysis is a kind of latent factor that is hidden in multivariate data and cannot be directly observed but affects or dominates measurable variables and estimates the degree of influence of latent factors on measurable variables, as well as the correlation between latent factors.

3.1. Fundamentals of Factor Analysis

To better illustrate the rationale, an example is given. Table 1 shows the correlation coefficients between the three indicators, among which, x_1 represents mathematics achievement, x_2 represents Chinese achievement, and x_3 represents English achievement.

Table 1. Correlation coefficient between 3 indicators

	x_1	x_2	x_3
x_1	1.00	—	—
x_2	0.77	1.00	—
x_3	0.72	0.86	1.00

Let ξ be the potential factor affecting these three indicator variables. Obviously, each indicator variable may be influenced by other factors besides this potential factor, which is denoted as γ. The degree to which ξ affects x_i is expressed as a_i. Mathematically, their relationship can be expressed as follows:

$$\begin{cases} x_1 = a_1\xi + \gamma_1 \\ x_2 = a_2\xi + \gamma_2 \\ x_3 = a_3\xi + \gamma_3 \end{cases} \tag{1}$$

Assuming that ξ is of variance 1, ξ and γ_i are independent of each other, and assuming that the index is normalized to the variable X_i of variance 1, then the covariance between the two variables can be simplified as follows:

$$\begin{aligned}\operatorname{cov}(X_1,X_2)&=\operatorname{cov}(a_1\xi,a_2\xi)+\operatorname{cov}(\gamma_1,\gamma_2)+\operatorname{cov}(a_1\xi,\gamma_2)+\operatorname{cov}(a_2\xi,\gamma_1)\\&=a_1a_2Var(\xi)+0+0+0\\&=a_1a_2\end{aligned} \tag{2}$$

And the index has been standardized, so $a_1a_2=0.77$, similarly:

$$\begin{cases}a_1a_3=\gamma_{13}=0.72\\a_2a_3=\gamma_{23}=0.86\end{cases} \tag{3}$$

A set of solutions can be obtained by solving the equation:

$$a_1=0.897,a_2=0.959,a_3=0.803$$

Thus, the relations between three standard indexes X_i and potential factor ξ and error terms γ_1, γ_2, γ_3 can be obtained:

$$\begin{cases}X_1=0.897\xi+\gamma_1\\X_2=0.959\xi+\gamma_2\\X_3=0.803\xi+\gamma_3\end{cases} \tag{4}$$

3.1.1. Mathematical Model of Factor Analysis

Suppose p indices, $X_1,X_2,\cdots,X_p$ are observed for n samples, and the observed data are obtained. What we need to do is to start with a set of observational data and analyze the correlations between indicators to find the underlying factors that govern them, so that these factors explain the correlations between indicators.

The factor analysis model is described as follows:

(1) $X=(X_1,X_2,\cdots,X_p)$ is an observable random vector, the mean vector $E(X)=0$, and the covariance matrix $\operatorname{cov}(X)$ is equal to the correlation matrix R.

(2) $F=(F_1,F_2,\cdots,F_p)$ is an unobtainable vector, the mean vector $E(F)=0$, and the covariance matrix $\operatorname{cov}(F)=1$, that is, the components of the vector are independent of each other.

(3) $E=(E_1,E_2,\cdots,E_p)$ is independent of F, and $E(E)=0$. The covariance matrix of E is diagonal, that is, the components E are independent of each other.

Then, the mathematical model of factor analysis can be expressed as:

$$\begin{cases}X_1=a_{11}F_1+a_{12}F_2+\cdots+a_{1m}F_m+e_1\\X_2=a_{21}F_1+a_{22}F_2+\cdots+a_{2m}F_m+e_2\\\vdots\\X_p=a_{p1}F_1+a_{p2}F_2+\cdots+a_{pm}F_m+e_p\end{cases} \tag{5}$$

Its matrix form is: $X=AF+e$

F is called the common factor or potential factor of X, matrix A is called the factor loading matrix, and e is called the special factor of X. $E_i\,(i=1,2,\cdots,p)$ denotes a unique factor affecting X_i. A_{ij} is the factor load, which is the weight of the i^{th} variable on the j^{th} principal factor, reflecting the relative importance of the i^{th} variable on the j^{th} principal factor. The basic problem of factor analysis is to determine the factor load.

ISSN: 2688-8653

3.1.2. Properties of Factor Models

The covariance matrix of X is as follows:

$$\sum_X = E(AF+e)(AF+e)' = AA' + \sum e \tag{6}$$

According to the assumptions of the factor analysis mathematical model, it can be known that:

$$\begin{cases} X_i = \sum_{k=1}^{m} a_{ik} F_k + e_i \\ 1 = VAR(X_i) = \sum_{k=1}^{m} a_{ik}^2 + \sigma_i^2 \end{cases} \tag{7}$$

The sum of the squares of the i^{th} row elements in the factor loading matrix A is denoted as h_i^2, which is called the commonness of the variable X_i, i.e

$$h_i^2 = \sum_{k=1}^{m} a_{ik}^2 \text{, have } h_i^2 + \sigma_i^2 = 1, i = 1, 2, \cdots, m.$$

h_i^2 is the contribution of all potential factors to the variance of the original index X_i, reflecting the influence of all potential factors on the variable X_i.

3.1.3. Estimated Factor Load

Since $AA' = V(X) = W'\Lambda W$, the estimated factor load a_{ij} of the factor analysis model is:

$$a_{ij} = w_{ij}\sqrt{\lambda_i} \tag{8}$$

Where, w_{ij} is the j^{th} component of orthogonal eigenvector W_i corresponding to eigenvalue λ_i.

Thus, the calculation procedure of factor load matrix A is as follows:

(1) Calculate the standard index vectors $x_1, x_2, \cdots, x_p$ of p measurable indexes.

(2) calculate the correlation coefficients between p standard index vectors $r_{ij} = Corr(X_i, X_j), i, j = 1, 2, \cdots, p$.

(3) Calculate p eigenvalues of correlation coefficient matrix $V(X) = (r_{ij})$ and order $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k$.

(4) p linearly independent eigenvectors corresponding to all eigenvalues are calculated: $L_1, L_2, \cdots, L_p$.

(5) p a feature vector units: $W_i = L_i / |L_i|$, including $|L_i|$ is a vector L_i mode.

(6) $W = (W_1, W_2, \cdots, W_p)$.

(7) $A = W\Lambda$.

3.1.4. Factor Rotation

Sometimes the factor load is relatively uniform, so it is not easy to see which index the potential factor has the greatest influence directly, so it is not easy to give the potential factor a reasonable variable meaning. In this case, the need to adopt some spinning method, which uses a linear transformation converts the initial potential factor into a new set of potential factors, makes the new potential factor of the absolute value of the factor loading of each indicator to 0 or 1 polarization, which sees that every potential factor influences on the targets of which one

ISSN: 2688-8653

is the largest. Then, a more reasonable explanation of the actual meaning of potential factors is given.

Linear transformation does not change the correlation between variables, so any rotation transformation will not change the number of hidden potential factors in a set of indicators and the actual significance of potential factors.

3.1.5. Factor Score

After the establishment of the factor analysis model, another important role is to use the factor analysis model to evaluate the status of each sample in the whole model, that is, to carry out the comprehensive evaluation. Factor scores refer to the values of potential factors in a model. Since the potential factor is a theoretical variable that cannot be observed directly, its value can only be obtained by means of measurable variables, and its calculation method is generally linear regression.

3.2. Step of Factor Analysis

3.2.1. The Basic Steps of Factor Analysis

(1) Confirm whether the original variables are suitable for factor analysis;

(2) Construct factor variables;

(3) Use rotation method to make factor variables more interpretable;

(4) Calculate the score of factor variables.

3.2.2. The Calculation Process of Factor Analysis

(1) Data standardization;

(2) Find the correlation matrix of standardized data;

(3) Find the eigenvalues and eigenvectors of the correlation matrix;

(4) Calculate variance contribution and cumulative variance contribution rate;

(5) Determining factor;

(6) Factor rotation;

(7) Calculate the score of each factor;

(8) Comprehensive score.

Taking the variance contribution rate of each factor as weight, the comprehensive evaluation index function is obtained from the linear combination of each factor.

4. Clustering Analysis

Cluster analysis is a statistical method to classify random phenomena. Cluster analysis, also known as group analysis and point group analysis, is a multivariate statistical method to study classification.

4.1. Theoretical Basis of Cluster Analysis

4.1.1. The Basic Idea of Cluster Analysis

Cluster analysis is a statistical method that classifies individual samples or index variables according to their characteristics. As the research sample or indicators (variables) between the similarity degree of different, so can according to a sample of multiple observation index to find some can measure the degree of similarity between samples or indicators statistics, with these statistics as the basis of typology, will get together for a big, similar to other higher similarity between each other for another, The closely related are aggregated into a small taxon, and the distance into a large taxon, until all the samples are aggregated.

Cluster analysis is usually divided into Q-type cluster analysis and R-type cluster analysis. Q cluster analysis is the classification of samples, also known as sample cluster analysis; R-type cluster analysis is the classification of indicators, called indicator cluster analysis. The purpose of Q-type cluster analysis is to classify the samples with unclear classification into several groups according to the degree of similarity in nature, so as to discover the commonness of similar samples and the differences between different samples. The purpose of R-type clustering is to divide the indexes that are not clearly classified into several groups according to the degree of similarity in nature, so as to replace the original multiple indexes with a small number of indexes under the condition of not losing information as much as possible.

4.1.2. Statistics for Cluster Analysis

No matter what kind of clustering, the key is how to define similarity. The first step in clustering requires the statistic of similarity between two indicators or samples.

(1) Distance coefficient

There are n samples and p indicators, and the data matrix is:

$$\begin{bmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{np} \end{bmatrix} \tag{9}$$

Since each sample has p indicators, each sample can be regarded as a point in the p-dimensional space, and n samples constitute n points in the p-dimensional space. Therefore, distance can be used to measure the degree of proximity between samples. Make $x_i=\left(x_{i1},x_{i2},\cdots,x_{ip}\right)^T$ is the i^{th} sample observations, $x_j=\left(x_{j1},x_{j2},\cdots,x_{jp}\right)^T$ is the i^{th} sample observations. Then the commonly used distance coefficients between these two samples are as follows:

① Minkowski Distance

$$d_{ij}(q)=\left[\sum_{t=1}^{p}\left|x_{it}-x_{jt}\right|^{q}\right]^{1/q} \tag{10}$$

when $q=1$, is the absolute distance; $q=2$, is the Euclidean distance; $q=3$, is Chebyshev distance. However, before calculating Ming distance, it is often necessary to standardize the data.

② Mahalanobis distance

Let Σ represent the covariance matrix of the index, i.e

$$\Sigma=\left(\sigma_{ij}\right)_{p\times p} \tag{11}$$

where,

$$\sigma_{ij}=\frac{1}{n-1}\sum_{a=1}^{n}\left(x_{ai}-\overline{x}_i\right)\left(x_{aj}-\overline{x}_j\right) \qquad i,j=1,\cdots,p$$

$$\overline{x}_i=\frac{1}{n}\sum_{a=1}^{n}x_{ai},\ \overline{x}_j=\frac{1}{n}\sum_{a=1}^{n}x_{aj}$$

If Σ^{-1} exists, then the Mahalanobis distance between the two samples is:

$$d_{ij}^{2}(M)=\left(x_i-x_j\right)'\Sigma^{-1}\left(x_i-x_j\right)$$

(2) Correlation coefficient

In cluster analysis, the correlation coefficient is exploited to describe the degree of similarity between two indicators.

$x_s=\left(x_{s1},x_{s2},\cdots,x_{sp}\right)^T$ is the s^{th} indicator variable, $x_i=\left(x_{t1},x_{t2},\cdots,x_{tp}\right)^T$ is the t^{th} indicator variable, then the correlation coefficient between x_s and x_t is as follows:

$$r_{st}=\frac{\sum_{i=1}^{n}\left(x_{is}-\overline{x}_s\right)\left(x_{it}-\overline{x}_t\right)}{\sqrt{\sum_{i=1}^{n}\left(x_{is}-\overline{x}_s\right)^2}\sqrt{\sum_{i=1}^{n}\left(x_{it}-\overline{x}_t\right)^2}} \tag{12}$$

(3) Similarity coefficient between classes

① Correlation coefficient

The correlation coefficient, as it is commonly called, is generally for the case between variables. Similar definition can also be given for the similarity relation between characterization samples. The correlation coefficient between the i^{th} and j^{th} samples is defined as follows:

$$r_{ij}=\frac{\sum_{a=1}^{p}\left(x_{ia}-\overline{x}_i\right)\left(x_{ja}-\overline{x}_j\right)}{\sqrt{\sum_{a=1}^{p}\left(x_{ia}-\overline{x}_i\right)^2\sum_{a=1}^{p}\left(x_{ja}-\overline{x}_t\right)^2}} \tag{13}$$

4.2. Clustering Methods

4.2.1. The Clustering Indicators

System clustering is generally used to cluster indicators. The steps of systematic clustering for k indicators are as follows:

Step 1. To determine the minimum proportion p of the total variation of intraspecial indicators in each category that could be explained by category components;

Step 2. All indicators are regarded as a class, and the proportion of total variation of indicators within the class explained by the class components is calculated. If the explained proportion is greater than or equal to p, the clustering is stopped, otherwise to **Step 3**;

Step 3. Split the class into two classes, the principle of classification is to make the total variance in each category indicator as is the class of the class composition is explained, and the correlation coefficient between the minimum, calculate each kind of class in the proportion of the total variance indicator was explained by class composition, if the proportion of explained is greater than or equal to p, the clustering to stop, otherwise to **Step 4**;

Step 4. Finally explain the smallest proportion of the class, and then continue to decompose;

Step 5. Repeat the above steps until the proportion of total variation of intra-class indexes explained by class components is greater than or equal to p for all classes.

4.2.2. The Clustering Samples

Both systematic clustering and stepwise clustering can be used to cluster samples.

The basic idea of clustering samples in the systematic clustering method is to first regard n clustering samples as n classes, then merge the two classes with the greatest similarity into one class according to the distance between classes, and then merge the two classes with the greatest similarity among all the classes into one class. This process is repeated until all of the classes are at a certain distance, or until all of the samples have been combined into one class.

The basic idea of clustering samples by step-up clustering method is to select several initial condensation points, which can be any samples of all samples or several observations of new samples determined at random. Then, each sample is classified into the initial class represented by the nearest condensation point of the sample according to the distance, and then the "center of gravity" of these initial classes is used as the new condensation point to classify the samples again. Repeat the above steps until no more changes are made to the separated classes.

ISSN: 2688-8653

5. The Example Analysis

5.1. Data Source and Processing

The data selected in this paper are the per capita expenditure data of urban residents in 31 provinces, municipalities and autonomous regions of China in 2003, 2008, 2013 and 2018, which are from The China Statistical Yearbook of the National Bureau of Statistics. The China Statistical Yearbook, published annually by the National Bureau of Statistics, is the most comprehensive and authoritative comprehensive statistical yearbook in China. It contains a large number of annual economic and social statistics of all provinces, autonomous regions and municipalities, as well as major statistical data of the whole country in important historical years.

The statistical yearbook has divided per capita expenditure into eight categories, namely food, clothing, housing, household goods and services, transportation and communication, education, culture and entertainment, health care and other goods and services, these eight categories of indicators are identical in units. On this basis, according to the relevant paper materials, food, clothing and housing three indicators are divided into subsistence consumption; The indicators of household goods and services, transportation and communication are divided into development-oriented consumption; and the indicators of education, culture and entertainment, health care are divided into enjoying consumption; then, other goods and services are separately divided into other consumption.

5.2. Data Analysis

5.2.1. Clustering Analysis

Cluster analysis was conducted for each province in China.

The classification of provinces and cities in these four years is summarized in the Table 2.

Consumption is closely related to income, and the correlation is generally positive. It can also be seen from the table that the consumption level of the economically developed areas in Beijing, Shanghai and Guangzhou is often similar, followed by the coastal economic open areas such as Zhejiang, Fujian, Jiangsu and so on. The consumption level of these areas can also be classified into one category. While the inland provinces and the northwest provinces are in the middle and lower level of economic development, and the consumption level is often low.

To further study the changes of consumption structure, the consumption composition ratio of provinces and cities in each year is averaged according to the clustering categories and divided into five categories.

The urban resident's consumption level of the first category of provinces and cities is the highest in the five provinces and cities category, and subsistence consumption only accounts for a little more than half of the total consumption, and hedonic consumption is also the highest among the five categories of provinces and cities. It can be seen that in 2003, the consumption of urban residents living in Beijing, Shanghai, Guangzhou, and Zhejiang could not only meet the basic needs of life but also have more consumption for their development and enjoyment, with less living pressure. For urban residents in Tibet, the pressure of subsistence is still great, and their consumption is already at a low level, and the expenditure on subsistence accounts for 63.8% of the total expenditure, only a small part of which can be used for their own hedonic.

ISSN: 2688-8653

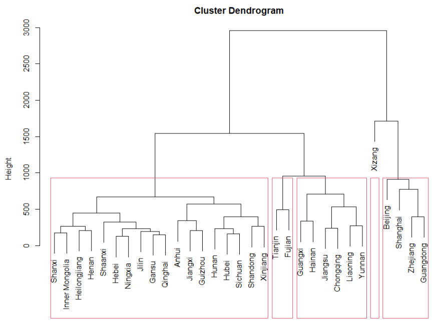

Figure 1. Clustering results of 2003 data

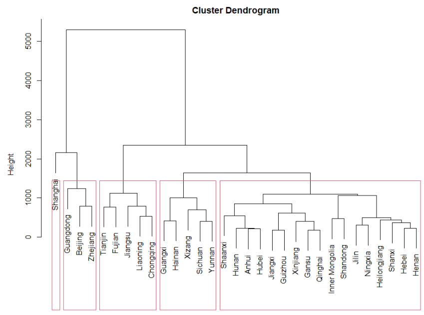

Figure 2. Clustering results of 2008 data

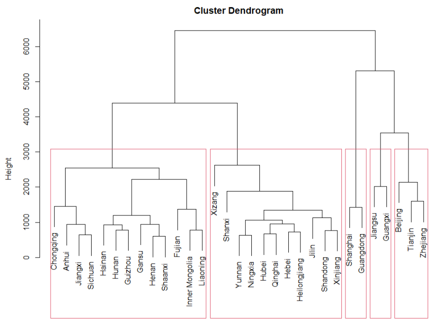

Figure 3. Clustering results of 2013 data

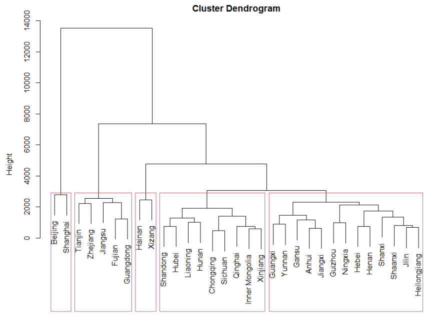

Figure 4. Clustering results of 2018 data

Table 2. Consumption clustering of provinces and cities in four years

	2003	2008	2013	2018
One	Beijing Shanghai Zhejiang Guangdong(10378.34)	Shanghai(19397.89)	Beijing Tianjin Zhejiang(23436.63)	Beijing Shanghai(44470.45)
Two	Tibet(8045.25)	Guangdong Beijing Zhejiang(15715.51)	Jiangsu Guangxi(20776.71)	Tianjin Zhejiang Jiangsu Fujian Guangdong(31156.92)
Three	Guangxi Hainan Jiangsu Chongqing Liaoning Yunnan(6199.01)	Tianjin Fujian Jiangsu Liaoning Chongqing(12055.88)	Shanghai Guangdong(22131.18)	Hainan Tibet(23015.3)
Four	Tianjin Fujian(7611.9)	Guangxi Hainan Tibet Sichuan Yunnan(9223.04)	Tibet Shanxi Yunnan Ningxia Hubei Qinghai Heilongjiang Jilin Shandong Xinjiang(14470.8)	Shandong Hubei Liaoning Hunan Chongqing Sichuan Qinghai Inner Mongolia Xinjiang(24396.67)
Five	Shanxi Inner Mongolia Heilongjiang Henan Shaanxi Hebei Ningxia Jilin Gansu Qinghai Anhui Jiangxi Guizhou Hunan Sichuan Shandong Xinjiang(5414.06)	Shaanxi Hunan Anhui Hubei Jiangxi Guizhou Xinjiang Gansu Qinghai Inner Mongolia Shandong Jilin Ningxia Heilongjiang Shanxi Hebei Henan(9260.74)	Chongqing Anhui Jiangxi Sichuan Hainan Hunan Guizhou Gansu Henan Shaanxi Fujian Inner Mongolia Liaoning(16558.77)	Guangxi Yunnan Gansu Anhui Jiangxi Guizhou Ningxia Hebei Henan Shanxi Shaanxi Jilin Heilongjiang(21364.68)

Note: The content in brackets is the average per capita consumption expenditure of these provinces and cities, and the unit is (yuan/person).

ISSN: 2688-8653

Table 3. Consumption expenditure of urban residents in 2003

	Subsistence consumption			Developmental consumption		Hedonic consumption		Other consumer
	X1	X2	X3	X4	X5	X6	X7	X8
One	0.356	0.073	0.106	0.066	0.131	0.162	0.071	0.035
Two	0.440	0.140	0.058	0.055	0.147	0.074	0.039	0.046
Three	0.402	0.086	0.106	0.058	0.110	0.132	0.074	0.032
Four	0.364	0.117	0.104	0.062	0.103	0.140	0.076	0.034
Five	0.399	0.076	0.130	0.060	0.104	0.130	0.069	0.033

Table 4. Consumption expenditure of urban residents in 2008

	Subsistence consumption			Developmental consumption		Hedonic consumption		Other consumer
	X1	X2	X3	X4	X5	X6	X7	X8
One	0.366	0.078	0.085	0.061	0.174	0.148	0.039	0.048
Two	0.360	0.087	0.093	0.058	0.155	0.138	0.071	0.039
Three	0.389	0.098	0.103	0.061	0.117	0.121	0.072	0.039
Four	0.458	0.094	0.091	0.052	0.130	0.089	0.055	0.031
Five	0.373	0.123	0.107	0.064	0.104	0.113	0.079	0.037

Compared with 2003, the consumption level of urban residents in all provinces and cities increased to some extent in 2008, indicating that living conditions have been partially improved. Among the five categories of provinces and cities, the fourth category (Guangxi, Hainan, etc.) has the highest living pressure. The consumption level of these provinces and cities is the lowest among the five categories. The annual per capita consumption of urban residents is 9223.04 yuan, of which subsistence consumption accounts for 64.3%, while hedonic consumption is only 14.4%. The first category of provinces and cities (Shanghai) has the highest consumption level, and subsistence consumption accounts for the smallest proportion, while hedonic consumption is the highest, indicating that the living standard of urban residents in Shanghai is relatively the highest.

Table 5. Consumption expenditure of urban residents in 2013

	Subsistence consumption			Developmental consumption		Hedonic consumption		Other consumer
	X1	X2	X3	X4	X5	X6	X7	X8
One	0.330	0.099	0.088	0.065	0.173	0.131	0.066	0.048
Two	0.450	0.073	0.078	0.059	0.137	0.129	0.046	0.028
Three	0.241	0.082	0.117	0.073	0.210	0.166	0.056	0.055
Four	0.339	0.124	0.105	0.064	0.132	0.123	0.076	0.038
Five	0.382	0.107	0.096	0.068	0.132	0.117	0.063	0.036

In 2013, for the first time, subsistence consumption accounted for less than half of the total consumption of urban residents in the first category of provinces and cities, and the consumption level of the third category of provinces and cities (Guangdong and Shanghai) was also in the forefront of the five categories.

Table 6. Consumption expenditure of urban residents in 2018

	Subsistence consumption			Developmental consumption		Hedonic consumption		Other consumer
	X1	X2	X3	X4	X5	X6	X7	X8
One	0.221	0.050	0.346	0.053	0.114	0.111	0.075	0.029
Two	0.291	0.060	0.258	0.057	0.136	0.106	0.066	0.026
Three	0.373	0.067	0.213	0.054	0.124	0.088	0.056	0.026
Four	0.283	0.082	0.195	0.067	0.141	0.116	0.088	0.028
Five	0.271	0.080	0.212	0.065	0.136	0.119	0.091	0.026

Comparing the data of the four years, from 2003 to 2018, the consumption level of urban residents in all provinces and cities in China has increased significantly, but the proportion of each consumption type in the total consumption has not changed much, and subsistence consumption has accounted for about 60% of the total consumption. In 2018, the per capita consumption of urban residents in the first category (Beijing and Shanghai) reached 44,470 YUAN, of which up to 60 percent was spent on subsistence consumption, exceeding the total consumption of urban residents in the third, fourth and fifth categories. Prices and housing prices in the north have always been among the highest in the country, which has brought great pressure on urban residents. At the same time, we should also see, in addition to the individual provinces and cities that exist and changes in the ownership category, but most of the provinces, cities and the classification of the situation is not a too big change, this suggests that the low level of consumption of provinces and cities have been in a low level of consumption, didn't get a good development to catch up with the other provinces and cities, and compared with other provinces and cities of consumption level gap has widened gradually. The level of consumption can largely reflect the level of income. From this point of view, it shows that although some of China's original poor regions have achieved considerable development in the past two decades, such development cannot make them catch up with those rich provinces, but the gap between the two has been gradually widened.

5.2.2. Factor Analysis

The more recent the data are, the more revealing the problem is. Therefore, the data of 2018 is selected for factor analysis. The number of factors was selected and the total variance interpretation table was obtained by exploiting R [10-11].

Table 7. Variance interpretation table

	Factor1	Factor2	Factor3
SS loadings	3.648	2.393	0.865
Proportion Var	0.456	0.299	0.108
Cumulative Var	0.456	0.755	0.863

The first row of the data in the table contains the eigenvalues associated with the factor, which are the normalized variance values associated with a particular factor. The second line represents the degree of explanation of each factor to the whole data set. It can be seen that the first factor explains 45.6% of the variance of the 8 consumption indicators, while the thrid explain 86.3% of each indicator.

As the key of factor analysis is to explain each factor, the explanation of each factor to each original observation index is not obvious at this time, and it is difficult to name and summarize, factor rotation is considered at this time so that each factor can explain the relevant information of the original data more clearly and effectively.

ISSN: 2688-8653

Table 8. Factor load matrix

	Factor1	Factor2	Factor3
X1	0.243	0.884	
X2	0.863		-0.501
X3	0.533	0.678	0.311
X4	0.691	0.516	
X5	0.606	0.614	
X6	0.595	0.459	0.502
X7	0.882		0.461
X8	0.764	0.542	0.204

Table 9. Factor load matrix after rotation

	Factor1	Factor2	Factor3
X1	0.916		
X2	0.145	0.206	0.965
X3	0.794	0.450	
X4	0.655	0.394	0.402
X5	0.734	0.345	0.305
X6	0.605	0.671	
X7	0.157	0.929	0.328
X8	0.705	0.549	0.347

Table 10. Comprehensive score and ranking of factors of each province

Provinces and cities	Factor score			Composite score	Rank
	Factor1	Factor2	Factor3		
Shanghai	2.79261	1.80510	0.06473	1.92416	1
Beijing	1.54521	2.26097	0.76801	1.60583	2
Tianjin	0.88124	1.08161	0.69406	0.9048	3
Zhejiang	1.45457	-0.15892	0.95049	0.85200	4
Guangdong	1.88486	-0.79880	-1.35690	0.38628	5
Liaoning	-0.13575	0.89227	0.64915	0.34394	6
Jiangsu	0.41025	0.27548	0.12299	0.30929	7
Xinjiang	-0.30435	0.02166	1.18466	0.10388	8
Inner Mongolia	-0.26834	-0.57818	1.95147	0.09438	9
Fujian	1.21139	-1.33868	-0.74516	0.01950	10
Chongqing	-0.20482	-0.22456	0.56794	-0.05133	11
Shandong	-0.11334	-0.40205	0.58580	-0.05815	12
Hunan	-0.19504	-0.02890	-0.03559	-0.11079	13
Jilin	-0.97538	0.76652	0.50980	-0.13052	14
Sichuan	0.10832	-0.40726	-0.30328	-0.13597	15
Qinghai	-0.83128	0.53695	0.48504	-0.13673	16
Shaanxi	-0.44053	0.33989	-0.30180	-0.17078	17
Ningxia	-0.75305	0.14356	0.40278	-0.23735	18
Gansu	-0.80523	0.29760	0.24028	-0.24860	19
Tibet	0.80197	-3.06071	1.47608	-0.25214	20
Heilongjiang	-1.26376	0.89802	0.22451	-0.28854	21
Hainan	-0.83729	0.55807	-0.32391	-0.30018	22
Hubei	0.23452	0.14901	-2.70754	-0.39949	23
Hebei	-0.66851	-0.28181	0.03037	-0.40471	24
Anhui	-0.08595	-1.19014	-0.27192	-0.46548	25
Jiangxi	0.18551	-1.59977	-0.31654	-0.46972	26
Yunnan	-0.51746	0.13019	-1.26530	-0.47182	27
Henan	-0.85502	-0.06403	-0.26309	-0.48841	28
Guizhou	-0.44561	-0.55950	-0.55527	-0.50344	29
Shanxi	-1.34536	0.33120	0.06196	-0.53677	30
Guangxi	-0.46437	0.20521	-2.52382	-0.68283	31

ISSN: 2688-8653

The explanatory aspects of each factor can be determined from the factor loading. The factor loads of main factor F1 in X1 (food), X3 (living), X4 (daily necessities and services), X5(transportation and communication), and X8 (other consumption) are large. The factor load of main factor F2 in X6 (education, culture, and entertainment) and X7 (medical care) was larger. The factor load of main factor F3 on X2 (clothing) is larger.

The main factor F1 reflects People's Daily consumption, which can be named as the daily consumption factor. The higher the score on this factor, the higher the daily consumption. The main factor F2 mainly reflects hedonic consumption, which can be named the hedonic factor. The higher the score on this factor, the more the expenditure on improving the quality of life. The main factor F3 mainly reflects the consumption of clothing. Although clothing is a necessity, it is not a daily consumer good, so it can be defined as the clothing consumption factor.

Further, according to the factor load coefficient obtained above, the factor load coefficient of each province is weighted and summed to obtain the comprehensive score of each province.

Draw factor score chart, can have more intuitive understanding to each province's consumption situation.

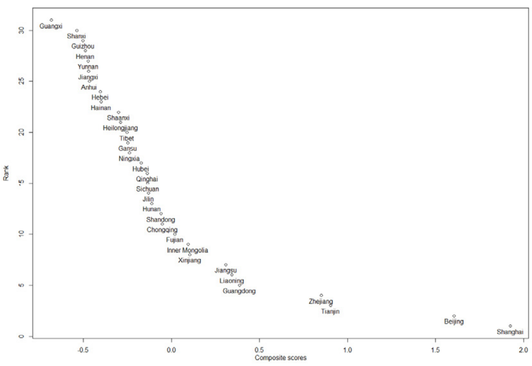

Figure 5. Factor score diagram

Eduction in various provinces and cities of Shanghai, the first one is the main factor in Shanghai F1 scoring well ahead of other provinces and cities, it shows that urban residents in Shanghai is the need to spend a lot of in the daily expenses, moreover also scored in Shanghai in the main factor F2 only slightly lower than Beijing, this shows that people in Shanghai also spend more on luxury consumption, and the main factor on the F3 score is very low, It shows that shanghainese spend little on clothing. In the cluster analysis, Beijing and Shanghai were categorized as spending more on clothing, and Beijing scored higher on hedonic spending than Shanghai.

In the second category, Tianjin has the highest overall score, followed by Zhejiang and Guangdong, which are economically developed provinces and cities. Guangdong's score on the main factor F1 is very high, ranking the second place among 31 provinces and autonomous regions, indicating that guangdong's daily consumption is at a high level. However, guangdong's score on the main factor F2 and F3 are both very low, and the score on F3 is only higher than that of neighboring Guangxi province, indicating that Guangdong people spend very little on clothing. This is closely related to the weather and climate of Guangdong and the consumption concept of Guangdong people. There is no particularly cold time in Guangdong all year round, so people spend little on clothes.

ISSN: 2688-8653

Among the third category of provinces and cities, The score of Tibet in factor F1 is close to That of Tianjin, but the score in factor F2 is indeed the lowest, which indicates that Tibet is in short of both medical resources and education resources, while the daily expenses are not low.

The distribution of other central and western provinces in the factor score chart is relatively concentrated, and the consumption structure is similar.

6. Conclusions and Suggestions

This paper analyzes the per capita consumption data of urban residents in 31 provinces, municipalities and autonomous regions in China in 2003, 2008, 2013 and 2018, hoping to have a better grasp of the change of consumption structure of Urban residents in China in recent years.

Based on the analysis of the article, it should be confirmed that the consumption level of urban residents in China has been greatly improved in the past two decades, which reflects the rapid development of per capita income. However, the gap in consumption level between different regions is widening. Housing prices and prices in Beijing, Shanghai, Guangzhou and economically developed coastal provinces are often higher than those in inland provinces, which makes a large proportion of per capita consumption still belong to subsistence consumption. However, these regions are rich in medical and educational resources, which can satisfy developmental consumption and hedonic consumption to a certain extent. In contrast, medical and education resources are still scarce in the northwest border region and some inland provinces, most notably In Tibet. In the factor analysis, Tibet has the lowest score in the hedonic consumption of the main factor F2, on the one hand, it is due to the underdeveloped economy and low per capita income of Tibet. On the other hand, compared with other provinces, the lack of medical and education resources in Tibet also leads to its lower score in F2.

The relationship between consumption structure and economic growth is often regarded as mutual causation, that is, the upgrading of residents' consumption structure will promote economic growth, and economic growth can increase per capita income, thus driving the growth of residents' consumption will achieve consumption upgrading. According to the analysis of four annual consumption data, the Engel coefficient of urban residents in China is declining, which indicates that the consumption structure of residents is in the process of continuous upgrading. This may be due to the "crowding out effect" of residential consumption on development-oriented and enjoyment-oriented consumption caused by the high housing price, that is, residents bear great pressure on mortgage loans, so they have no more consumption for their own development and enjoyment.

In order to better promote consumption upgrade, the most important point is to raise income. The low-income group should be gradually transferred to the middle-income group, so as to release the potential consumption demand. The middle-income group is also a relatively stable class in the society. The expansion of the middle-income group is of great significance for narrowing the gap between the rich and the poor and expanding domestic demand. In addition, all provinces and cities should strengthen supply-side structural reform. Developed provinces have abundant educational resources, but they need to curb the rise of housing prices so that residents can spend more on development and enjoyment, while, the central and western provinces need to strengthen their own medical education resources on the basis of increasing residents' income.

ISSN: 2688-8653

References

[1] J.M. Xue and M. Jin (2019). Consumption upgrading and high-quality economic development: An empirical study based on China's inter-provincial panel data. Business economics research, vol.22, no.1, p.42-46.

[2] T. Qiu (2019). Analysis on consumption level and consumption structure of urban residents in Mianyang city. The Chinese market, vol.31, no.1, p.20-23.

[3] R.Y. Sun (2018). Factor Analysis of the change of consumption structure of urban residents in Anhui Province. Economic Forum, vol.04, no.1, p. 27-30.

[4] X. Zhang (2021). Comparison of regional differences in urban and rural consumption structure in Jiangsu Province. Operation and Management, vol.03, no.1, p. 188-192.

[5] Y.B. Zhang, F. Zhang and X.X. Li (2021). Clustering analysis in the application of evaluation index analysis of China's urban residents consumption. Modern marketing (school edition), vol.08, no.1, p.12-14.

[6] X.Yang, F.G. Tong and F. Wang (2018). Town residents' consumption structure of henan province study. Journal of north China water conservancy and hydropower university (social science edition), vol.5, no.1, p.46-50.

[7] Z.M. Gong and H.Y. Li (2019). Research on regional consumption structure differences in China since the 21st century. Journal of shaoyang university (natural science edition), vol.16, no.1, p.15-16.

[8] J.K. Yang (2021). Research on the Structure of Chinese residents' consumption expenditure: Based on Factor Analysis and Cluster Analysis. Modern Commerce, vol.14, no.1, p.7-9.

[9] S.W. Shen and W. Zou (2021). An empirical study on the change of rural residents' consumption structure: Based on the Data of Chongqing Municipality from 2004 to 2018. Journal of Shandong Agricultural University (Social Science Edition), vol.23, no.3, p.10-14.

[10] B.H. Wang (2016). Multivariate statistical analysis and R modeling. Guangzhou: Jinan University Press.

[11] H.B. Wang (2018). R statistical analysis and application. Beijing: Posts and Telecommunications Press.