ISSN: 2688-9323

An Empirical Study and Policy Recommendations on the Quality of Life of Residents based on Statistical Models

Lejun Zhu

School of Economics, Hebei GEO University, Shijiazhuang Hebei 050031, China

Abstract

As a major driving force of the country's economic development, consumption is the most direct way to reflect people's aspirations for a better life. This paper selects the data of urban residents of Chinese provinces, cities, and autonomous regions in China on the consumption status of various aspects of life in 2020, and uses the statistical analysis method to conduct a study. There are eight research indicators, and the per capita consumption levels of these provinces, cities and autonomous regions aranalyzeded and evaluated through SPSS 21, and based on the obtained factor scores and comprehensive scores a heat map is drawn using R language to determine the regional boards The study was carried out by using R language to draw a heat map based on the obtained factor scores and composite scores to determine the links and differences between the scores of each regional segment.

Keywords

Empirical Analysis; Factor Analysis; R; Heat Map.

1. Introduction

As a major driving force of the country's economic development, consumption is the most direct way to reflect people's aspirations for a better life. As residents' income steadily rises and regional ties become closer, residents' consumption capacity is gradually enhanced and people are getting more and more well-off. China's vast territory, the pace of development and construction of provinces in different regions is quite different, and the spending concerns of different provinces are also great differences, the country has implemented several policies such as the revitalization of the old industrial base in the northeast and the western development strategy for the coordinated development of each region. This paper uses factor analysis to study the consumption behavior of urban residents in each province, explore the influence of geographical location and the degree of urban development on consumption expenditure, and provide a useful reference for government decision-makers in the formulation of policies such as investment attraction.

2. Literature Review

In recent years, with the widespread use of multivariate statistical tools, a large number of scholars at home and abroad have conducted studies on the structure of urban consumer spending in China. Clements et al (2006) studied the consumption and income data of residents in various countries around the world, analyzed a wide range of consumption patterns of residents, and classified the consumption patterns of different countries [2]. Among domestic scholars, Zeng Guang [3] (2012) identified the main determinants of consumption levels in various regions in terms of two types of factors, namely economic development and climate, and made corresponding comments; Ji Rongfang [4] (2007) analyzed the changing trends of the consumption structure of residents in Tai'an City, taking the data on residents' consumption expenditure as an example, and gave corresponding evaluations from the findings; Ke Jian [5]

(2004) used a combination of cluster analysis and factor analysis to analyze the trends of the consumption structure of residents in Tai'an City. a combination of cluster analysis and factor analysis to analyze the consumption structure of each province and put forward suggestions for guiding consumption and effectively initiating it. Ge Nan, Zhang Kexin & Hao Zixu's[7] analysis of the effects of GDP, urban disposable income and urban consumption expenditure using time series econometric methods. Huang Chunxia & He Chen[8] analyses and examines the differences and convergence of consumption levels of urban residents in China by combining the relevant statistical indicators and measurement methods of economic convergence studies and consumption studies. Peng C.S. & Yang L.Y.[9] Based on the panel data of 41 cities in the Yangtze River Delta region from 2010 to 2019, a fixed-effects model was used to analyse the impact of house prices on urban residents' consumption. The Yangtze River Delta region was classified into high-income cities and low-income cities by the K-means method. Zhang Manlin[10] further explored the main factors influencing consumption power, mainly through correlation analysis, linear regression fitting and principal component analysis of the data.Hua Lina[11] have analysed the main factors affecting urban residents' consumption expenditure with the help of multiple linear regression.

All the above-mentioned articles took cross-sectional data directly for factor analysis. Based on the previous studies, this paper first uses correlation tests to analyze the feasibility, making the experimental results more scientific, and uses heat maps to visualize the research results and enhance the comprehensibility of the study.

3. Data Sources and Selection of Indicators

The data selected for this paper are from the China Statistical Yearbook 2021[6] on per capita consumption expenditure of urban residents by region, including 31 provinces, municipalities, and autonomous regions in eight areas, namely food, tobacco and alcohol (x_1), clothing goods (x_2), housing (x_3), household goods and services (x_4), transportation and communications (x_5), entertainment, education and cultural services (x_6), health care (x_7) and other miscellaneous items (x_8), the indicators were selected with reference to the China Statistical Yearbook 2020 [6]. As shown in Table 1.

Table 1. Raw data

Region	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
Beijing	8488.5	2229.5	15751.4	2387.3	4979.0	4510.9	3739.7	1151.9
Tianjin	8983.7	1999.5	6946.1	1956.7	4236.4	3584.4	2991.9	1154.9
Hebei	4675.7	1304.8	4301.6	1170.4	2415.7	1984.1	1699.0	435.8
Shanxi	3997.2	1289.9	3331.6	910.7	1979.7	2136.2	1820.7	396.5
Inner Mongolia	5517.3	1765.4	3943.7	1185.8	3218.4	2407.7	2108.0	597.1
Liaoning	5956.6	1586.1	4417.0	1275.3	2848.5	2929.3	2434.2	756.0
Jilin	4675.4	1406.8	3351.5	948.3	2518.1	2436.6	2174.0	564.7
Heilongjiang	4781.1	1437.6	3314.2	844.8	2317.4	2444.9	2457.1	514.4
Shanghai	10952.6	2071.8	15046.4	2122.8	5355.7	5495.1	3204.8	1355.9
Jiangsu	6847.0	1573.4	7247.3	1496.4	3732.2	2946.4	2166.5	688.1
Zhejiang	8928.9	1877.1	8403.2	1715.9	4552.8	3624.0	2122.6	801.3
Anhui	6080.8	1300.6	4281.3	1154.3	2286.6	2132.8	1489.9	411.2
Fujian	8095.6	1319.6	6974.9	1269.7	3019.4	2509.0	1506.8	619.3
Jiangxi	5215.2	1077.6	4398.8	1128.6	2104.3	2094.2	1264.5	367.3

ISSN: 2688-9323

Shandong	5416.8	1443.1	4370.1	1538.9	2991.5	2409.7	1816.5	440.8
Henan	4186.8	1226.5	3723.1	1101.5	1976.0	2016.8	1746.1	354.9
Hubei	5946.8	1422.4	4769.1	1418.5	2822.2	2459.6	2230.9	497.5
Hunan	5771.0	1262.2	4306.1	1226.2	2538.5	3017.4	1961.6	395.8
Guangdong	9369.2	1192.2	7329.1	1560.2	3833.6	3244.4	1770.4	695.5
Guangxi	5031.2	648.0	3493.2	944.1	2384.7	2007.0	1616.0	294.2
Hainan	7122.3	697.7	4110.4	932.7	2578.2	2413.4	1294.0	406.2
Chongqing	6666.7	1491.9	3851.2	1392.5	2632.8	2312.2	1925.4	501.3
Sichuan	6466.8	1213.0	3678.8	1201.3	2576.4	1813.5	1934.9	453.7
Guizhou	4110.2	984.0	2941.7	873.8	2405.6	1865.6	1274.8	324.3
Yunnan	4558.4	822.7	3370.6	926.6	2439.0	1950.0	1401.4	311.2
Tibet	4792.5	1446.3	2320.6	847.7	2015.2	690.3	519.2	397.4
Shaanxi	4671.9	1227.5	3625.3	1151.1	2154.8	2243.4	1977.4	413.3
Gansu	4574.0	1125.3	3440.4	945.3	1972.7	1843.5	1619.3	358.6
Qinghai	5130.9	1359.8	3304.0	953.2	2587.6	1731.8	1995.6	481.8
Ningxia	4605.2	1476.6	3245.1	1144.5	3018.1	2352.4	1929.3	525.5
Xinjiang	5042.7	1472.1	3270.9	1159.5	2408.1	1876.1	1725.4	441.7

From the data in the above graph, the extreme differences under each indicator for different provinces can be obtained as shown in Table 2 below

Table 2. Extreme deviations for each indicator

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
Range	6955.4	1581.5	13430.8	1542.5	3383	4804.8	3220.5	1061.7

As can be seen from the graph above, the same indicator varies widely between regions, i.e. there is a large variability in consumption expenditure among urban residents in China, indicating the relevance of the project's research.

4. Research Design

4.1. Validity Testing

Whether the KMO test and Bartlett's spherical test pass or fail is a prerequisite for the factor analysis method. When the author used SPSS 21 to perform the two tests mentioned above, he obtained the results in Table 3 below.

Table 3. Results of the KMO test and the Bartlett spherical test

KMO and Bartlett's test		
The Kaiser-Meyer-Olkin metric of sampling adequacy.		0.858
Bartlett's test for sphericity	Approximate cardinality	335.070
	df	28
	Sig.	0

From the KMO test, KMO>0.8; Bartlett's test value, sig. <0.05, according to the KMO taken to select the original indicators for factor analysis of the judgment criteria, can be concluded that

ISSN: 2688-9323

suitable for factor analysis. The correlation coefficients were then tested and the correlation coefficients were tabulated in Table 4 below.

Table 4. Table of correlation coefficients

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
x_1	1.000	.545	.816	.816	.873	.813	.535	.827
x_2	.545	1.000	.668	.773	.741	.659	.754	.832
x_3	.816	.668	1.000	.893	.896	.900	.726	.868
x_4	.816	.773	.893	1.000	.905	.861	.768	.874
x_5	.873	.741	.896	.905	1.000	.899	.734	.919
x_6	.813	.659	.900	.861	.899	1.000	.822	.888
x_7	.535	.754	.726	.768	.734	.822	1.000	.820
x_8	.827	.832	.868	.874	.919	.888	.820	1.000

Based on the above table it can be seen that the correlation coefficients are all above 0.3, making them suitable for factor analysis. And from the obtained common factor variance table it can be seen that the variables have a minimum common degree of 0.726, which is high overall, indicating that the original variables have less information missing and almost all are included in the common factor, which is suitable for factor analysis, and the results are shown in Table 5 below.

Table 5. Table of common factor variance

Common factor variance		
	Initial	Extraction
x1	1.000	.738
x2	1.000	.667
x3	1.000	.871
x4	1.000	.899
x5	1.000	.921
x6	1.000	.889
x7	1.000	.713
x8	1.000	.934
Extraction method: Principal component analysis.		

4.2. Factor Extraction Process

After the factor analysis was known to be applicable to the data used in this study, the author conducted a principal component analysis for the contribution of each component variance and the results are shown in Table 6 below.

As can be seen from the above figure, the initial eigenvalues of the first two factors are greater than 1,which indicates that the first two factors can cover most of the information of the original data, which means that the extracted two factors can make a more adequate explanation of the original eight indicators. The evaluation system is more concise and clear after factor extraction.

Table 6. Table of variance contribution of each component

Total variance explained									
Ingredients	Initial Eigenvalue			Extraction of squares and loading			Rotate square and load		
	Total	% of variance	Cumulative %	Total	% of variance	Cumulative %	Total	% of variance	Cumulative %
1	6.630	82.875	82.875	6.630	82.875	82.875	4.074	52.630	52.630
2	.616	7.697	90.572	.616	7.697	90.572	3.171	35.427	88.057
3	.320	4.005	94.576						
4	.161	2.014	96.590						
5	.109	1.360	97.950						
6	.073	.914	98.864						
7	.057	.715	99.580						
8	.034	.420	100.00						
Extraction method: Principal component analysis.									

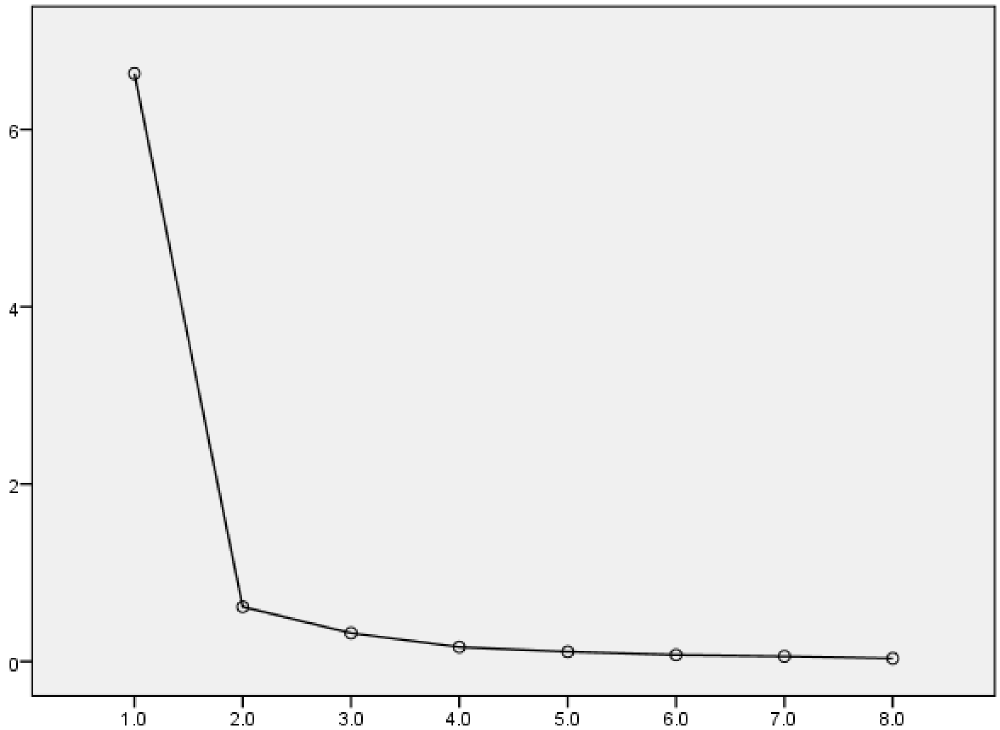

Figure 1. Gravel map

4.3. Factor Loading Analysis and Common Factor Naming

In this paper, the maximum variance method was chosen for the orthogonal rotation, and after processing using SPSS 21 software, the rotation component matrix was obtained as shown in Table 7 below.

As can be seen, Component 1 has a greater correlation with the five indicators of expenditure on food, tobacco and alcohol, housing, transportation and communication, education, culture and entertainment, and miscellaneous goods and services, which generally do not change significantly throughout the year, but are closely related to the geographical environment in which residents live, and are therefore named the geographical factorF_1 ; Component 2 has a greater correlation with the three indicators of clothing, household goods and services, and health care, which are closely related to These indicators are closely linked to seasonal factors and climatic factors in each region, and are therefore named climate-based factorsF_2 , and are briefly summarised in Table 8.

ISSN: 2688-9323

Table 7. Matrix of rotated components

Rotating component matrix[a]		
	Ingredients	
	1	2
x1	.939	.225
x2	.324	.875
x3	.812	.487
x4	.737	.597
x5	.814	.525
x6	.772	.548
x7	.378	.855
x8	.695	.674
Extraction method :Main ingredient. Rotation method :Orthogonal rotation method with Kaiser standardisation.		
a. Rotation converges after 3 iterations.		

Table 8. Explanatory table of indicators and meanings of the common factor loadings

Common factor	F_1	F_2
Load indicators	Food, tobacco and alcohol, housing, transport and communications, education, culture and entertainment, miscellaneous goods and services expenses	Clothing, household goods and services, health care
Common factor meaning	Geographical factors	Climatic factors

4.4. Calculation of Common Factor Score and Corresponding Ranking

In order to explore the characteristics of the structure of consumer spending, this paper compares each region by calculating the score of the common factor, firstly, the score coefficient matrix of the common factor to derive the score function of each common factor, the score coefficient matrix is shown in Table 9 below.

Table 9. Matrix of component score coefficients

Component score coefficient matrix		
	Ingredients	
	1	2
x1	.566	-.459
x2	-.385	.637
x3	.276	-.105
x4	.138	.059
x5	.250	-.069
x6	.200	-.015
x7	-.329	.578
x8	.049	.167

A common factor score function can be derived from this.The weightings are then aggregated using the weight of the variance contribution of each factor to the total variance contribution of the two factors to arrive at a composite score for each region.

ISSN: 2688-9323

The corresponding calculations were completed and ranked with the help of the statistical software SPSS21, as shown in Table 10 below.

Table 10. Table of scores for each common factor, overall scores and corresponding rankings

Region	SF_1 Corresponding rankings	SF_2 Corresponding rankings	SF rankings
Beijing	4	1	2
Tianjin	7	19	6
Hebei	20	14	20
Shanxi	30	4	16
Inner Mongolia	21	5	12
Liaoning	18	7	11
Jilin	29	6	14
Heilongjiang	31	3	15
Shanghai	1	10	1
Jiangsu	8	11	7
Zhejiang	3	23	4
Anhui	11	27	27
Fujian	5	26	9
Jiangxi	10	15	10
Shandong	16	18	13
Henan	28	9	23
Hubei	17	20	17
Hunan	12	29	29
Guangdong	2	30	5
Guangxi	9	31	31
Hainan	6	17	3
Chongqing	14	21	19
Sichuan	13	25	25
Guizhou	19	28	30
Yunnan	15	24	26
Tibet	25	16	24
Shaanxi	23	22	28
Gansu	22	12	21
Qinghai	27	8	18
Ningxia	24	13	22
Xinjiang	26	1	8

From the data shown in the table above, separate blocked regional heat maps were created for each metric factor, showing the heat effects of different provinces on the map. This is represented in Figures 2 and 3 below, where the darker the blue the higher the regional factor score and vice versa the lower. The distribution of colours can be used to determine the links and differences between the regional blocks. The combined scores are then plotted against the blocked regional heat maps, as shown in Figure 4 below.

ISSN: 2688-9323

Figure 2. First common factor score plotting heat map

As can be seen from the chart above, with the exception of Beijing and Tianjin, the scores are roughly decreasing from the more rapidly developing economic regions along the eastern coast to the northern inland regions, with provinces in the northeast such as Jilin and Heilongjiang, and central regions such as Shanxi and Henan scoring lower on the first common factor, and the rest of the wider region scoring moderately.

Figure 3. Second common factor score for plotting the heat map

This is in line with the climate-based meaning of the second common factor, as the colder and drier climate in northern China means that people spend more on clothing, household goods and services, and healthcare. However, in the east-west direction, with the exception of Xinjiang, the eastern region scores slightly higher than the western region. This paper suggests that this is due to the influence of clothing trends, which make residents in the eastern regions with higher economic levels focus more on fashionable clothing, rather than all on keeping warm.

ISSN: 2688-9323

5. Conclusion and Policy Advice

5.1. Conclusion

This paper has used the method of factor analysis to explore the structural characteristics of China's per capita consumption expenditure levels in 31 provinces, cities and autonomous regions in 2020, and it can be seen that the main factor determining the per capita income of residents is still mainly the income level, but in addition to this different geographical and climatic environments also have a great influence on consumption levels. In this paper, using the maximum variance method for rotation, two common factors were extracted, which were named as geographic-type factor and climate-type factor, and the following conclusions were obtained.

1.With the exception of Beijing and Tianjin, the first public factor scores are generally decreasing from the more rapidly developing eastern coastal regions to the northern inland regions, i.e. for the more rapidly developing regions there is a higher proportion of expenditure on food, tobacco and alcohol, housing, transport and communications, education, culture and entertainment, and miscellaneous goods and services.

2.The distribution of scores for the second common factor is generally decreasing from the north to the south, with residents in the north spending more on clothing, household goods and services, and health care. However, due to clothing trends, residents of the more economically advanced regions in the east are more interested in fashionable clothing rather than warmth.

3.Beijing, as the capital city, scores high on both public factors, while Tianjin, as a neighbour of the capital city Beijing and in the same economic circle, scores much higher on the public factor and the overall score than the neighbouring regions of Hebei, Shanxi and Shandong, which are not part of the capital city's economic circle. This shows that the economic circle plays an extremely important role in the development of the city.

5.2. Recommendations

In formulating policies for different regions, policy makers should decide on the introduction of relevant policies according to the characteristics of each region, for the central region, the "halo effect" of large cities should drive the economic development of the surrounding underdeveloped regions, forming a perfect For the eastern coastal regions, the government should pay more attention to the provision of higher-level consumer goods such as education, culture and entertainment, and guide the development of popular industries so as to expand the scale of consumption in the region; for the western regions, the government should invest more in the folklore and cultural tourism industries and industries with local characteristics, so as to raise the income level of local residents and stimulate the general consumption demand in less economically developed regions. For the western region, the government should invest more in the folklore and cultural tourism industries and industries with local characteristics to raise the income levels of local residents and stimulate general consumption demand in less developed regions.

References

[1] Karl E. Case, John M. Quigley, Robert J. Shiller. Comparing Wealth Effects: The Stock Market versus the Housing Market[J]. Advances in Macroeconomics,2005(1):1235-1253.

[2] Kenneth W. Clements,Yanrui Wu,Jing Zhang.Comparing international consumption patterns[J]. Empirical Economics,2006,31(1).

[3] National Bureau of Statistics. China Statistical Yearbook 2021 [M]. Beijing: China Statistics Press, 2021.

ISSN: 2688-9323

[4] Zeng Guang, Research on the structure of urban residents' consumption expenditure in China: a factor analysis based on inter-provincial cross-sectional data in 2010[J]. Journal of Jinan (Philosophy and Social Science Edition), 2012(09).

[5] Ji Rongfang, Factor analysis model and application of consumption structure of urban residents[J]. Journal of Shandong Agricultural University (Natural Science Edition), 2007(02).

[6] Ke Jian, A comparative study on the consumption structure of urban residents in China by region[J]. Exploring Economic Issues, 2004(10).

[7] Ge Nan, Zhang Kexin & Hao Zixu, Factors influencing consumption expenditure of urban residents in China. Wealth Times, 2020(07): 131-132.

[8] Huang Chunxia & He Chen, Analysis of consumption differences and convergence among urban residents in China[J]. Business Economics Research, 2019(20): 72-75.

[9] Peng C.S. & Yang L.Y., Research on the impact of housing price fluctuations on urban residents' consumption - based on the Yangtze River Delta region[J]. Journal of Heilongjiang Institute of Technology (Comprehensive Edition), 2022. 22(02): 97-102.

[10] Zhang Manlin, Analysis of factors influencing urban residents' consumption expenditure - based on SPSS[J]. Time Finance, 2019(17): 93-94.

[11] Hua Lina, Statistical analysis of urban residents' consumption expenditure in China[J]. Guangxi Quality Supervision Herald, 2019(05): 227-228.