Product competitiveness analysis for e-commerce platform of special agricultural products

Fucheng Wan\textsuperscript{1}, Ning Ma\textsuperscript{1}, Dongwei Yang\textsuperscript{2}, Zhangyuan Xiong\textsuperscript{1}

\textsuperscript{1}Northwest University for Nationalities, National languages Information Technology research institute, Lanzhou city, Gansu province, China
\textsuperscript{2}Northwest University for Nationalities, School of Journalism and Communication, Lanzhou city, Gansu province, China

Abstract. On the basis of analyzing the influence factors of the product competitiveness of the e-commerce platform of the special agricultural products and the characteristics of the analytical methods for the competitiveness of the special agricultural products, the price, the sales volume, the postage included service, the store reputation, the popularity, etc. were selected in this paper as the dimensionality for analyzing the competitiveness of the agricultural products, and the principal component factor analysis was taken as the competitiveness analysis method. Specifically, the web crawler was adopted to capture the information of various special agricultural products in the e-commerce platform \texttt{chi.taobao.com}. Then, the original data captured thereby were preprocessed and MYSQL database was adopted to establish the information library for the special agricultural products. Then, the principal component factor analysis method was adopted to establish the analysis model for the competitiveness of the special agricultural products, and SPSS was adopted in the principal component factor analysis process to obtain the competitiveness evaluation factor system (support degree factor, price factor, service factor and evaluation factor) of the special agricultural products. Then, the linear regression method was adopted to establish the competitiveness index equation of the special agricultural products for estimating the competitiveness of the special agricultural products.

1 Introduction

In recent several years, in order to further stimulate domestic economic development, the government proposed the slogan of “Internet +”, “Mass Innovation” and “Mass Entrepreneurship”. In fact, the mode of “Internet + Traditional Agriculture” has become an important method for converting present agricultural development mode and promoting the efficient allocation of market resources. Under such situation, the e-commerce of agricultural products has accordingly become a new growth point of the “Agricultural Economy (related to agriculture, farmer and rural area)” in China. At present, the two major e-commerce leaders in China ----- Jingdong (“Fresh and Specialty Sectors”) and Taobao (“\texttt{chi.taobao.com}”) are energetically promoting the agricultural products on e-commerce platform; meanwhile, some particular e-commerce platforms for agricultural products, such as \texttt{www.womai.com}, \texttt{www.tootoo.cn}, \texttt{www.sfbest.com}, \texttt{www.benlai.com}, \texttt{www.taocz.com}, \texttt{www.fieldschina.com}, Long Bao Tracing Mall and \texttt{www.cndlbz.com}, are widely established. In order to further promote the development of the e-commerce platforms particularly for special agricultural products, this paper is mainly focused on researching the competitiveness analysis system on the e-commerce platform of the special agricultural products in order to provide relevant services to the e-commerce platform of the special agricultural products.
2 Data Source
In consideration of the prematuration of most e-commerce platforms of the special agricultural products in China, a small amount of samples for merchants, product sales volume and other relevant data are insufficient for the product competitiveness analysis. In this paper, the research data are sourced from the commodity data of the special agricultural products in “chi.taobao.com”. Specifically, the web crawler was adopted to capture the commodity attribution data in the product details webpage. The commodity attribution data for over 1,000 agricultural products of 7 different product types have been captured for research, including leisure foods, biscuits & pastries, fresh fruits & vegetables, oils & dried foods, tea & drink, traditional tonics, domestic and foreign famous wines.

3 Principal Component Factor Analysis Method ---- Theoretical Analysis

3.1 Introduction to principal component factor analysis method
In practical problem research, problem variables may have certain correlation, namely: the information described by the variables may be overlapped with each other \(^{[17]}\). In order to solve variable correlation problem, it is necessary to adopt less variables to replace original variables for variable conversion under the precondition of not losing or only losing very few original variable information. Therefore, the factor analysis method is adopted to solve such problem.

The factor analysis method is a dimension reduction method for the problem describing variables. In this method, the internal correlation of the variables is mathematically analyzed and only several abstract variables (interpretable) are adopted to describe all or main information of original variables, wherein such interpretable abstract variables are called as factors in the factor analysis method. Original variables can be directly obtained from the observed values of the experiment samples, and the abstract variables are interpretable in practical problems.

In the factor analysis method, the most universal factor extraction method is the principal factor analysis method. Specifically, the original variables of the samples are standardized to convert the original variable group into one group of irrelevant factor variables; then, the principal factors are selected according to the completion degree of the factor group for describing the variable information to reduce the dimension of the original variables while not losing the information of the original variables.

3.2 Characteristics of principal component factor analysis method
(1) Original variables are correlated to each other, but the factor variables are not significantly correlated to each other, thus favorable for deeply analyzing the problem;
(2) The number of the factor variables must be less than that of the original variables in order to simplify problem analysis.
(3) The factor variables must be able to describe most information of the original variables in order to maximally reduce original information loss and improve problem analysis accuracy.
(4) The factor variables must be interpretable, namely: the factor variables can describe a certain aspect of the practical problem.

4 Product Competitiveness Analysis of E-commerce Platform of Special Agricultural Products

4.1 Model analysis
p original variables, \(x_1 x_2 \ldots x_p\) (namely, p influence factors of the observed data of each special agricultural product), are set in this paper, and each original variable includes m aspects of the commodity competitiveness, recorded as \(F_1 F_2 \ldots F_m\) \((m < p)\), wherein m aspects are called as the common factors of p original variables, and the aspects that cannot be interpreted by m common factors are called as the special factors of the original variables, recorded as \(\varepsilon_1 \varepsilon_2 \ldots \varepsilon_p\). Therefore, the factor analysis model is established as follows:
\[
\begin{align*}
x_1 &= a_{11}F_1 + a_{12}F_2 + \cdots + a_{1m}F_m + \varepsilon_1 \\
x_2 &= a_{21}F_1 + a_{22}F_2 + \cdots + a_{2m}F_m + \varepsilon_2 \\
\vdots & \quad \vdots \\
x_p &= a_{p1}F_1 + a_{p2}F_2 + \cdots + a_{pm}F_m + \varepsilon_p 
\end{align*}
\] (formula 4-1)

Where \( a_{ij} \) \((i = 1,2,\ldots,m; j = 1,2,\ldots,p)\) is the factor load, called as the load of the \( i \)th variable on the \( j \)th common factor, namely, the degree of the \( i \)th variable for interpreting the \( j \)th aspect of the commodity competitiveness.

The matrix expression of this model is as follows:
\[
X = AF + \varepsilon 
\] (Formula 4-2)

Description:
\[
X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix}, \quad A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pm} \end{bmatrix}, \quad F = \begin{bmatrix} F_1 \\ F_2 \\ \vdots \\ F_m \end{bmatrix}, \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_p \end{bmatrix}
\]

The following conditions should be met:

1. \( m \leq p \), namely: the number of the common factors must be less than that of the influence factors in original data;

2. \( \text{Cov}(F, \varepsilon) = 0 \), namely: the correlation between the common factors and the special factors must be 0;

3. \( D(F) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_m \), namely: all common factors are not correlated to each other, and the variance is 1.

4. \( D(\varepsilon) = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \\ \vdots & \vdots \\ 0 & \sigma_p^2 \end{bmatrix} \), namely: all special factors are not correlated to each other, but the variance may not be the same.

### 4.2 Process analysis
The product competitiveness analysis process of the e-commerce platform of the special agricultural products is as shown in Fig.4.1:
Fig. 4.1 Product Competitiveness Analysis Process of E-commerce Platform of Special Agricultural Products

1. Standardize original data to obtain standard matrix X: various data in original variables are not uniformly measured, so it is necessary to standardize the original data to convert them into standard matrix in order to reduce analysis difficulty and implement uniform data calculation;

2. Calculate correlation coefficient matrix R according to the standard matrix: set $|R-\lambda E|=0$, calculate the characteristic value, the contribution rate and the accumulative contribution rate of $R$, and determine the number of the common factors according to the principle that the characteristic value is not less than 1 or the accumulative variance contribution rate is not less than 85%.

3. Calculate the characteristic vector and initial factor load matrix A;

4. Estimate factor scores by regression method, and take the specific value of the variance contribution rate of each factor and the total variance contribution rate of the factors as the weight value for the weighted summary in order to obtain the factor analysis model;

5. In case of unobvious factor meanings, adopt the maximum variance method for the orthogonal rotation of the initial factors in order to obtain rotated common factor solution $B$;

6. Comprehensively analyze and evaluate the factors according to the sample factor scores.

4.3 Empirical analysis

The factor analysis function provided by SPSS tool was adopted and 6 data indexes ---- commodity price, commodity sales volume, user evaluation, postage included service, store reputation and commodity collection quantity were taken as the original variable. Then, the principal component factor analysis method was adopted to convert the original variables and calculate the factor load matrix. The common factors obtained at the first time were not obviously interpretable, so the
maximum variance method was adopted to rotate the load matrix. Finally, the regression method was adopted to calculate the influence factor scores.

According to the results of KMO and Bartlett sphericity tests (see Tab.4.1), KMO value is more than 0.6, thus indicating high variable intercommunity. Meanwhile, the significance level (0.00) of Bartlett sphericity test is less than 0.05 and the hypothesis of Bartlett sphericity test is refused, so the original data of the special agricultural product samples collected thereby are applicable to factor analysis.

| Tab.4.1 KMO and Bartlett Tests |
|--------------------------------|
| Kaiser-Meyer-Olkin Measurement for Sampling | 0.609 |
| Bartlett Sphericity Test | Approximate Chi-square | 114.642 |
| df | 15 |
| Sig. | 0.000 |

According to the common factor variance analysis (see Tab.4.2), the variance extraction rate of the collected commodities is 0.547 while those of other variables are mainly above 0.7. Obviously, the variable ---- commodity collection quantity has the largest information loss, and other variable information can be well interpreted.

| Tab.4.2 Common Factor Variance |
|--------------------------------|
| Commodity Price | Initial | 1.000 |
| | Extracted | 0.930 |
| Commodity Sales Volume | Initial | 1.000 |
| | Extracted | 0.699 |
| Store Level | Initial | 1.000 |
| | Extracted | 0.702 |
| Commodity Collection Quantity | Initial | 1.000 |
| | Extracted | 0.547 |
| Good Reputation Rate | Initial | 1.000 |
| | Extracted | 0.994 |
| Transportation Expense | Initial | 1.000 |
| | Extracted | 0.957 |

According to the total variance table (see Tab.4.3) interpreted by the experimental calculation result data of the principal factors extracted by the principal component factor analysis method, four principle component factors are successfully extracted and the accumulative contribution rate of the principal components reaches 80.469%, namely: 80.469% of the information in original variables are extracted. Therefore, it is indicated that each variable of original commodity information loses very few information, and the four common factors can well describe these indexes, and the factors are ideally extracted.

| Tab.4.3 Interpreted Total Variance |
|-----------------------------------|
| Component | Initial Characteristic Value | Extraction of Square and Load | Rotation of Square and Load |
| | Total | Variance(%) | Accumulation(%) | Total | Variance(%) | Accumulation(%) | Total | Variance(%) | Accumulation(%) |
| 1 | 1.861 | 31.010 | 31.010 | 1.861 | 31.010 | 31.010 | 1.775 | 29.583 | 29.583 |
| 2 | 1.255 | 20.922 | 51.932 | 1.255 | 20.922 | 51.932 | 1.040 | 17.335 | 46.918 |
According to the factor load matrix of the experiment result (see Tab.4.4), the interpretations of the four common factors in the variables are dispersed and have unobvious factor expressions, so the factors cannot be directly used for interpretation. Subsequently, the maximum variance method was adopted for the orthogonal rotation of the initial factors, without changing the accumulative contribution rate of the rotated principal components (see Tab.4.3). Therefore, the sample data information of the special agricultural products will not be further lost.

According to the experimental analysis result, the first common factor is the support degree factor, orderly followed by the price factor, the service factor and the evaluation factor. The degree of the influence of the factors on the competitiveness of the agricultural products is as follows: support degree > price > service > evaluation.
The rotation component matrix can reflect the proportion of each principal factor, and we can obtain the following principal factor model for the product competitiveness:

\[ x_1 = -0.039F_1 + 0.955F_2 + 0.955F_3 + 0.025F_4 + \varepsilon_1 \]
\[ x_2 = 0.758F_1 - 0.304F_2 + 0.176F_3 + 0.022F_4 + \varepsilon_2 \]
\[ x_3 = 0.830F_1 - 0.028F_2 + 0.104F_3 - 1.018F_4 + \varepsilon_3 \]
\[ x_4 = 0.707F_1 + 0.144F_2 - 0.130F_3 + 0.101F_4 + \varepsilon_4 \]
\[ x_5 = 0.063F_1 + 0.023F_2 + 0.077F_3 + 0.992F_4 + \varepsilon_5 \]
\[ x_6 = -0.073F_1 + 0.122F_2 + 0.965F_3 + 0.079F_4 + \varepsilon_6 \]

(1)

The original information of the special agricultural products mainly includes commodity price, commodity sales volume, user evaluation, postage included service, store reputation and commodity collection quantity. In this paper, the principal component factor analysis method was adopted to select four principal factors (support degree factor, price factor, service factor and evaluation factor) to describe the competitiveness of the special agricultural products. In this way, only a few abstract variables were adopted to describe all information in order to summarize the main influence factors of the competitiveness of the agricultural products.

| Component Score Coefficient Matrix |
|-----------------------------------|
| Component | 1     | 2     | 3     | 4     |
|-----------|-------|-------|-------|-------|
| Commodity Price | 0.084 | 0.939 | -0.017 | -0.044 |
| Commodity Sales Volume | 0.421 | -0.254 | 0.273 | -0.056 |
| Store Level | 0.479 | 0.077 | -0.050 | -0.077 |
| Commodity Collection Quantity | 0.412 | 0.234 | -0.129 | 0.052 |
| Good Reputation Rate | -0.047 | -0.038 | -0.071 | 1.009 |
| Transportation Expense | 0.023 | -0.035 | 0.972 | -0.065 |

Common factors (F1, F2, F3 and F4) are variables, but they are different from such original variables as commodity sales volume and commodity price and cannot be directly used as statistical information. However, the factor scores, namely the score of each sample on each common factor, can be calculated to measure the competitiveness of the commodity on the support degree factor, the price factor, the service factor and the evaluation factor. According to the component score coefficient matrix (see Tab.4.6) obtained by regression method, the regression equation of the common factor scores can be obtained as follows:

\[
\begin{align*}
F_1 &= 0.084x_1 + 0.421x_2 + 0.479x_3 + 0.412x_4 - 0.047x_5 + 0.023x_6 \\
F_2 &= 0.939x_1 - 0.254x_2 + 0.077x_3 + 0.234x_4 - 0.038x_5 + 0.035x_6 \\
F_3 &= -0.017x_1 + 0.273x_2 - 0.050x_3 - 0.129x_4 - 0.071x_5 + 0.972x_6 \\
F_4 &= -0.044x_1 - 0.056x_2 - 0.077x_3 + 0.052x_4 + 1.009x_5 - 0.065x_6 
\end{align*}
\]

(2)

The factor scores can be obtained according to the regression equation of the principal factors, and the weighted mean of the contribution rate of the common factors to original information is taken as the weight value to calculate the comprehensive score of the product competitiveness of individual sample data for the comprehensive evaluation of the competitiveness of the agricultural products.

The competitiveness of the special agricultural products is jointly influenced by four factors. In order to objectively measure the product competitiveness, the specific value of the variance contribution rate of each factor and the total variance contribution rate of the factors is taken as the weight value to calculate the comprehensive score as the product competitiveness index. The
comprehensive evaluation model for the competitiveness of the special agricultural products is established as follows:

\[ F = \frac{29.583\%}{80.469\%} F_1 + \frac{17.335\%}{80.469\%} F_2 + \frac{16.852\%}{80.469\%} F_3 + \frac{16.699\%}{80.469\%} F_4 \]

5 Conclusion

In this paper, the principal component factor analysis method was adopted to establish the product competitiveness analysis model for the e-commerce platform of the special agricultural products, and SPSS statistical software was adopted as the analysis tool to research and analyze product competitiveness. According to the research and analysis result, the competitiveness of the special agricultural products is mainly influenced by four factors, namely support degree factor, price factor, service factor and evaluation factor. Different products are differently influenced by the four factors, thus presenting the competitiveness difference of the special agricultural products on different factors. Meanwhile, the comprehensive evaluation model for the competitiveness of the special agricultural products was also established in this paper to analyze the competitiveness index of the products in the e-commerce platform of the special agricultural products so as to measure the competitiveness of the special agricultural products.

Acknowledgement

This research is supported by the Natural Science Foundation of Gansu Province (NO. 1610RJZA105).

References

[1] Liu L, Wu L, Ren W, et al. Analysis of Competitiveness of Beijing Agricultural Product Processing Industry[J]. Academic Periodical of Farm Products Processing, 2006.
[2] Shen A H. Analysis of Competitiveness of the Agricultural Processing Industry in Subei[J]. China Business & Market, 2007.
[3] Hao W, Wang H, Zhang P. Internal Competitive Competency Analysis of Regional Industries in Inner Mongolia[J]. Scientific Management Research, 2009.
[4] Wang Z, Yang Z. A Comparative Study of the Six Major Inner Mongolia Industries with Local Advantages and Characteristics on the Basis of Factor Analysis[J]. Journal of Inner Mongolia Finance & Economics College, 2010.
[5] Li W L, Zhong T M. Characteristics and Comparative Analysis of the Chaihe Volcanic Relics in Zalantun, Inner Mongolia[J]. Acta Geoscientica Sinica, 2013, 34(6):749-756.
[6] Bartuněk, J. Analysis of the competitiveness of the Czech wood-processing industry. [J]. Environmental Politics, 2000, 17(5):828-834.
[7] Takei F. Product competitiveness evaluation— quantitative analysis for development strategy[J]. Technological Forecasting & Social Change, 1985, 28(2):123-139.
[8] Prasad B. COMPETITIVENESS ANALYSIS OF EARLY PRODUCT INTRODUCTION AND TECHNOLOGY INSERTION[C]// ASME International Mechanical Engineering Congress and Exposition, Chicago, Illinois, November 6-11, 1994, Sposored by ASME Production Engineering Division. 1994.