An Analytical Research on Current Development Situations of Social Organization Standards of Guangdong Province on the Basis of Big Data Analysis Method

Tingjun Huang[1], Youhong Yang[1(✉)], Zhengyuan Wang[1], Jingyang Mao[2], and Hanying Zhi[3]

[1] School of Standardization, China Jiliang University, Hangzhou, Zhejiang, China
byangyouh@163.com
[2] Faculty of Science, China Jiliang University, Hangzhou, Zhejiang, China
[3] School of Economics and Management, China Jiliang University, Hangzhou, Zhejiang, China

Abstract. People from all walks of life are paying more and more attention to social organization standards, which play an important role in meeting market and innovation demands. This paper uses the clustering analysis method and the entropy weight method based on big data to analyze the current development situations of social organization standards. Research results show that the quantity of social organization standards for "manufacturing industry" issued in Guangdong Province occupies nearly half of the quantity of social organization standards issued in the whole province. The quantity of social organization standards for "manufacturing industry" is far higher than that of the "agriculture, forestry, animal husbandry and fishery" which ranks the second. As far as the development situations are concerned, top three industries in terms of the comprehensive score are "manufacturing industry", "wholesales and retail industry" and "leasing and business service industry" respectively while the industry with the lowest comprehensive score is "water conservancy, environment and public utilities management industry". This is because public infrastructure construction belongs to "public products" and its industry standards are mainly set by the government, so social organization standards organizations only play a small role there.

Keywords: Social organization standards of Guandong Province · clustering analysis · entropy weight method

1 Introduction

Scheme for Deepening Standardization Work Reform printed and issued by the State Council in 2015 formally put forward the concept of "social organization standards". In 2017, the newly revised *Standardization Law of the People's Republic of China*, established the legal status of social organization standards for the first time. Till now, social organization standards of China has developed for five years in China. Statistics from the

National Social Organization Standards Information Platforms show that there are more than new 4,000 social organization standards and 700 new social groups carrying out standardization work on average each year. Social organization standards have positive effects on standardizing product quality, ensuring product compatibility, promoting technological innovation, protecting consumer interests and promoting the modernization transformation of China's standard system (Liu Ke, 2019) [1]. At the same time, they can also greatly stimulate the vitality of market entities, meet standard demands of China's markets, provide sufficient impetus to China's economy in its improving quality and efficiency and promote the transformation of China's government functions (Wu Huixia, 2018) [2]. Zhejiang, Guangdong, Shandong and Shanghai have also carried out activities to improve quality of products and develop regional brands according to social organization standards (Xiao Yining, 2019; Ma Xiaoou, 2019; Alyson, 2018) [3–5]. Although the above scholars have made a comprehensive study on current development situations, challenges and development countermeasures of social organization standards of China, there is still a lack of studies on targeted analyses of current development situations of social organization standards in different provinces. Moreover, there are also few studies which systematically elaborate and analyze the development status of social organization standards in Guangdong Province. First, this paper uses big data analysis ideas, and collects relevant data based on the China Standards Information Service Network, the National Social Organization Standards Information Platform, the Guangdong Provincial Institute of Standardization, and China Standardization. Secondly, by digging and sorting out the social organization standard data of Guangdong Province, using the clustering method, a quantitative analysis of the construction and development of various industry social organization standards in Guangdong Province is carried out. Finally, this paper concludes current development situations of social organization standards in Guangdong province and related existing problems.

2 Research Design

2.1 Data Collection

This paper collects the data of social organization standards in Guangdong published on national social organization standards information platform from 2015 to 2020. The standardized policy text is obtained by referring to the policy document of Guangdong mentioned in *China Standardization* and searching the websites of governments at all levels of Guangdong. As of Nov 21, 2020, the platform has operated for more than four years. The statistics show that there are totally 589 social groups in Guangdong that register or publish on national social organization standards information platform. They have totally published or issued 2238 social organization standards. 348 social groups have published standards, which accounts for 59.08% of the number of all social groups.

2.2 Research Method

This paper uses clustering analysis method. This method classifies each sample to be researched, so that there will be a great similarity between the research samples in each

category. It is aimed to maximize the heterogeneity between each category and maximize the individual homogeneity in each category. With social group as the basic unit, this paper conducts statistics for 589 social groups in Guangdong registered on the platform one by one and carries out clustering analysis from three dimensions of "industrial distribution of national economy", "distribution of social groups" and "distribution of prefecture-level cities" with purpose of more intuitively analyzing the phenomenon behind the data.

Entropy weight method is a common objective evaluation method and is often used for weight defining in multi-object and multi-index objective and comprehensive evaluation. It is able to better avoid the influence of human subjective factors, so that the evaluation results can be more accurate. The way of defining the index weight with entropy weight method and evaluating the quality and status of the target object has been widely used in social management and other fields.

3   Analysis on Development of Social Organization Standards in Guangdong Province

When using entropy weight method, this paper reasonably screens and evaluates all optional indexes based on the method of selecting index researched by ZENG Yao (2016) [6], LIU Yuanyuan, et al. (2020) [7] and in combination with this research topic, and finally sets the indexes of using entropy weight method in this paper as the "number of corporate units", "annual GDP" and "the number of social organization standards". It evaluates the social organization standards development of each industry from the three dimensions and defines the weight according to the information provided by the data of each index. The bigger the index variation is, the information entropy is smaller, and the distributed weight is bigger. It determines the comprehensive score through the weight and finally judges the social organization standards development of the industry based on the comprehensive score. The evaluation process is as follows:

3.1   Index Selections

In order to evaluate the development status scientifically, this paper selects three representative indexes: "number of corporate units", "annual GDP" and "number of social organization standards" which can directly and effectively reflect the development scale and level of the industry. Therefore, this section collects information around the three representative indexes for 19 industries including "agriculture", "forestry", "animal husbandry and fishery", "manufacturing", "education", etc. (Table 1).

3.2   Data Normalization

In consideration of the dimensional difference between different indexes which will possibly affect the evaluation results, it is necessary to conduct normalization processing for the original data matrix. Because all the indexes selected for research are forward

Table 1. Information statistics of three representative indexes

Industry	Number of corporate units (Piece)	Annual GDP (100 million Yuan)	Number of Social Organization Standards (Piece)
financial industry	29733	9906.99	12
Agriculture, forestry, animal husbandry, fishery	45448	4769.99	230
Construction industry	72214	4651.50	61
…	…	…	…
Real estate	124429	10625.65	2

Note: the information statistics is from Chins Statistical Yearbook, stats.gd.gov.cn and national social organization standards information platform

indexes, it is unnecessary to conduct isotropic processing. The mode of normalization processing is as follows:

$$t_{ik} = \frac{t_{ik} - min\{t_{1k}, \ldots, t_{sk}\}}{max\{t_{1k}, \ldots, t_{sk}\} - min\{t_{1k}, \ldots, t_{sk}\}} \quad i = 1, 2 \ldots s$$

where, "k" is index (k = 1, 2, 3), "i" is industry (s = 1, 2, …,15), t_{ik} is the value of "i" corresponding to "k".

3.3 Entropy Calculation

After normalization processing for data matrix, it is necessary to calculate the proportion of industry "i" in index "k" to get the entropy of index "k", the formula is as follows:

$$p_{ik} \frac{t_{ik}}{\sum_{i=1}^{s} t_{ik}}, i = 1, \ldots, s, k = 1, \ldots m \tag{1}$$

$$e_k = -k \sum_{i=1}^{s} p_{ik} ln(p_{ik}), \quad k = \frac{1}{ln(s)} > 0, \ e_k \geq 0 \tag{2}$$
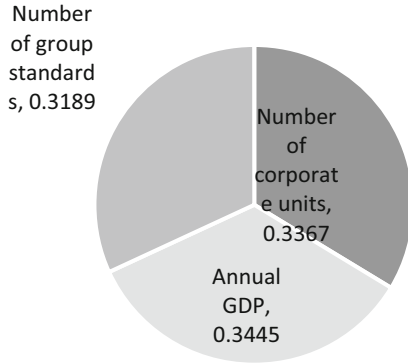
where, p_{ik} is the proportion of index "k" of industry "i" in total index, e_k is the information entropy of index "k".

3.4 Weight Calculation

$$d_k = 1 - e_k \tag{3}$$

$$w_k = \frac{d_k}{\sum_{k=1}^{m} d_k} \tag{4}$$

where, d_k is the variation coefficient of index "k". The bigger d_k is, the index is more important; w_k is the weight of index "k".

Fig. 1. The weights of "number of corporate units", "annual GDP" and "number of social organization standards"

3.5 Calculate the Comprehensive Score of Each Industry Development

$$g_i = \sum_{k=1}^{m} w_k p_{ik} \tag{5}$$

The weights of three indexes are calculated with matlab software, as shown in Fig. 1:

The figure shows that the indexes are equally important and can all reflect the development state of industry. Now we calculate the comprehensive score with the normalized data according to the weight, and put them in order from large to small according to the score, as shown in Table 2:

Table 2. Information statistics of three representative indexes

No.	Type of industry	Comprehensive score
1	Manufacturing industry	0.860 840 088
2	Wholesales and retail trade	0.466 079 66
3	Leasing and business service	0.232 875 1
4	Information transmission, software and Information Technology services	0.160 069 827
5	Real estate industry	0.137 844 968
6	Agriculture, forestry, animal husbandry, fishery	0.136 149 537
7	Transportation, storage and postal services	0.127 401 444
8	Electricity, heat, gas and water production and supply	0.116 824 477
9	Construction industry	0.106 285 067
10	Financial industry	0.099 910 433
11	Scientific research and technical services industry	0.094 513 963
12	Education	0.077 721 993

(continued)

Table 2. (*continued*)

No.	*Type of industry*	*Comprehensive score*
13	Health and social work	0.063 308 297
14	Public administration, social security and social organization	0.062 335 403
15	Residential services, repairs and other services	0.048 343 672
16	Hotels and catering services	0.045 215 77
17	Mining industry	0.032 310 72
18	Culture, sports and entertainment industry	0.029 510 49
19	Water conservancy, environment and public utilities management	0.015 642 514

The table shows that the industries with comprehensive scores of social organization standards development ranking the first three places are "manufacturing industry", "wholesales and retail industry" and "leasing and business service". The industry scoring the highest is "manufacturing industry" with score of 0.860840088, so the social organization standards development level of manufacturing industry in Guangdong is high, that's because Guangdong is a major province of manufacturing industry and possesses a manufacturing industry system with complete range and perfect supporting facilities. Guangdong has 31 main categories of industries (totally 31 in the whole country), more than 50 thousand industrial enterprises above designated size which ranks first in the country, and the most advanced manufacturing industry base in China and even the whole world. The industry scoring the lowest is "water conservancy, environment and public utilities management". The score is only 0.015642514 which is lower than the comprehensive score of manufacturing industry by about 55 times. We can see that the social organization standards development level of "water conservancy, environment and public utilities management industry" in Guangdong is low. It is mainly because the public infrastructures construction is mainly dominated by the government, the marketization degree of standards construction is low, and the social groups intervene less in the social organization standards construction of the field due to the influence of status, target of making profits and limited resources.

4   Conclusions

This paper uses the cluster analysis method and the entropy weight method to evaluate standard development situation of industry groups in Guangdong Province from the three dimensions of the number of corporate units, GDP each year and the standard number of the group. This paper draws the following conclusions:

1) The top three industries in Guangdong Province in terms of standard development situations of groups in Guangdong Province are: "manufacturing industry", "wholesale and retail industry" and "leasing and business service industry", with manufacturing industry scoring the highest.

2) The quantity of "manufacturing industry" social organization standards issued in Guangdong province accounts for nearly half of the total quantity of social organization standards issued. The quantity of "manufacturing industry" social organization standards is much higher than that of "agriculture, forestry, animal husbandry and fishery", ranking at the second place. This indicates that social organization standards of Guangdong Province are not developed in a balanced way. Instead, they are largely different, meaning that the development structures remain to be optimized and improved.

3) The industry with the lowest comprehensive evaluation score is "water, environment and utilities management" because public infrastructure belongs to "public products". Standards of this industry are mainly formulated by the government while social groups play a role in the industry.

The big data method is used in this paper to explore and analyze the data. However, due to the diversity of data sources and channels, the authority and objectivity of the data need to be checked. Therefore, this paper mainly conducts analysis according to the data from the standardized information platform of national groups. In the future, big data technology can be used to identify, sort out and use data from many channels.

References

1. Liu Ke. Research on the Legal Issues of Group Standards in my country [D]. Beijing: Central University of Finance and Economics, 2019.
2. Wu Huixia. Research on Government Governance of High-tech Industry Association Standards [D]. Changsha: Hunan University, 2018.
3. Xiao Yining, Zhu Peiwu, Fu Wenjing, et al.(2019) Development status and countermeasure analysis of "Made in Zhejiang" group standards [J]. Quality Exploration, 16 (2): 53-59.
4. Ma Xiaoou. (2019) Cultivating the organization of group standards and building the brand of group standards-Practice and thinking on the construction of group standards in Shandong Province [J]. China Standardization, (9): 71-76.

5. Zhang Yu. (2018) Group standards promote the construction of "Shanghai Brand" [J]. Quality and Standardization, (12): 40-41.
6. Zeng Yao. Higher education performance evaluation based on TOPSIS method and entropy method [D]. Guangzhou: Jinan University, 2016.
7. Liu Yuanyuan, Wang Shaoqiang, Wang Xiaobo, et al. (2020) Flood disaster risk assessment in Bangladesh, India and Burma based on AHP_entropy method [J]. Geographical Research, 39(8): 1892-1906.