Analysis of Factors Influencing People's Well-being Based on Multi-source Data Fusion

Xue Xu[1, *], Haidong Liu[1]

[1]College of Science, Tianjin University of Commerce, Tianjin, China
*Corresponding author

Abstract: As socialism with Chinese characteristics enters a new era, the people's needs for a better life are growing, and more and more people are starting to pay attention to their own well-being. The issue of well-being has received widespread attention. Traditional well-being questionnaires rely mainly on objective questions, while people's comments can describe their subjective well-being more specifically and in greater detail. This article collects relevant comment data on well-being through open-ended questionnaires, and uses text analysis methods such as Chinese word segmentation and LDA topic modeling to identify the thematic features that affect people's well-being. The logical relationships between the various topics are also organized. On this basis, the results are compared and analyzed with the empirical test results of well-being based on CGSS (2015) social survey data, in order to obtain more comprehensive results through the integration of multi-source data. This provides a useful way to enrich the theoretical and methodological research on well-being.

Keywords: people, well-being, multi-source data fusion, questionnaire survey, LDA topic model

1. Introduction

In recent years, with the development of Chinese economy and the increase in people's income, people's living standards have continued to improve. More and more people have begun to pay attention to their own well-being. The improvement of people's livelihood has become an objective need in Chinese current stage of development. The issues of sense of gain and well-being have received widespread attention. Well-being is people's subjective reflection on their real life based on the objective conditions of life. It is closely related to the objective conditions of individual life and reflects individual needs and values[1]. Subjective well-being is a kind of positive psychological experience of individuals towards their own existence and development status, generated by the joint action of these factors. Well-being can be divided into objective well-being and subjective well-being. The former emphasizes the objective authenticity of well-being, while the latter emphasizes subjective experience[2]. Compared with objective well-being, subjective well-being has become the mainstream research object in academia due to its greater inclusiveness. However, related research cannot completely abandon objective factors, and still conducts comprehensive analysis of both objective and subjective well-being[3].

Traditional well-being questionnaires rely mainly on objective questions, while people's comments can describe their subjective well-being more specifically and in detail. Based on this, this paper designs open-ended questionnaires and conducts statistical surveys on people's well-being, using text mining methods to identify factors influencing well-being. On this basis, an empirical test of the factors affecting well-being is conducted based on CGSS (2015) survey data, and the results of empirical test are compared and analyzed with the identified results of influencing factors of well-being. By integrating multi-source data, more complete results are obtained, which provides a useful way to enrich the theoretical and methodological research on well-being.

2. Related Work

Regarding research on well-being and its influencing factors, domestic and foreign scholars mainly take the relationship between economic income and subjective well-being as the research basis, and gradually introduce factors such as demographic factors, social security, public policy, and ecological environment into the research.

Demographic factors generally include age, gender, marital status, education level, health status, etc. In terms of age, most studies believe that well-being shows a U-shaped trend over a person's entire life cycle[4][5]. In terms of gender, women are usually happier than men[6]. Regarding the impact of health, marriage, and family on subjective well-being, it is believed that good health, marriage, and family harmony can enhance subjective well-being, while divorce, separation, bereavement and other factors can significantly reduce subjective well-being[7][8]. As for whether education level affects well-being, some scholars believe that education level is positively correlated with well-being; but some scholars believe that the well-being of highly educated groups is not necessarily higher than that of uneducated people, and may even have a negative impact[9].

A person is the sum of all social relations, and well-being depends on society, the surrounding environment and people[10]. Scholars mainly study the relationship between social factors and well-being from aspects such as living environment, social security, government public expenditure, and social relations. Research unanimously agrees that a good living environment, social security level, public service level, social fairness, and social relations promote well-being[11][12]. In terms of government public expenditure, related research believes that government size and residents' well-being show an inverted "U" relationship. Increasing public expenditure on education and health care can improve residents' well-being, especially for middle-income groups[13]. Zhou et al.[14], based on Chinese Livelihood Index survey data, believe that China has entered a stage where the role of economic growth in improving national well-being is gradually weakening. Improving public services such as compulsory education, medical care, ecological environment, and social security can significantly improve individual life satisfaction.

Regarding the impact of economic factors on well-being, domestic and foreign scholars mainly focus on the impact of income level, inflation, unemployment rate and other economic factors on national well-being. Earlier studies mostly believed that the higher the income level, the stronger the well-being. The mainstream view now is that income is highly correlated with national well-being, but there are large differences between countries at different development levels[15][16]. Regarding inflation, the mainstream view is that both expected and unexpected inflation have negative effects on residents' well-being[17]. Most studies believe that unemployment can reduce people's subjective well-being[18].

3. Identification of Factors Influencing People's Well-being Based on Open Questionnaire Data

3.1. Research Design

In order to systematically and comprehensively grasp the real situation of people's well-being, this paper collects 1,018 relevant comment data on well-being through open-ended questionnaires, and uses text analysis methods such as Chinese word segmentation and LDA topic model to identify the thematic features and probability distribution that affect people's well-being. Logical relationships between topics are also organized.

3.1.1. Design And Data Collection of Open-ended Questionnaires

(1) Determine the dependent variable. The explained variable or dependent variable in this topic is well-being, which is measured by directly asking respondents about their overall feeling of well-being in life. The corresponding question is "Overall, do you think your life is happy?" with the following options: very unhappy, relatively unhappy, hard to say happy or not, relatively happy, very happy.

(2) Questionnaire content. Through reviewing relevant literature, we know that the source of well-being comes not only from one aspect, but is the result of many factors. Different age and gender groups perceive well-being differently. Our questionnaire is divided into objective and subjective sections. The objective section includes people's level of well-being and basic information such as gender and age. The subjective section is: what makes you feel happy or unhappy? How do you think is happier? (Prompt: Comment around family, health, economy, work, study, emotions and interpersonal relationships.) This subjective question explores people's true feelings of well-being. The prompt information basically reflects the indicator content of factors influencing people's well-being.

3.1.2. Data Preprocessing

Chinese word segmentation, as the basic process of Chinese information processing, its accuracy will directly affect the results of Chinese text mining. Since the corpus of word segmentation comes from relevant comment data on people's well-being, there are many domain vocabularies related to well-being. This paper constructs a domain dictionary by extracting keywords from the literature, and

adds it as a supplementary dictionary to the general dictionary for word segmentation of the corpus of people's well-being. Stop words with high frequency but no information value for this topic, such as "but", "if", "oh", are removed based on the stop word dictionary.

3.1.3. Identification of Well-being Topics Based on LDA Model

The LDA topic model is a Bayesian probability model with a three-layer structure of words, topics and documents proposed by Blei et al. By introducing the topic dimension between documents and words, it reduces the dimension of vector space and can process large-scale corpora[19]. This model uses the bag-of-words approach, treating each document as a word frequency vector. α and β are hyper-parameters of the document's topic distribution θ and the distribution φ of words in the topic, respectively, following the prior Dirichlet distribution. This paper uses the LDA topic model to analyze the open questionnaire data. Through model training and parameter tuning, it identifies the topic features and their probability distributions contained in the well-being comment data. The top k words under each topic are selected and displayed in word clouds.

3.2. Identification of Factors Influencing People's Well-being

The statistics show that a total of 357 related comments on well-being from people aged 18-30, 165 from people aged 30-55, and 496 from people over 55 were collected. It can be seen from the review of relevant research literature that the influencing factors of well-being may differ among groups of different age levels. Therefore, when using the LDA topic model to identify the influencing factors of well-being in this section, the above three groups of data are analyzed separately by age.

3.2.1. People Aged 18-30

For the well-being data of the 18-30 age group, it is proven that when the number of topics is set to 3, the topic division is clearer. The top 20 representative words with higher occurrence probabilities under each topic are selected and displayed in word clouds, as shown in Figure 1.
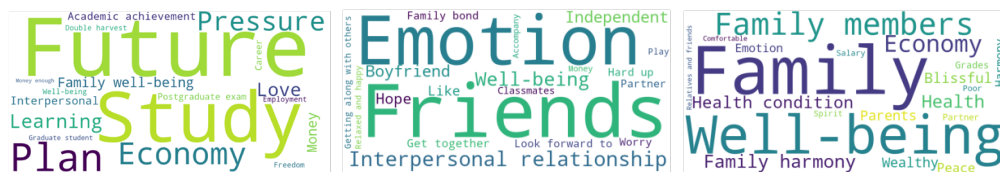

Figure 1: Word clouds of topics for people aged 18-30.

As shown in Figure 1, the 3 topics from left to right can be summarized as study, emotions and social interactions, and family.

From the word cloud of the study topic, words like "study", "learning", "academic achievement", "plan", "future", "pressure", "love", "family well-being", "money", "employment", "freedom", "career" have larger fonts. This indicates that for the 18-30 age group, the academic status may have an impact or pressure on their future and planning, thus affecting their well-being; family well-being, economic conditions, and their studies influence each other, which in turn affects well-being; in addition, academic and romantic double harvest is the beautiful aspiration that this group expects, which has an important impact on their well-being.

From the word cloud of the emotions and social interactions topic, words like "interpersonal relationship", "friends", "emotion", "boyfriend", "like", "independent", "get together", "partner", "hard up" have larger fonts. This indicates that for the 18-30 age group, their social contacts are mainly classmates and boy/girlfriends. Whether the interpersonal interactions with classmates are relaxing and pleasant, whether the feelings for their boy/girlfriends are sincere, whether the personality is independent, etc., all these will have an important impact on their well-being. In addition, from words like "hard up" and "money", it can be seen that getting together with classmates and emotional interactions also rely on their economic conditions to some extent, which in turn affects their well-being.

From the word cloud of the family topic, words like "family", "family harmony", "well-being", "economy", "health", "harmony", "blissful", "grades" have larger fonts. This indicates that for the 18-30 age group, a happy, perfect, harmonious family has an important impact on their well-being. It can be inferred from the review of the influencing factors of well-being in the second part that the economic and health conditions of a family will affect the well-being and perfection of the family, thereby

affecting their well-being. In addition, most of the respondents aged 18-30 are students, so academic performance will also affect family harmony and well-being to some extent.

3.2.2. People Aged 30-55

For the well-being data of the 30-55 age group, it is proven that when the number of topics is set to 4, the topic division is clearer. Due to the small amount of valid comment data for the 30-55 age group, the top 20 representative words with higher occurrence probabilities under each topic are selected and displayed in word clouds, as shown in Figure 2.

Figure 2: Word clouds of topics for people aged 30-55.

As shown in Figure 2, the 4 topics from left to right can be summarized as: family, spiritual culture, economic conditions, and career.

From the word cloud of the family topic, words like "physical health", "well-being", "family well-being", "marriage", "children", "family harmony", "health", "parents" have larger fonts. This indicates that for the 30-55 age group, the physical health of family members, family well-being and harmony, and marital status have an important impact on their well-being. For middle-aged groups, the healthy growth of children and parents' health also enhance their well-being. In addition, under the current environment of high housing prices where young people have to struggle for housing for most of their lives, having a car and a house enhances the well-being of middle-aged people to some extent.

From the word cloud of the spiritual culture topic, words like "life", "friends", "spirit", "travel", "culture", "pressure", "time", "vacation", "abundance", "hard up" have larger fonts. This indicates that for the 30-55 age group, tourism, leisure, and getting along with friends can enrich their spiritual and cultural lives and enhance well-being. Abundant economy, sufficient time or vacations are important guarantees for spiritual and cultural life, which indirectly affects their well-being.

From the word cloud of the economy conditions topic, words like "income", "wealthy", "stable", "family", "house", "money", "economy", "education level", "rest time" have larger fonts. For the 30-55 age group, whether emotionally or economically, they are the pillars of children, parents and the whole family. As an important support for family life, stable income, high and considerable salary, savings, satisfactory housing will bring them well-being. Economic income is also closely related to education level and educational background, which in turn affects their well-being.

From the word cloud of the career topic, words like "smooth", "career", "learning", "family", "future", "colleagues", "opportunities", "plan" have larger fonts. For the 30-55 age group, whether the career is going smoothly, planning for the future, harmonious work relationships with colleagues and bosses, whether there are intrigues, and whether the career is on the rise will affect their well-being. High education level and academic success have an important impact on career, thereby indirectly affecting their well-being. In addition, family, emotions and other factors will also affect their career to some extent, which in turn affects their well-being.

3.2.3. People over 55 Years Old

Figure 3: Word clouds of topics for people over 55 years old.

For the well-being data of people over 55 years old, it is proven that when the number of topics is set to 3, the topic division is clearer. The top 30 representative words with higher occurrence probabilities under each topic are selected and displayed in word clouds, as shown in Figure 3.

As shown in Figure 3, the 3 topics from left to right can be summarized as: health, family, and economy.

From the word cloud of the health topic, words like "children", "filial piety", "health", "medical insurance", "housing" etc. have greater weights, indicating that these factors have a greater impact on the quality of life of the elderly. For the elderly, physical health is one of the main factors affecting their well-being. Health becomes a prerequisite for the elderly to live a happy life. Healthy elderly people are more likely to experience the joys of life, so their subjective well-being will be higher. When suffering from illness, they experience more pain and misfortune, with greatly increased psychological pressure, so subjective well-being will decrease. When the elderly have lower pension income or poorer economic conditions and health, they need to rely on their children for financial support and living care. Therefore, in late-life, filial children can greatly enhance the elderly's well-being. Due to the special characteristics of the elderly population, their health conditions make them have higher demands on housing area, floor, elevator, etc., so the housing environment directly affects their quality of life and well-being.

From the word cloud of the family topic, words like "marriage", "well-being", "children", "economy", "health" have larger fonts, indicating a greater impact on the elderly's well-being. With the prevalent social phenomena of "low fertility" and "empty nest", a happy marriage is especially important in late-life. As they age, the elderly's health deteriorates, self-care ability declines, social circles gradually shrink, and it is easy to feel lonely. Marriage not only provides opportunities for mutual care, but also relieves the elderly's frequent emotional lows and fragile emotions mentally. Therefore, married elderly generally do not feel lonely and have higher well-being. The health of the elderly affects their working ability, which in turn affects the income of rural elderly. When the elderly are in poor health, the increased financial burden of medical treatment causes difficulties in life, which negatively affects their well-being. Children's care and financial support means an extra sense of security that greatly enhances the elderly's well-being.

From the word cloud of the economy topic, words like "well-being", "rich", "health", "medical insurance", "family" have a more significant impact on the elderly's well-being. Elderly people with better economic conditions have higher life satisfaction, which greatly enhances their well-being. When the elderly are in good health, they can participate in certain social activities and increase income, without the financial burden of medical expenses. The elderly in poor health cannot obtain income through labor, causing life difficulties and needing family care and support, which directly or indirectly reduces their well-being. Since social medical insurance is an important economic guarantee for the elderly and their families, having medical insurance can reduce the financial burden of medical expenses for the elderly, thus significantly increasing their well-being.

4. Empirical Test of Factors Influencing People's Well-being

4.1. Data Source and Processing

This part uses data from the 2015 Chinese General Social Survey (CGSS). The data was collected using a multi-stage stratified random sampling method, covering 28 provinces, municipalities and autonomous regions in China, a total of 478 villages, with 10,968 valid samples and extensive representation. After removing missing and abnormal values, 9,506 valid samples were retained.

Based on the questionnaire data and literature review results, the dimensions for measuring people's well-being include demographic factors, health and medical conditions, economic conditions, family and intergenerational support, social activities, hobbies, housing environment, personal perceptions, etc. Since there are many factors affecting well-being, including ordinal classification variables, unordered classification variables, and quantitative variables, this paper uses the Spearman's Rank Correlation Coefficient method to test the correlation between variables. The final variable selection results are shown in Table 1. Among them, A30 hobbies is the average value of A30-1 to A30-12 in the questionnaire. The smaller the value, the more hobbies.

Table 1: Preliminary screening results of explanatory variables.

Variable	Questionnaire Content	Code
People's well-being	A36. How happy do you feel about your life? Very unhappy 1, Relatively unhappy 2, Hard to say 3, Relatively happy 4, Very happy 5	a36
Demographic factors	A2. Respondent's gender: Male 1, Female 2	a2
	A31. Age	a31
	A7a. Your current education level: No education 1, Private school 2, Primary school 3, Junior high school 4, Vocational high school 5, High school 6, Secondary school 7, Technical secondary school 8, Junior college (adult higher education) 9, Junior college (formal higher education)10, Undergraduate (adult higher education) 11, Undergraduate (formal higher education) 12, Graduate school or above 13, Other 0	a7a
	A69. Your current marital status is: Unmarried 1, Cohabitation 2, First marriage 3, Remarriage 4, Separated 5, Divorced 6, Widowed 7	a69
Health status	A15. How do you feel about your health now? Very unhealthy 1, Relatively unhealthy 2, General 3, Relatively healthy 4, Very healthy 5	a15
Social security	A6. Number of social security benefits	a6
Interpersonal relationships	A30. Hobbies (the smaller the value, the more hobbies)	a30
	A311. In the past year, have you done the following in your spare time (socializing, visiting)? Never 1, Rarely 2, Sometimes 3, Often 4, Very frequently 5	a311
Economic conditions	A8a. What was your personal income last year (10,000 yuan)?	a8a
	A62. What is your family's total annual income (10,000 yuan)?	a62
	A65. How many properties does your family currently own?	a65
	B1. Compared with your peers, what do you think your socioeconomic status is? Higher 1, About the same 2, Lower 3	b1
Family and social	A35. Do you think today's society is fair? Very unfair 1, Relatively unfair 2, Hard to say 3, Relatively fair 4, Very fair 5	a35
	A68. Number of children	a68

4.2. Empirical Test

Table 2: Empirical test of factors influencing young people's well-being

a36	under 30 years old			30-55 years old			55 years old and above		
	Coef.	Std.Err.	P>|z|	Coef.	Std.Err.	P>|z|	Coef.	Std.Err.	P>|z|
a2	0.009	0.113	0.933	0.176	0.063	0.005	0.306	0.068	0.000
a31	-0.085	0.019	0.000	-0.003	0.005	0.597	0.008	0.005	0.098
a7a	0.055	0.018	0.003	-0.008	0.013	0.559	0.015	0.015	0.321
a8a	-0.003	0.002	0.103	0.003	0.002	0.085	0.004	0.004	0.348
a15	0.503	0.071	0.000	0.397	0.033	0.000	0.396	0.033	0.000
a30	-0.289	0.130	0.026	-0.367	0.068	0.000	-0.504	0.072	0.000
a311	0.172	0.057	0.003	0.026	0.031	0.412	0.026	0.031	0.398
a35	0.302	0.059	0.000	0.485	0.032	0.000	0.586	0.037	0.000
a6	0.050	0.067	0.458	0.082	0.040	0.040	0.037	0.052	0.484
a62	0.001	0.003	0.697	0.000	0.001	0.805	0.000	0.001	0.768
a65	0.107	0.065	0.098	0.321	0.057	0.000	0.163	0.061	0.008
a68	0.230	0.097	0.018	0.081	0.041	0.045	0.153	0.027	0.000
a69	0.261	0.075	0.001	-0.185	0.035	0.000	-0.053	0.021	0.013
b1	-0.393	0.111	0.000	-0.647	0.062	0.000	-0.710	0.063	0.000
/cut1	-2.532	0.850		-2.375	0.489		-1.846	0.550	
/cut2	-0.748	0.843		-0.932	0.487		-0.438	0.548	
/cut3	2.429	0.845		2.408	0.488		2.828	0.550	

When we use subjective indicators to measure well-being, people are often asked to subjectively quantify their own well-being: unhappy, happy, very happy, etc. Well-being as a response variable is a

discrete and ordered classification variable. The ordered logit model can be used to evaluate well-being.

Considering that people of different ages may have different influencing factors for well-being, this paper divides the 9,506 valid samples after screening the CGSS (2015) data into three groups by age. The youth group under 30 years old has 1,375 observations; the middle-aged group over 30 and under 55 has 4,328 observations; the elderly group aged 55 and above has 3,803 observations. The ordered logit model is selected to verify the significance of various influencing factors on the well-being of different age groups. The regression results are shown in Table 2.

By analyzing Table 2, we can make the following observations:

(1) Factors like marital status, number of children, health status, economic conditions, hobbies, social fairness, and socioeconomic status have significant impacts on youth, middle-aged, and elderly groups.

(2) Social security factors have a significant impact on the well-being of the middle-aged group, but no significant impact on youth and elderly groups.

(3) Factors like education level, age, and social frequency have a significant impact on the youth group's well-being, but no significant impact on middle-aged and elderly groups.

4.3. Comparative Analysis of Identification and Empirical Test Results

The identification results of factors influencing people's well-being based on open-ended questionnaire data are compared and analyzed with the empirical test results as follows:

(1) For the elderly group's well-being comment data, the identification results of topic features show that factors influencing their well-being mainly include economic conditions, medical insurance, spiritual satisfaction, and interpersonal relationships. The empirical test shows that factors influencing this group's well-being include health status, marital status, interpersonal relationships, and social security. The two analysis results are basically consistent.

(2) For the middle-aged group, they not only support the family but also have very complex social identities. Their comments on well-being cover many aspects, including self-development, elderly, children, health, study, life, etc. The word cloud analysis shows that factors influencing their well-being include marital status, health status, personal income, number of family properties, interpersonal relationships, education level, etc. The empirical test shows factors like health status, hobbies, marital status, number of properties, etc. The two analysis results are basically consistent.

(3) For the youth group, the identification results of well-being comment topics show that the representative words are mainly "study", "learning", "academic achievement", "postgraduate entrance exam", "graduate student", "family", "emotion", "friends", "interpersonal interaction", etc. Their comments on well-being mainly focus on study, emotions, family and interpersonal interactions. The empirical analysis shows that the main influencing factors are health status, hobbies, education level, marital status, age, social frequency, etc. The two results are basically consistent. However, health status is not very significant in the identification results, which may be related to the particularity of the youth group's age and the small sample size.

5. Conclusion

This paper designs open-ended questionnaires and conducts statistical surveys on people's well-being, using text mining methods to identify factors influencing well-being and obtaining relatively clear well-being topic features. On this basis, an empirical test of influencing factors of well-being is conducted based on CGSS (2015) survey data, and the results are compared and analyzed with the identification results, providing useful insights for relevant research on well-being.

Due to the limited data obtained, there are relatively few valid evaluation data on well-being collected in the paper. It can be seen from the identification results of well-being topic features that the feature words of each topic are relatively sparse. Therefore, there is room for improvement in obtaining valid questionnaires. With the rapid development of social networks, relevant information on well-being is emerging from platforms like news, media, and social networks, which can provide new analytical perspectives for the study of people's well-being. In future research, it is hoped to find relevant data from authoritative online platforms and incorporate them into this study using appropriate

methods.

Acknowledgement

References

[1] Zhanjun Xing. Construction of a Short Version of Subjective Well-Being Scale for Urban Residents in China [J]. Chinese Journal of Behavioral Medicine Science, 2003(6):703-705.

[2] Haiying Wang, Ying Xia, Dongsheng Sun, Lin Zhang, Hao Wei. Meta-Analysis of Influencing Factors of Subjective Well-Being of Chinese Farmers [J]. China Agricultural Resources and Regional Planning, 2021, 042(006): 203-214.

[3] Dongdong Li. Review of Domestic and Foreign Research Progress on Subjective Well-Being in China from the Perspective of Economics [J]. Technology Economics and Management Research, 2019(9): 103-108.

[4] Wise J. Well-being is U shaped, highest in the teens and 70s, survey shows[J]. BMJ, 2012, 344:1.

[5] Blanchflower D G, Oswald A J. Is Well-being U-shaped over the life cycle[J]. Social Science & Medicine, 2008(8):1733-1749.

[6] Jing Z. Construction and Application of Piano to Intelligent Teaching System Based on Multi-Source Data Fusion[J]. Journal of Circuits, Systems and Computers, 2023, 32(04).

[7] Yanjie Bian, Yang Xiao. A Comparative Study on Subjective Well-being between Chinese and British Residents [J]. Sociological Studies, 2014(2):22-42.

[8] Wijekoon W, Ranjan W, Vitharana S. A Review of Gardening as a Recreation Impact on People's Psychological wellbeing[J]. IOSR Journal of Sports and Physical Education, 2021(11):68-74.

[9] Ting Li. Which Generation Is Happier? - The Evolution of Subjective Well-being of Chinese Residents from the Perspective of Age, Period and Cohort Analysis [J]. Population and Economics, 2018(1):90-102.

[10] Marx and Engels Selected Works (Vol. 3) [M]. Beijing: People's Publishing House, 2012:1002.

[11] Di Zhu. Market Competition, Collective Consumption and Environmental Quality - An Analysis of Urban Residents' Life Satisfaction and Its Influencing Factors [J]. Sociological Studies, 2016(3): 193-217.

[12] Yanping Wang. New Progress in the Research on Well-being Economics [J]. Economic Perspectives, 2017(10):130-146.

[13] Kotakorpi K, Laamanen J P. Welfare state and life satisfaction: Evidence from public health care [J]. Economica, 2010(307):565-583.

[14] Shaojie Zhou, Wang Hongchuan, Su Yang. How Can Chinese People Have A Higher Level of Well-being - Based on China Well-being Index Survey [J]. Management World, 2015(6):8-21.

[15] Diener E, Seligman M E P. Beyond money: toward an economy of Well-being [J]. Psychological Science in the Public Interest, 2004(1):1-31.

[16] Veenhoven R. Inequality of Well-being in nations [J]. Journal of Well-being Studies, 2005(4): 351-355.

[17] Graham C, Felton A. Inequality and Well-being: insights from Latin America [J]. The Journal of Economic Inequality, 2006(1):107-122.

[18] Daly M, Delaney L. The scarring effect of unemployment throughout adulthood on psychological distress at age 50: Estimates controlling for early adulthood distress and childhood psychological factors [J]. Social Science & Medicine, 2013, 80:19-23.

[19] Blei D M, Ng A Y, Jordan M I. Latent Dirichlet Allocation [J]. The Journal of Machine Learning Research, 2003, 3(1): 993–1022.