Research on Social Stability Based on PCA Analysis and Logistic Regression

Weiming Xiong [#, *], Xiangyu Han [#], Yucheng Tang [#]

China University of Petroleum-Beijing at karamay, Xinjiang, China, 834000

* Corresponding Author Email: 18248078632@163.com

Abstract. This paper investigates a series of regime change movements named after colors that occurred in the CIS and Central Asian regions in the early 21st century, conducted in peaceful and non-violent manners. Firstly, this study collects nine highly representative warning indicators in three aspects: "politics, economy, education," and assigns names to the three main components as "political indicators, economic indicators, cultural indicators" through principal component analysis, thereby constructing an indicator system affecting social stability. Secondly, stepwise regression analysis is used to screen out the significantly influential warning indicators on whether society is turbulent and a logistic model for social turbulence warning is constructed. Finally, relevant suggestions are once again offered in the three aspects of "politics, economy, education," providing effective resolution strategies for policy makers. The study concludes that the warning model of turbulence rate is related to the unemployment rate and the proportion of account balance, so the relevant departments need to timely adjust the unemployment rate and the proportion of account balance to maintain the safety level of social turbulence rate with the greatest efficiency.

Keywords: Logistic model, Variance inflation factor, Backoff test.

1. Introduction

At the end of the last century, peaceful and non-violent regime change movements named after colors took place in several countries and regions. At the beginning of this century, they achieved their goals by attacking the weakest link affecting social stability [1]. Multi-factors are originally an organic whole, but once the ratio of the forces between the various factors is unbalanced, society will be deformed, and social stability will be shaken. Then, before social stability problems appear, if you can find some "traces" in some precursors and intervene in time, you may avoid social risks and prevent unnecessary social turmoil [2].

2. Materials and methods

2.1. Data source

This article uses the data of the 2023 "Huasu Cup" International College Student Mathematical Contest in Modeling (https://www.saikr.com/vse/mcmicm/2023) data for research and analysis.

3. Model building and solution

3.1. Establishment of comprehensive evaluation system based on PCA analysis.

Principal component analysis is a comprehensive evaluation method that has the advantage of performing dimensionality reduction analysis on a large number of variables. Based on these characteristics, we decided to use nine early warning indicators to construct an assessment of social risks based on principal component analysis. We choose to accumulatively account for at least 75% of the variance explanation rate of early warning indicators, and the three principal components generated at this time will be used as the main reference dimensions of the indicator system.

3.1.1 KMO and Bartlett test

Before principal component analysis, we conducted KMO and Bartlett's test. The results are shown in Table 1.

Table.1. KMO and Bartlett test

KMO value		0.652
Bartlett test for sphericity	Approximate chi-square	192.174
	df	36
	p value	0.000

Table 1 shows that: KMO value is between 0.6 and 0.7, indicating that PCA analysis can be performed; the corresponding p value of Bartlett test is less than 0.05, which also indicates that it is suitable for principal component analysis.

3.1.2 Calculation of eigenvalues and variance interpretation rate

For the proposed 9 variables, it needs to be compressed into as few new variables as possible, and these new variables retain as much original information as possible in reflecting the information of the indicators (variance explanation rate>=75%). At the same time, according to the actual needs, a comprehensive variable of three mutually irrelevant variables is taken out, and this step can be solved by variance explanation. The results are shown in Table 2.

Table.2. Extraction of characteristic root/principal component

serial number	characteristic root	Variance explained %	Cumulative percentage%
1	2.424	26.929	26.929
2	1.767	22.629	46.558
3	1.157	22.855	75.413
…	…	…	…

Three principal components were constructed (the first three components were used to calculate the cumulative variance), and the characteristic root values were all greater than 1, which were 2.424, 1.767, and 1.157 in turn. The variance explanation rates of these three principal components were 26.929%, 22.629%, and 22.855%, respectively, and the cumulative variance explanation rates were 75.413 %.

3.1.3 Naming and weight calculation of principal components

In order to seek the affiliation relationship between the latest three principal components and the original early warning indicators, we obtained the table of factor loading coefficients after rotation by means of factor analysis, and we can find that:

The first principal component is related to "account balance ratio, deposit interest rate, per capita GDP", which can be summarized by the term "economic level". The economic order is chaotic, the gap between rich and poor is widening, unemployment is serious, and there are many poor people. The unemployed and poor have become active participants in "street politics"[3].

The second principal component is related to "corruption perception index, unemployment rate, GDP per capita", we can use "political level" to summarize. The ruling party is corrupt, and voters lack confidence in them; the greed of those in power and "people who work for the welfare of the people" is a devastating blow to a country.

The third principal component is related to "illiteracy rate, year-on-year education expenses", which can be summarized by "education level". The uneducated people are ignorant and backward in thinking, and are easily distorted by the opposition, which in turn fuels the flames and stimulates more people to join the social unrest.

Assuming that the three principal components are respectively PC_1, PC_2, PC_3, and introducing a comprehensive concept function $PC_{\text{comprehensive}}$, after selection, the three principal components are:

$$\begin{cases} PC_1 = 0.321x_1 - 0.321x_2 - 0.040x_3 + 0.233x_4 + 0.001x5 - 0.294x_6 + 0.307x_7 - 0.091x_8 - 0.058x_9 \\ PC_2 = -0.048x_1 + 0.094x_2 - 0.402x_3 + 0.368x_4 + 0.131x_5 - 0.011x_6 + 0.113x_7 + 0.455x_8 + 0.163x_9 \\ PC_3 = -0.081x_1 + 0.009x_2 + 0.038x_3 + 0.197x_4 + 0.154x_5 + 0.024x_6 + 0.169x_7 - 0.297x_8 - 0623x_9 \end{cases} \quad (1)$$

The contribution rate of these three principal components is weighted respectively (the results are shown in Figure 1), and the comprehensive evaluation model of the principal components is constructed, and the expression is obtained after solving by Matlab software:

$$PC_{comprehensive} = 0.5665 \times PC_1 + 0.265 \times PC_2 + 0.1685 \times PC_3 \quad (2)$$

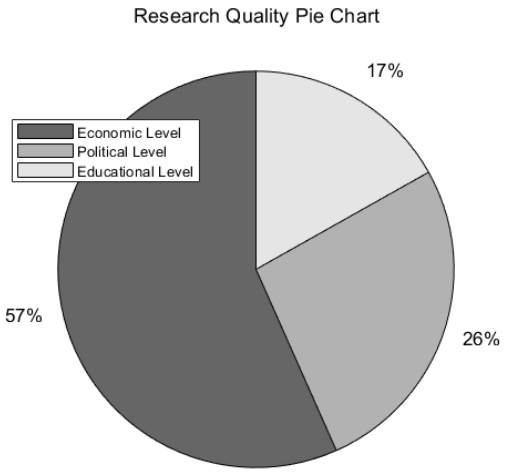

Figure 1. Pie chart of principal component weights

The comprehensive evaluation function is obtained by the linear addition of three principal components, and the total score represents the level of social stability. In order to improve social stability and maintain the smooth operation of the economy and society, we need to make a fuss about the three aspects of "economy, politics, and education". At the same time, since the "economic level" is given a weight of 56.65%, we should devote more energy to economic construction while maintaining the "triangular balance" of the principal components, so as to maintain social stability and maximize efficiency.

3.1.4 Principal components and causality of social stability

(1) Economic level

Economic development can be defined as a steady and orderly continuous growth of the economic system over a long period of time. The key lies in ensuring the stability of the growth rate and avoiding the excessive expansion of the gap between the rich and the poor, which may threaten the stable development of the economy and cause extreme fluctuations. From the point of view of historical materialism, the development of productive forces is considered to be the most revolutionary driving factor, which affects and shapes the political system and all aspects of society of the superstructure. Therefore, economic development has a profound impact on political stability and the stability of the whole society.

Especially for some countries that are in the process of modernization and development, there are prone to unfair resource distribution, possession and use at this stage, as well as the crazy rent-seeking of power in resource allocation, which often leads to the continuous widening of the gap between the rich and the poor and the huge differentiation of social classes, which in turn leads to social unrest, interest group struggles and group violence and other political instability that affect sustainable economic development[4], eventually leading to social chaos .

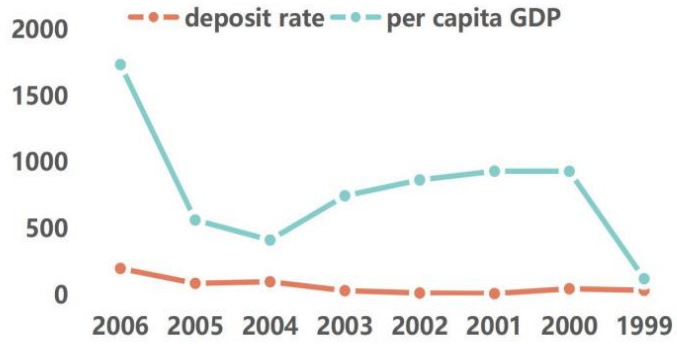

Figure 2. Trends of related economic indicators in Zimbabwe

From Zimbabwe's economic data at the beginning of the 21st century (as shown in Figure 2) , a remarkable phenomenon can be observed: the deposit interest rate during this period was significantly higher than the growth trend of its per capita GDP. This data phenomenon implies that the country's currency may have suffered a significant depreciation, causing the central bank to face challenges in the formulation and implementation of monetary policy, and it is difficult to carry out effective macro-control on the domestic market. This situation has further led to an imbalance in the allocation of social resources, the problem of social class differentiation has gradually intensified, and social unrest has worsened. Through the results of principal component analysis, we can further verify this observation, emphasizing the strong association between economic indicators and social stability. This provides an important reference for policy makers in macroeconomic management and social governance, emphasizing the intrinsic link between economic development and social stability.

(2) Political level

Political stability can be defined as the relative stability or moderate changes in a country's political system, subjects of political power, political life, policies and regulations, and political order. It also includes the orderly operation of the political system and the ability of the political system to operate normally in accordance with established procedures. Under this framework, the government should basically have the ability to control social conflicts within an acceptable security range. For countries that are in the process of modernization, maintaining political stability is a major challenge they must face and solve[5].

The famous political scientist Samuel Huntington once pointed out : "The main problem is not the pursuit of freedom, but the establishment of a legal public order. People can have order without freedom, but they cannot enjoy freedom without order"[6] .This point of view further emphasizes the core position of political stability, highlights the priority of order, and provides a profound theoretical perspective for us to understand and evaluate the importance of political stability.

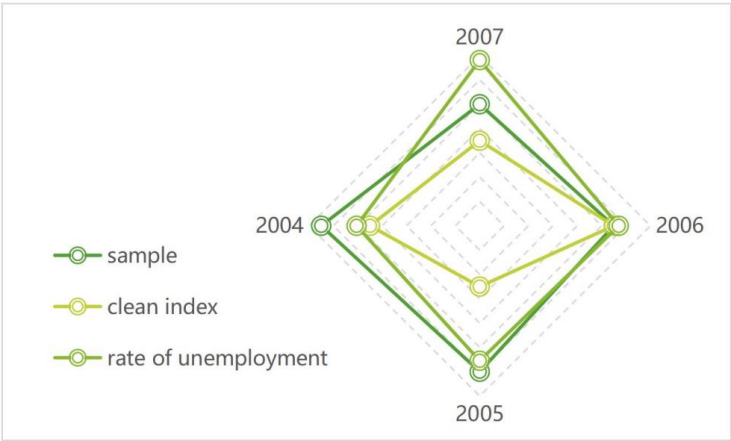

Figure 3. Radar map of warning indicators during the Ukrainian Color Revolution

As shown in Figure 3, the early warning indicator chart, the relevant data of Ukraine from 2004 to 2007 shows that the Corruption Perception Index fluctuated significantly during this period; at the same time, the unemployment rate also increased year by year . The results of PCA further strengthen this observation, especially in transitional countries like Ukraine, where political stability plays a central role in maintaining social order, promoting economic development, and improving citizens' well-being. We can therefore infer that political stability has a profound effect on the maintenance of social stability, which further supports Samuel Huntington's point.

(3) Education level

Marx emphasized the importance of education to the stable development of society in his "Theses on Feuerbach". Specifically, education provides key support for economic development by cultivating and developing talents, promoting knowledge innovation and technological progress. At the same time, education is also conducive to the maintenance of social order and the guarantee of social stability by improving citizens' social cognition and participation ability.

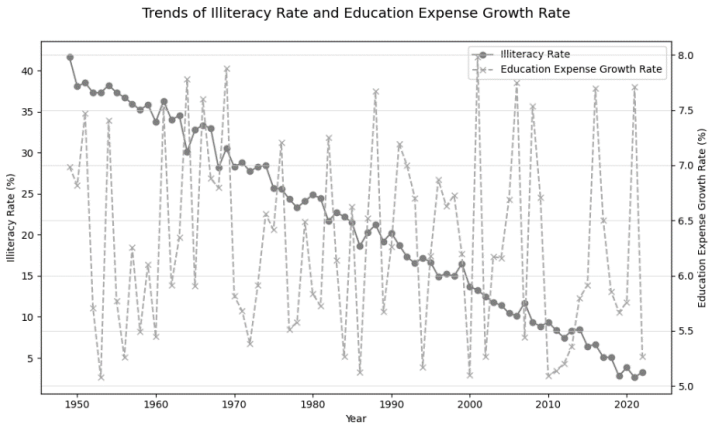

Figure 4. Trends in the rate of illiteracy and the growth rate of educational expenses

Since the founding of the People's Republic of China in 1949, the level of education has continued to improve, and the illiteracy rate has dropped from 40% to 2%. With the continuous improvement of the socialist education system, my country's economy has also continued to flourish, which has enhanced my country's overall national strength, continuously improved the people's happiness index, and made the society more stable and harmonious. (As shown in Figure 4)

3.2. Social early warning based on logistic regression.

3.2.1 Screening of early warning indicators

In this paper, stepwise regression is used to delete the less significant early warning indicators until the two most representative early warning indicators are found. Under common sense, the expression of the logistic function is kept as refined as possible. The screening results of early warning indicators are shown in Table 3:

Table. 3. Results of stepwise regression analysis (n=100)

	B	standard error	Beta	T	p	VIF
unemployment rate	0.040	0.006	0.362	6.654	0.000**	1.879
account balance ratio	0.010	0.004	0.125	2.773	0.007**	1.293
R^2	0.852					
F	F (5, 94) = 107.926, p = 0.000					

Table 3 shows that after the automatic identification of the model, two early warning indicators of "unemployment rate x_2" and "account balance ratio x_6" are finally left.

3.2.2 Introduction and solution of logistics model

Logistic is a generalized linear regression analysis model. The main idea is: to establish a regression equation on the decision boundary according to the existing data, and then map the regression equation to the classification function to achieve classification[7] .

In this paper, we transform the early warning problem of social unrest into a probability value to quantify social risk. We believe that the higher the rate of social unrest, the lower the social stability, and the easier it is for "color revolutions" to break out.

A logistic regression model is established through relevant data, recorded as the social unrest rate, and the social stability rate is 1- p_i. The correlation coefficient of the indicators in the regression equation is fitted through the early warning indicators, and the model is:

$$\ln \frac{p_i}{1-p_i} = \beta_0 + \beta_1 X_1 + ... + \beta_k X_k \tag{3}$$

After Matlab software fitting, the parameters are obtained as follows： β_0=-7.470, β_2=0.751, β_6=0.321.The final model of this problem is:

$$p_i = \frac{1}{1+\exp^{-(0.751\times unemployment\ rate+0.321\times Account\ balance\ ratio-7.40)}} \tag{4}$$

3.2.3 Fitting quality of logistic model

Through the above analysis, the result of Logit regression is obtained.

Table.4. Results of likelihood ratio test of binary Logit regression model

Model	-2x log-likelihood	chi-square value	df	p	AIC value	BIC value
intercept only	110.216					
final model	29.561	80.655	2	0.000	35.561	43.377

From Table 4 that the original assumption of the model test here is that the quality of the model is the same whether or not the independent variable (unemployment rate, account balance ratio) is put in; here the p value is less than 0.05, thus indicating that the original assumption is rejected, that is, when the model is built this time, the independent variable put in is valid[8], and this model construction is meaningful .

3.2.4 Sensitivity analysis of logistic model

Sensitivity analysis is a method to study and analyze the sensitivity of a system (or model) state or output change to changes in system parameters or surrounding conditions[9].This paper constructs a logistic model for early warning of social unrest:

$$p_i = \frac{1}{1+\exp^{-(0.751\times unemployment\ rate+0.321\times Account\ balance\ ratio-7.40)}} \tag{5}$$

In order to explore the difference between the two early warning indicators on social stability, and to give an advanced early warning effect through observation data, the results calculated by MATLAB are shown in Figure 5, Table 5 and Table 6 .

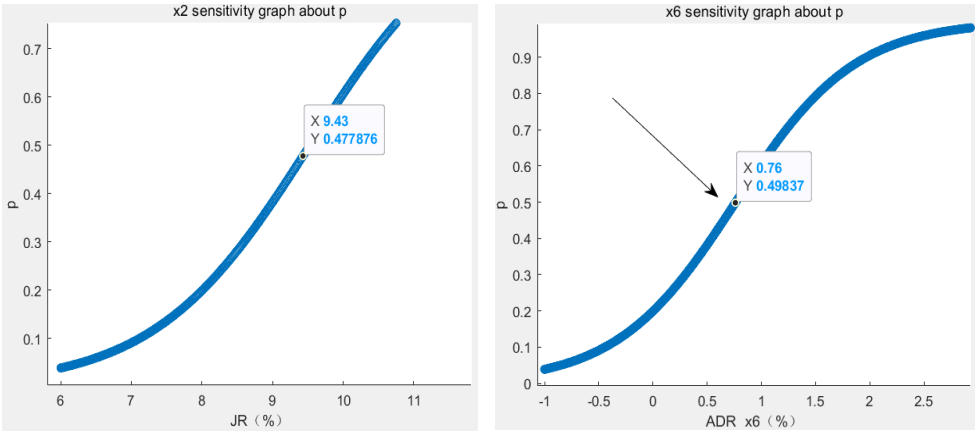

Figure 5. Sensitivity Analysis

Table.5. x_2 change table for sensitivity data p

unemployment rate x_2	social unrest rate p	unemployment rate x_2	social unrest rate p
6.0	0.0386051	9.15	0.414893
6.12	0.042876	9.25	0.43717
6.17	0.0447858	9.43	0.477876

It can be concluded from the graph: when $x_2$the range is between [-1,-0.775], the unemployment rate $x_2$changes by 1%, the social unrest rate $p$will change by about 0.08%, and the $x_2$relative $p$sensitivity in this range is very weak; when $x_2$the range is [0.61, 1], the unemployment rate changes by 1%, the social unrest rate $p$will change by about 6.2%, and the $x_2$relative $p$sensitivity in this range is very strong, and social unrest is easy to break out. Unemployment fell to safe levels.

Table.6. x_6 change table for sensitivity data p

account balance ratio x_6	social unrest rate p	account balance ratio x_6	social unrest rate p
-0.905	0.0455722	0.61	0.430461
-0.775	0.057064	0.65	0.459719
-0.42	0.103618	0.76	0.49837

It can be concluded from the graph: when $x_6$the range of [-1,-0.3], $x_6$every 1% change in the account balance ratio, the social unrest rate will change by about 0.2%, and the $p$relative sensitivity $p$in this range $x_6$is very weak; when $x_6$the range is [0.61,0.9], every 1% change in the account balance ratio, the social unrest rate will change by about 6.3%, the relative sensitivity in this range is very strong, and social unrest is easy to break out. According to this early warning, the government needs to introduce relevant monetary policies in time to alleviate the economic pressure of the people and prevent large-scale people from creating riots because of "poverty".

3.3. Recommendations to prevent color revolutions

Based on the research of this paper, based on the basis of the social stability index system, through the processing and classification of the national data in the "color revolution", we found that the main factors leading to social instability are the instability of "politics, economy and education"[10].

How to avoid the outbreak of "color revolution"? It needs to rely on the mutual balance and restriction among the key early warning indicators. Therefore, the early warning index system is to maintain the triangular balance of "economic level PC_1, political level PC_2 and educational level PC_3".

(1) Economic point of view

The results of the study show that the proportion of account balances in the color revolution countries is relatively high, which to some extent reveals the widening gap between the rich and the

poor in society. To alleviate this social contradiction, policy makers need to consider intervening through economic means. Specific strategies may include improving the living standards of grassroots workers and adopting macro-control policies to redistribute wealth in order to reduce the gap between the rich and the poor and alleviate social conflicts.

(2) Political angle

During the color revolution, the government corruption index of the relevant countries was generally low, which may reflect serious problems in the bureaucratic system, further exacerbating political turmoil and social disorder. Therefore, this phenomenon needs to be solved by improving the national judiciary and the construction of a clean government, as well as establishing a complete bureaucratic supervision mechanism. In addition, in order to achieve continuous supervision of risk management at the government level, it is necessary to enhance the government's scientific management capabilities, break the limitations of government entities, and reduce the phenomenon of "organized irresponsibility", so as to enhance public trust in the government.

(3) Educational perspective

It is crucial to realize the combination of the risk assessment of social stability and the current mass line education practice, which can enhance the communication awareness, methods and ability of public officials and the public. It is recommended that schools at all levels and administrative management schools incorporate social stability risk assessment into the curriculum system, clarify its role in maintaining social stability, understand its operating principles and steps, and understand its connection between different departments. In addition, strengthening the training of leading cadres at all levels to make social stability risk assessment an important working method is also the key.

4. Conclusion

The "color revolution" has dealt an immeasurable blow to the normal life of the people of various countries and the functioning of the economy and society. In order to protect the sanctity of state power and maintain the stable development of social economy, the early warning of "color revolution" is particularly important. Firstly, this paper constructs a comprehensive evaluation system of social stability in three aspects of "politics, economy and education" : $PC_{comprehensive} = 0.5665 * PC_1 + 0.265 * PC_2 + 0.1685 * PC_3$. Secondly, a logistic early-warning model is constructed after stepwise regression is used to eliminate the indicators of insignificant correlation:

$$p_i = \frac{1}{1+\exp^{-(0.751\times unemployment\ rate+0.321\times Account\ balance\ ratio-7.40)}} \tag{6}$$

On this basis, this paper again puts forward targeted suggestions from the three perspectives of "politics, economy, and education" and provides policy makers with effective solutions.

References

[1] Cao Jialu. Research on America’s Central Asia Strategy from the Perspective of GeostrategicExternalities [D]. LanzhouUniversity, 2022.DOI: 10.27204/d.cnki.glzhu .2022.002836.

[2] Niu Weigan, Tan Dongchen. Research on "Color Revolution" from the Perspective of the Legacy of the Cold War [J]. World Economy and Politics, 2022(04):4-29+156.

[3] Li Lan.Research on "Color Revolution" in Commonwealth of Independent States [D]. Nankai University, 2013.

[4] Wang Wenwen. The Fundamental Elements of Maintaining Social Stability and Its Important Support [J]. People's Forum, 2021(08): 34-37.

[5] Wang Baofeng. Discrimination and Analysis of the Stability Problem in Rawls' Theory of Justice [D]. Liaoning University, 2019.

[6] Samuel Huntington. "Clash of Civilizations" [J]. People's Rule of Law, 2019(17): 108.

[7] Zhang Hanxia. Scenario Analysis for Linear Regression and Logistic Regression [J]. Automation and Instrumentation, 2022 (10): 1-4+ 8. DOI: 10.14016/j.cnki .1001-9227.2022.10.001.

[8] Liu Jiantao. The Influence of Public Security Law Enforcement Credibility on Public Security—An Empirical Study Based on Logit Regression Analysis [J]. Journal of Wuyi University,2020,39(11): 45-50.DOI: 10.14155/j.cnki.35-1293/ g4.2020.11.008.

[9] Wang Bingjie. Empirical Research on Credit Risk of China's Listed Companies Based on Logit Model [D]. Shanxi University of Finance and Economics, 2016.

[10] Hu Qiyang.Application of Linear Programming Algorithm in Parametric BIM Model Optimization [D]. Tianjin University, 2018.