

10th International Conference on Information Technology and Quantitative Management

Research on economic development trend of reform and opening up: based on big data modeling analysis method

Anzhou Wang^a, Yichen Yang^b, Shuo Zhang^{c,*}, Yikun Du^d, Yaping Zeng^b, Tian Wu^b

^aFaculty of Natural, Mathematical & Engineering Sciences, King's College London, London WC2R 2LS, United Kingdom

^bSchool of Economics and Management, Tsinghua University, Beijing 100084, China

^cFaculty of Arts & Humanities, King's College London, London WC2R 2LS, United Kingdom

^dCollege of Information Science and Technology, Beijing University of Chemical Technology, Beijing 100029, China

Abstract

The development of information technology proves the possibility of acquiring large samples of policy text information and promotes innovative research based on the policy informatics approach of fusing data from multiple sources. This paper constructs a new public policy model based on existing data on economic system reform, and uses multi-types of textual big data such as government website, official weblog, and social media as the object of collection, to conduct a big data measurement on policy concerns related to the theory and practice of economic system reform since the reform and opening up. And this paper constructs a fine-grained measurement and portray the evolutionary trend of policy concerns and policy implementation. Based on the database, this paper helps to analyse the economic development trend since the reform and opening up, evaluate and monitor the implementation effect of different incentive policies and establish an iterative feedback mechanism, and provides a reference for relevant decision-making departments.

© 2023 The Authors. Published by Elsevier B.V.

This is an open access article under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0>)

Peer-review under responsibility of the scientific committee of the Tenth International Conference on Information Technology and Quantitative Management

Keywords: Economic Reform and Opening Up; Economic System Reform; Policy Informatics; Text Mining; Big Data

1. Introduction

The 40-year history of China's economic reform and opening up is magnificent. In promoting the development of the theory and practice of economic system reform, the state has formulated and promulgated a large number of economic policies and economic projects. Combining the interdisciplinary fields of management, data science, systems engineering, statistics, and other areas, in the context of big data-driven development of the network, from

* Corresponding author. Tel.: +44 7799382712

E-mail address: shuo.2.zhang@kcl.ac.uk, wanganzhou2001@126.com

the perspective of policy informatics, the measurement of policy concerns and public perceptions, integrating big network data, is of great value on the theoretical and practical significance to incorporate the measurement of crucial policy indicators, constructing a new public policy model, and conducting a big data measurement of policy concerns related to the theory and practice of economic system reform since the reform and opening up.

The scholars have conducted a wealth of quantitative and qualitative research on the results of reform and opening up from different directions. Regarding regional economic development trends, there were currently three main theories in the academic community regarding the long-term trend of regional economic differences. The first was the convergence hypothesis. In the long run, the regional economic development gap gradually has decreased and tended to converge. The second was the "inverted U theory" by Jeffrey G. Williamson. Williamson has pointed out that, with economic development, the differences between regional economies have shown an inverted U trend. The differences in regional economic development have shown a gradual increase in the early stage of economic development. At a particular stage of development, the differences in regional economic development will have shown a gradual decrease (the "inverted U" theory). Regional economic development differences will have shown to decrease progressively [1]. Some scholars have also used this theory to explain the changes in European countries' economic situation after the European Union's establishment [2]. The third theory, known as the Amos hypothesis, was proposed by Amos, who found through empirical research that Williamson's theory was somewhat flawed and that regional economic development differences in fact showed a "widening - narrowing --widening-shrinking" trend of "inverted U + U" [3]. Kim has used China's modified GDP variation data to verify the validity of the three basic theories. The study has found that since 1952, China's regional economic development gap has shown a cyclical pattern of "inverted U + U"; since the economic reform and opening up, China's regional economic development gap has mainly shown a narrowing trend, and this convergence trend mainly occurred in 1978-1990; in addition, the regions that played a significant role in the narrowing of the regional economic development gap differed in different periods, with low-development regions playing a substantial role in the 1980s, high development regions in the 1990s, and low-development regions again playing a significant role since the 21st century [4]. Some scholars have argued that the regional economic development gap has gradually widened during China's economic development and will eventually become unmanageable. Xiaoxia Fu has studied the contribution of institutional change to economic growth by constructing institutional change variables. An has used the stochastic frontier analysis developed from the transcendental logarithmic production function to calculate TFP for 30 Chinese provinces and measured economic development fluctuations through TFP [5]. Ding and Li have used the Solow residual model to calculate the growth rate of TFP of China since the reform and opening up [6]. The results of the empirical study have pointed out that institutional innovation at the beginning of the reform and opening up drove China to make significant technological progress in the short term. Still, otherwise, China's TFP growth rate has been on a declining trend since the reform and opening up. Tan and Peng have focused on the impact of financial development on economic growth. They have classified the factors that cause economic growth into quantitative and qualitative categories and use the Ordinary Least Squares (OLS) method to analyse the relationship between financial development indicators and economic growth and its factors with the inclusion of control variables [7]. The quantitative factors have included savings, investment, and capital accumulation, while the qualitative factors include investment efficiency and TFP. Exports and investment mainly have driven economic growth in China, while technological innovation does not contribute enough to China's economic development.

Other scholars have focused on quantitative methods to study the role of institutionalised shocks on economic development in China. Pi has attempted to empirically analyse the impact of short-term economic fluctuations on long-term growth trends in China through the AABM model. The empirical study has shown that China's short-term economic fluctuations and China's long-term economic development trends generally exhibit an inverted U-trend [8]. This trend has been more pronounced in the central and western parts of the country, while in the east, the relationship shows a U-shaped movement. The government has ought to adopt short-term policy interventions in response to economic fluctuations, and in the long run, to esteem the major role in the economy market and reduce administrative interventions. Zhang and Jin have focused on economic cycles and conducted an impulse response analysis of economic cycle fluctuations and transformational institutional shocks through a VAR model, using GDP to measure economic fluctuations. They have proposed that since the reform and opening up, China's economy has been transformed, institutional change shocks having a clear impact on macroeconomic fluctuations. And this impact has been persistent and marginal decreasing but does not have an immediate impact but requires a certain time lag. And among the various types of transformational change, the most significant contribution to output volatility has been made by denationalisation shocks [9]. The process of denationalisation has been an important event in the reform and

opening up, still, some scholars have continued to argue that, apart from economic factors, the judgement of social stability and the sustainability of policies at the time of policy formulation is the reason why the system is not effective in stimulating certain economic indicators, and that these policies often took until economic growth plateaued to reveal themselves [10].

The scientific measurement of national policies is currently at a preliminary stage of exploration due to the complexity of policy content, the difficulty of uniformity of form, and the difficulty of collating and quantifying it. Although some studies have adopted a few quantitative methods, they lack systematicity. Li have constructed a database by collecting 683 technological innovation policies from 1990–2014 in BRICS countries to select 25 variables to scientifically describe technological innovation policies and used a combination of cluster analysis and factor analysis methods to construct and validate a three-dimensional comparative research framework for BRICS technology innovation policies [11]. Accordingly, the layout of technological innovation policies in BRICS countries has been compared. And often, the policies have processed a cross-acting character, which is not only reflected in one policy and event but also recurring, increasing the complexity of the data statistics, especially when it comes to international trade, and the policies in most areas also appear to have a phased distribution [12]. The division of phases has been also universal in analysing the impact of each area on economic development and other policies, with policies in each area showing 3–4 phases in the process of reform and opening up being common and in line with most scholars' classification of the overall phases of reform and opening up [13].

2. Policy informatics-based modelling analysis of multi-source data fusion

Policy informatics has been an interdisciplinary and comprehensive research methodology for increasingly complex public management and public policy issues, based on solutions developed by information and communication technology and data science, integrating policy process change and management modelling innovation, and its basic research methods have included system science simulation, social network analysis, text mining and topic modelling [14]. Due to the complexity of policy content, the difficulty of uniformity of form and the difficulty of collation and quantification, scientific measurement methods for policy are currently at a preliminary exploration stage. Although some studies have adopted a few quantitative methods, they lack systematicity. Based on this, this study takes the full record of China's Reform and Opening Up (1978–2018) (<http://www.reformdata.org/records/>) as the textual basis and builds a policy informatics methodology covering system science simulation, social network analysis, text mining, and complex semantic understanding techniques to cover the period 1978–2018 in China. The database will be used to analyse the impact of policies and events related to the theory and practice of economic system reform on China's economic and social development (later referred to as the "economic development database"). The database covers all important policies, economic events, and related economic data and indicators in the history of reform and can be used to construct a fine-grained measure of policy concerns and policy implementation and to map out the evolutionary trends.

2.1 Construction of the database

The core task of database construction is processing and processing textual information, on economic policies and events, to transform unstructured data into structured data. Based on this, we first evaluate and quantify the content of the database in three perspectives: the attributes of the event and policy, the area of influence, and the scope of influence, which are used to characterise the content of the textual information in the policy text database. Secondly, in the data processing, the same policy and data are presented in different entries, which are compared and filtered by keyword comparison, and the textual information that is not relevant to the theory and practice of economic system reform is cleaned up. In the third step, we classify and annotate the data manually by keywords. For the data packets containing only the date of occurrence and the content of the policy event, we assign values and labels to the entries and data in the database under three perspectives mentioned before. For content attributes, the three broad categories of policy, event and planning are subdivided into six categories: economic policy, non-economic policy, economic event, non-economic event, economic project, and non-economic planning, based on whether they are economically relevant or not. Based on the existing six categories, the six categories are further subdivided into 12 more refined attributes, such as major economic policy, economic policy, major non-economic policy and non-economic policy, to facilitate the determination of attribute values for each entry and data in terms of content attributes,

based on whether they are important and whether their impact is significant. For the areas of influence, we subdivide them into 17 areas based on existing research: finance, trade, livelihood, environment protect, agriculture, education (and competent professionals), culture, administrative construction, high-tech, infrastructure, combat corruption, new-area (development), regional policy, upgrade industry, major events, conference, regulation. Finance includes all pan-financial sectors, including banking, insurance, etc., and trade includes domestic and foreign trade. People's livelihood mainly concerns social welfare, health care, and pensions. Environment protect includes environment protection and sustainable development. Agriculture represents the three rural issues in China, agriculture, farmer, and rural area. Education includes compulsory education and higher education, as well as vocational education and the building of human resources on competent professionals. Culture represents all areas of culture, including but not limited to the arts, tourism, and national culture. Administration represents the optimisation of government structures and the development of laws. High-tech stands for high technology and management technology. Combat corruption stands for the building up of the party and the government and the integrity of institutions and enterprises. New-area represents the various levels and types of new regions established after the reform and opening up. Regional planning represents the planning and development of specific regions. The upgrade industry represents the optimisation of various industries, the elimination of backward production capacity, and urbanisation. The major event represents important or historically significant entries and data. Conferences and regulations correspond to meetings held by various policy and regulation making bodies and regulations and relevant provisions. It is important to note that an entry or data does not only have a single area of influence but often has a combination of attributes. Regarding the sphere of influence, entries, and data are assigned four attributes based on the sphere of influence: international, national, regional, and municipal. Figures 1 and 2 show the categories and the percentage of each domain in the database.

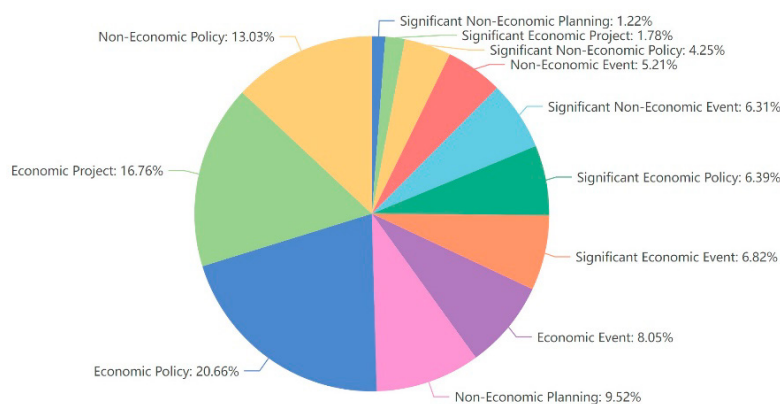


Fig. 1. Percentage share in data attributes distribution

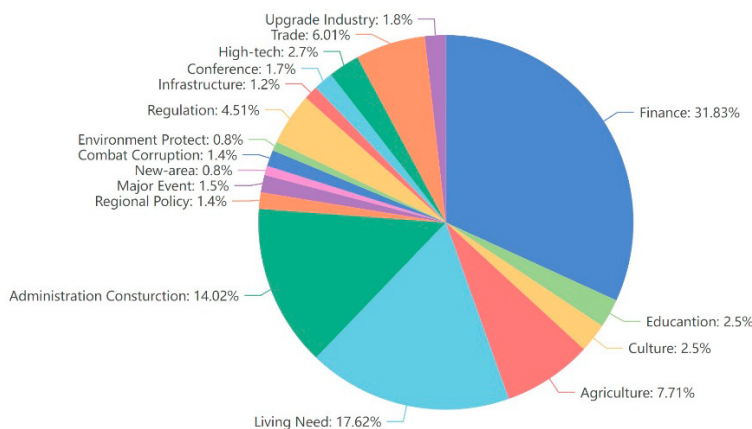


Fig. 2. Percentage share in areas of policy influence distribution

2.2 Construction of feature model

Quantitative research is based on policy database usually integrates computer technology and statistical methods to analyse and generalise large volumes of policy texts, thereby identifying the themes, content and influencing factors of policies. Simultaneously, the implementation of policies is tracked and evaluated through time-series analysis methods to capture the key influencing factors and the changing trends of policy effects.

Presently, quantitative research on policy database mainly includes the following aspects:

1. Policy impact factor research. This research focuses on analysing the impact factors of different policies and exploring the social, economic, and political factors related to policy implementation and the interactions between various variables.
2. Policy effect evaluation research. This research is mainly based on the indicators and data involved in the process of policy implementation to analyse and evaluate the actual effects of policies to provide a reference basis for the government to formulate more scientific and effective policies.
3. Research on policy classification systems. This study focuses on establishing a complete set of policy classification systems and summarising and analysing various reforms and opening up policies according to different categories.
4. Research on policy text analysis. This research mainly adopts computer technology and natural language processing technology to analyse and mine policy texts and identify policies' themes, contents, and influencing factors.

In this study, we investigate the intrinsic links between policy and economic and social development in China's economic reform and opening up process by conducting research on policy influencing factors, policy classification systems, and policy text analysis in the economic development database.

Feature extraction is an essential step in the process of quantitative analysis of the economic development database. By selecting relevant policy factors, including information on policy content, coverage and effects, and transforming them into comparable and computable feature expressions, valuable information can be extracted from the significant policy database for more refined analysis. In the feature extraction process, we combine cleaning, normalisation and standardisation operations to ensure the accuracy and comparability of the data. Simultaneously, we consider the interrelationships between different policy factors, such as synergies between policies and the impact of policies on different groups. We extract the frequency of each area of policy influence and each category and weighted the data according to the importance of the policy and the scope of its influence so as to obtain a comprehensive score for each influence area and construct a characteristic model for analysis. It is important to note that in our policy system classification study, we manually categorise the policies to ensure the accuracy and reliability of the model construction.

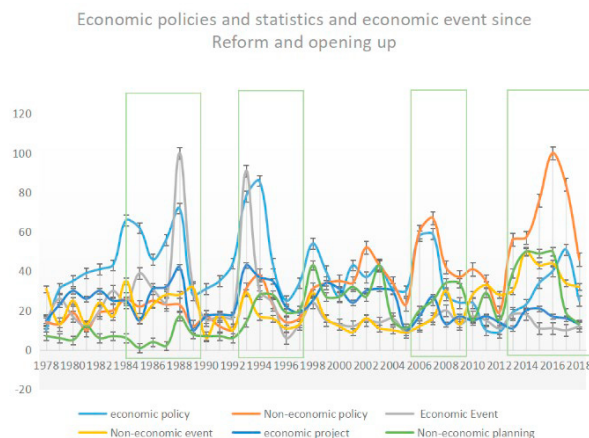


Fig. 3. Economic policies and statistics and economic event since reform and opening up

Based on the distribution of the economic policy, economic project, and non-economic policy, non-economic planning, and non-economic event in the collated database, we can identify four periods of high policy and regulatory promulgation: 1984-1990 (Seventh Five-Year Plan period 1986-1990), 1992-1996 (Eighth Five-Year Plan period 1991-1995), 2006-2010 (Eleventh Five-Year Plan period 2006-2010), and 2014-2018 (Second half of the 13th Five-

Year Plan period 2016-2020)

During the National Economic and Development Plan period and the early years of its commencement, economic policy, events, and plans were introduced and announced at a high level, and national-level policy guidelines set the direction for economic development. Since the reform and opening up, especially since the 21st century, the growth rate of non-economic policy has been much less than that of economic policy, and the frequency of economic policy and plans has soared, indicating a rapid increase in domestic economic market control and attention. The first Five-year Plans after the 21st century led to a rise in economic policy fluctuations dominated by economic policy and event; after the 21st century, it is more economic project that dominates, with planning to facilitate the coordination and co-ordination of the national economy and creating space for the development of an innovative and diversified economic system.

2.3 Empirical analyses

After constructing the characteristic model, this study uses the OLS method to conduct an econometric analysis of the relationship between the areas of policy influence and social development factors in China. We use the composite score (frequency of different categories and calculate the weighted average) of the policy classification as the independent variable and select China's GDP, capital deposit and fixed capital deposit between 1978 and 2018 as the dependent variables, respectively, for the analysis.

The linear regression equation is as follows:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_n X_n + \xi$$

Where β_0 is the intercept, $\beta_1 \dots \beta_n$ are the coefficients of each impact area of the policy, $X_1 \dots X_n$ are the combined scores of each impact area of the policy, and ξ is the error term.

The results of the fit show in the table below:

Table 1. Regression result of areas of policy influence and GDP

Areas of Policy Influence	Coefficient	P-value
Major Events	0.1160342	0.1444102
Administrative Construction	-0.8350148	0.0133346
Finance	-0.2045644	0.1224198
Education	-0.5033204	0.01067434
Culture	-0.9303495	0.1283115
Agriculture	0.2362423	0.0842819
Livelihood	0.52023123	0.0826596
Upgrade Industry	-0.3374771	0.0969572
Trade	0.1114234	0.0693225
High-tech	-0.6677230	0.0760594
Conference	-0.9015475	0.1385778
Infrastructure	0.1192341	0.0597142
Regulation	-0.2261864	0.0514257
Environmental Protect	-0.7549959	0.1601300
Combat Corruption	0.3042324	0.2708622
New-area	-0.5623394	0.0819408

Table 2. Regression result of areas of policy influence and capital deposit

Areas of Policy Influence	Coefficient	P-value
Major Events	0.1550122	0.1280284
Administrative Construction	-0.8476082	0.0246199
Finance	-0.1864722	0.0670563
Education	-0.5328917	0.0125553
Culture	-0.8638614	0.1418154
Agriculture	0.2270123	0.0880482
Livelihood	0.5090124	0.0858148
Upgrade Industry	-0.4700794	0.0982627
Trade	0.1710123	0.0564156
High-tech	-0.7841075	0.0733539
Conference	-0.1242942	0.1475907
Infrastructure	0.1300129	0.0581426
Regulation	-0.2843554	0.0436358

Environmental Protect	-0.9172795	0.1571122
Combat Corruption	0.2140335	0.2993879
New-area	-0.7490852	0.0772345

Table 3. Regression result of areas of policy influence and influence and fixed capital deposit

Areas of Policy Influence	Coefficient	P-value
Major Events	0.1381234	0.1408350
Administrative Construction	-0.8026943	0.0261485
Finance	-0.1827855	0.0672061
Education	-0.5253602	0.0125340
Culture	-0.7416559	0.1452971
Agriculture	0.2152312	0.0896784
Livelihood	0.4454124	0.0909526
Upgrade Industry	-0.2108071	0.0921038
Trade	0.1400123	0.0631905
High-tech	-0.7257209	0.0749026
Conference	-0.1004767	0.1257582
Infrastructure	0.1382134	0.0552968
Regulation	-0.2298240	0.0521879
Environmental Protect	-1.1579242	0.1568111
Combat Corruption	0.1072342	0.2668751
New-area	-0.8127624	0.0779866

The results of statistics show that agriculture, livelihood, trade, and infrastructure are positive numbers, which indicates they have a positive relationship with China's GDP, Capital Deposit, and Fixed Capital Deposit. Because of the country's development phase, the government attaches great importance to these areas throughout the 40 years of opening up, the policy inclination plays an important role in economic growth, which accords with the empirical evidence. However, the results of statistics contain negative numbers for education, environment protection, high-tech, and administrative construction, which has the following reasons: Firstly, the government formulates education and talent attraction policies mainly after the year 2000, the plenty cost of early education investment needs to be spread evenly over 40 years. Moreover, environmental protection policies are set up mainly after 2012, which may conflict with short-term economic goals but will guarantee sustainable development in the long term. In addition, high-tech and administrative construction include reforming state-owned and private enterprises and market economy institution construction, which require plenty of upfront investment, and the influence has a lag effect. To sum up, the early policy-related areas are more likely to be positive numbers, while later policy-related areas are generally negative numbers.

3. Discussion

Using a case study to discuss how to apply the economic and policy development database constructed in this study to a specific case and do the analysis. Taking education (and competent professionals) as an example, the Regulations of the People's Republic of China on Academic Degrees and the Circular on the Enrolment of Postgraduate Students for Doctoral Degrees in 1981 were promulgated in 1980 and 1982, respectively, which are milestones in China's education policy. The span of the policy's effect was 25 years, from 1983, when China produced its first independently trained doctoral students, to 2008, when it overtook the United States to become the world's largest doctoral-granting country. In the three decades between 1978 and 2008, China's education policy and development was more about expanding student numbers and exploring new education systems. Therefore, by 2008, China's higher education showed a trend of large numbers but a declining quality of graduates. In contrast, from 2010 to the present, more emphasis placed on building the quality of education, with emphasis on the all-round development of students at the primary level, and graduate education gradually moving from a phase of outward expansion to internal development, and education policy achieving a shift from quantity and external learning to quality and self-sufficient exploration [15].

Table 4. Regression result of policy in education (and competent professionals) and GDP

Areas of Policy Influence	Coefficient	P-value
---------------------------	-------------	---------

Education (and Competent Professionals)	-0.5033204	0.01067434
Education 1978-2008	-0.5863861	0.00932435
Education 2008-2018	0.2270123	0.21804823

In data, the year 2008 was used as a cut-off point for examining the single policy of education, which had a more pronounced impact on China's GDP between 1978 and 2008; the effect of education policy and planning on GDP was diminished, after 2008. Before 2008, both higher education and compulsory education were lacking, so the effect of policies and plans on economic development was significant. However, the effect of education policy on economic growth continues to not be positive feedback. After 2008, China's education planning was initially realised, and the original model no longer significantly affected the economy, but the effect has become a positive return, so that the process of generating positive feedback from education policy on economic outcomes can be long. This aligns with our general inference of China's education planning, as one of internal development and education policy, achieving a shift from quantity and external learning to quality and self-sufficient exploration.

4. Conclusion

Based on the textual information in the database of China's Reform and Opening Up (1978-2018), and combining multiple types of textual big data from government website, official weblog, social media, etc. This paper constructs a structured database for analysing the impact of policies and events related to the theory and practice of economic system reform on China's economic and social development from the perspective of policy informatics and conducts an empirical analysis of the intrinsic links between the areas of policy influence and the factors of China's social development. In future research, the study will be used as a basis for the analysis of the impact of policies and events on China's economic and social development. Furthermore, in future research, based on the economic and policy development database constructed in this paper, we will further optimise data-based policy adjustment recommendations and policy effect monitoring and iteration mechanisms and propose policy recommendations to promote China's economic development and economic system reform for the reference of relevant policy-making departments.

References

- [1] Williamson, J.G. (1965) "Regional inequality and the process of national development: a description of the patterns", *Economic Development and Cultural Change* 13: 51–84.
- [2] Tirado, D. A., Diez-Mingueta, A., and Martinez-Galarraga, J. (2016) "Regional inequality and economic development in Spain, 1860–2010", *Journal of Historical Geography* 54: 87–98.
- [3] Amos, O. M. (1988) Unbalanced regional growth and regional income inequality in the latter stages of development. *Regional Science and Urban Economics* 18(4): 549–566.
- [4] Kim, S. (2010) "Trends in regional economic development disparities in China and their characteristics - based on GDP revised data", *Nankai Economic Studies* 2010(01). <https://doi.org/10.14116/j.nkes.2010.01.001>
- [5] An, Ke, Shan, Y., and Shi S. (2022) "Impact of Industrial Intelligence on Total Factor Productivity", *Sustainability* 14(21):14535.
- [6] Ding, Chun, and Li, Junyang. (2014) "Analysis over factors of innovation in China's fast economic growth since its beginning of reform and opening up", *AI & society* 29: 377-386.
- [7] Yanzhi, Tan, Wenping, Peng. (2003) "Analysis of factors of financial development and economic growth", *Shanghai Economic Studies* 2003(10): 3-12.
- [8] Pi, J., Zhou, J., Yang, S., and Xing, W. (2017) "Inverted U Shape Relationship between China's Economic Growth and Fluctuations: the Perspective of Leverage Ratio Asymmetrical Change Mechism", *China Soft Science*.
- [9] Jianhui, Zhang, Tao, Jin. (2011) "Transitional institutional shocks and macroeconomic volatility in China: An empirical study based on 30 years of reform and opening-up data", *Journal of Economics* 606(08): 63-69.
- [10] Alexander, Eckstein. (2009) "The quest for economic stability", *China's Economic Revolution*, 159-190.
- [11] Li, F., Zhang, D., Liu, P., and Li, N. (2016) "Construction and Application of the Three-Dimensional Comparative Framework on Technology Innovation Policies: Based on the Analysis of BRICS' Policy Texts", *Science and Technology Management* 37(03): 3-12.
- [12] Hausmann, R., and Hidalgo, C. A. (2011) "The network structure of economic output", *Journal of economic growth*, 309-342.
- [13] Zhang, S., Liu, Y., and Huang, D. (2020) "Contribution of factor structure change to China's economic growth: evidence from the time-varying elastic production function model", *Economic Research-Ekonomska Istraživanja* 33(1): 2919-2942.
- [14] Desai, A., and Kim, Y. (2015) "Symposium on policy informatics", *Journal of Policy Analysis and Management* 34(2): 354-357.
- [15] Qi, Z., and Du, Y. (2022) "What is a Good Evaluation of the Effect of Education Policy", *Journal of East China Normal University* 40(2): 29-42.