A Survey of Regional Economic Research Based on Domain Big Data

Hongwu Qin[a], Yajun Li[b,*], Zhenglin Yang[c], Xiuqin Ma[d]

College of Computer Science & Engineering, Northwest Normal University, Anning District, Gansu, China

[a] qinhongwu@nwnu.edu.cn, [b] 1932900050@qq.com, [c] 1074335089@qq.com, [d] maxiuqin@nwnu.edu.cn

Keywords: Big data, domain big data, regional economy

Abstract: Regional economy is the microcosm of national economy and the basis of steady growth of national economy as a whole. Its traditional research mainly relies on the way of economic census to obtain the corresponding economic indicators, but the whole process is costly, time-consuming and low real-time. At present, the characteristics of economic digitization in China are becoming more and more prominent. Big data has gradually become an important driving force in the study of economic situation because of data availability, data collection, real-time and abundant data. Regional economic research based on big data has become one of the research hotspots. This paper summarizes the achievements of China's regional economic research based on big data from five aspects of Internet, social interaction, education, transportation and tourism, and probes into the existing problems and future development trends.

1. Introduction

The concept of regional economy originated from the agricultural location theory proposed by the German economist Thunen[1]. As the epitome of national economy, its development is closely related to national prosperity and people's life, which is the foundation and key to the overall steady growth of national economy. Regional economic research is of great practical significance to promote the development of regional economic integration and the high-quality development of regional economic in China[2]. In traditional economic research, people mainly obtain data by economic census and other ways, and then conduct economic analysis on the obtained samples, structured data and statistical data. The whole process is time-consuming, costly and the real-time is low. The lagging data makes the research results difficult to meet the real-time demand of economic decision-making, which seriously restricts the development of China's economic research. At present, big data research has involved various fields. Its application prospect in economic research has attracted extensive attention of scientific researchers and become the preferred method to solve economic-related problems[3], making cross-field research possible[4].In recent years, the research on regional economy with big data has become a focus in the era of digital economy. Compared with the traditional economic research, the biggest advantage of this method lies in the large number, timeliness, wide coverage, low cost and high authenticity of the data. Researchers can use these data to conduct a certain degree of empirical, statistical correlation analysis and provide valuable reference for economic policymakers.

The paper of [3] summarizes the economic situation in the context of big data from both macro and micro perspectives. The papers of [5-6] summarize the theoretical formation and development of regional economy, the quality of development, and evaluation methods, and point out the shortcomings of current evaluation methods. The papers of [7-9] review the relevant research on the imbalance of regional economic development in China. The papers of [10-11] review the relevant research on the relationship between education and regional economy. The papers of [12] reviews the relationship between tourism and regional economy, which provides new directions and ideas for researchers in the future. The articles of [13-15] depict the methods of predicting macroeconomics with big data and the methods of processing unstructured data. At present, the development of big data has become a national strategy. The report of the 19th National Congress of the Communist Party of China states that it is necessary to accelerate the development of the deep integration of big

data and real economy. However, from the existing literature, there is no review literature on China’s regional economic research based on big data for the different domains. This paper comprehensively analyzes and summarizes the achievements of China’s regional economic research based on the five fields of Internet, social networking, education, transportation and tourism. It also points out the existing problems and possible research directions in the future.

2. Typical fields and Regional Economic Research

At present, researchers mainly use big data in the fields of the Internet, social networking, education, transportation, and tourism to carry out China’s regional economic research. We will describe them separately below.

2.1. Internet

With the contemporary Internet technology and the rapid development of China's economy and society, the digital technology represented by the Internet is accelerating its deep integration with the economy and society[16] and has become an important engine to promote China's economic development. This paper[16] combines the panel data and spatial econometric models of 31 provinces and cities in China to analyze the relationship between the Internet and regional economic development level and structure, and finds that there is a significant interaction effect between the Internet penetration rate and regional economic development.

The paper[17] combines the spatial panel data of 31 provinces (districts and cities) in China from 2005 to 2015 and analyzes the mechanism of the Internet's role in China's regional economy by using the Dubin model, and find that the Internet has a significant role in promoting regional innovation. The paper[18] uses panel data from 31 provinces to empirically analyzes the relationship between the Internet and economic growth, and finds that Internet development has a significant role in promoting economic growth. The role of the Internet in promoting economic growth in eastern regions is greater than in central and western regions. Literature [19] studies the relationship between Internet and regional economic development by using time series data from 31 provinces, districts and cities (excluding Hong Kong, Macao and Taiwan province) for ten consecutive years, and finds that is was a significant linear correlation between the two. Literature [20] constructs the Internet resource index, using the production function model to measure the Internet's contribution to GDP growth, finds that the rapid development of the Internet has driven and promoted China's economic growth.

2.2. Communication

Everyone has their own social network (both online and offline). Relying on the Internet platform, the communication between people has become more efficient and convenient while expanding social circles. The prosperity of social economy comes from the gathering of human activities, and the proportion of social network in contemporary economic development is gradually increasing[21]. This simple and effective method of inferring economic development provides a new way for real-time prediction of macro-economy in the future. Therefore, it is of great significance to reveal the relationship between social and economic through the exploration of online social.

The current research on social and economics is mainly to analyze the online social data to achieve zero-cost economic development. Literature [22] analyzes the data of China's Sina Weibo and the National Bureau of statistics, investigates the potential relationship between social activities and economic development, and finds that there is a strong correlation between online social activities and local economic development and industrial structure. Literature [23] analyzes the relationship between the national communication network records and the community economic welfare data of the national census, and finds that the diversity of the relationship between individuals is closely related to the development of the community economy. Literature [24] analyzes the online social network data and the anonymous resume data of the higher education degree candidates, and finds that the structural characteristics of the two are related to the economic status, among which the relatively small rules model's talent flow network has a stronger ability to

predict GDP. Literature [25] analyzes the data of millions of Tencent users, and reveals the relationship between the activity on QQ and the urban economic indicator GDP.

2.3. Education

In the field of education, regional economic research based on big data mainly reflects in two aspects: talents and universities, which are described below.

2.3.1. Talent

The actual situation of China's regional development shows that economic growth and the development of scientific and technological talents are a mutually beneficial process. However, with the rapid development of China's economy, the gap between regional economies and the imbalance in regional talent resources are becoming increasingly serious. On the one hand, talents are more inclined to travel to coastal cities in the east; on the other hand, the relatively slow development of the central and western regions has also become a brain drain. Does the imbalance of regional talents affect the development of regional economy directly or indirectly? Literature [26] analyzes the influence of talents on regional economy by constructing a multivariable linear regression model, and points out that in order to promote the development of regional economy, China needs to allocate scientific and technological talents reasonably, coordinate the coordinated development of the central and western regions, and then narrow the economic gap between regions. Literature [27] analyzes the relationship between talents and regional economic development, and gives some suggestions for obtaining high-quality talents to promote regional economic development. Based on the 2018 yearbook data of Ningbo City, literature [28] verified the conclusion that the accumulation of scientific and technological talents is conducive to the development of regional economy through the design of talent aggregation model.

Literature [29] conducts an empirical study on the impact of human capital on China's economic growth, and finds that higher education plays a greater role than primary and secondary education in the economic impact of re-education, and the more economically developed regions benefit more from higher education.

2.3.2. Universities

Talents, as the link between the economy and universities, are the most active factors in economic and social development. Universities, as important resources in the country's economic and social development, are the basic forces for regional economic and technological development[30]. With the continuous development of big data concept and economic society, big scholarly data has been widely recorded and collected[31], and gradually applied to economic research. In terms of the research on the relationship between universities and regional economy, literature [32] analyzes the advantages and constraints of universities serving the local economy, studies the relevant mechanisms suitable for universities' talents serving the economic and social development, and puts forward the effective ways for universities to serve the local economic and social development. Literature [33] based on the relevant index data of scientific research input and economic growth of regional colleges and universities in China, this paper analyzes the contributions of talent training, scientific research and production and cultural construction in colleges and universities to regional economic growth, and finds that higher education can basically keep pace with regional economic growth. In terms of the coordination degree with regional economic development, the national coordination status presents a normal distribution. In the research direction of university financial investment and regional economy, literature [34] empirically analyzes the scale of universities and educational expenditure in each region, and finds that there is a significant correlation between regional economic development and local university development. Literature [35] uses panel data to establish a dynamic VECM model to analyze the synergy between local higher education financial investment and economic growth, and finds that there is a unidirectional short-term or long-term relationship between regional economic development and local higher education financial input. The emergence of big data enables people to make more scientific and reasonable educational decisions based on the results of experimental

data, and promotes the development of regional economy better served by colleges and universities.

2.4. Traffic

"If you want to get rich, build roads first". As an important part of the development of the national economy, transportation plays a solid role in safeguarding economic and social development. The relationship between it and regional economic development is an issue that economists focus on from beginning to end. The "The Belt and Road" strategy has greatly promoted the development of China's road traffic. Whether there is an inevitable connection between the national economic progress and the development of road traffic and what kind of relationship exists between economic and social development and the development of road traffic have become the focus of the majority of scholars. The convenience of transportation gave birth to the transformation of the logistics industry. Liu Guangdong[36] and others basing on the data of Anhui Province on the impact of regional logistics on the regional economy, and using Geoda and Stata software to perform spatial statistics and space analysis on regional logistics and regional economy econometric analysis explores the spatial distribution and the relationship between logistics and regional economy.

As part of China's infrastructure construction, high-speed railways have become a powerful engine for the country's economic and social development. The construction of the high-speed railway has led to the rapid development of a series of secondary industries, as well as the development of urban tourism and other tertiary industries along the route. In the relevant research on the impact of high-speed rail on regional economic development, it is mainly reflected in the relationship between transportation and regional economic development. For example, literature [37] takes prefecture level cities as the research object, introduces the accessibility index that decreases with the increase of travel time and ticket price, empirically studies the impact of high-speed rail on regional economic growth, and finds that high-speed rail has a positive impact on urban GDP growth. Literature [38-39] based on urban panel data, combining with panel quantile regression method and DID evaluation method, explores the impact of high-speed rail on regional economic development, and finds that there is a significant impact between them. The development of high-speed rail helps to narrow the gap between innovation levels between cities. Literature [40] uses spatial syntax and gravity model to analyze the spatial and temporal evolution pattern and correlation of road network accessibility and economic connection in different cities of Wuhan, and concludes that there is a significant correlation between the two.

2.5. Tourism

In the digital information era, the rapid development of artificial intelligence has made smart tourism and digital tourism become the main development direction of the current tourism industry. With the continuous improvement of people's living standards, tourism is gradually becoming an important part of consumer demand, and its development plays an important role in promoting regional economic development. It is of great significance to analyze and explore the relationship between big data of this domain and regional economy, and to provide decision-makers with suggestions by data support, for the development of national tourism and economic growth in the future. Literature [41] analyzes the causal relationship between national tourism income and economic growth, and finds that there is a one-way causal relationship between tourism income and GDP growth in China.

Literature [42] based on statistical yearbook data of Anhui Province for ten consecutive years, carries out a regression analysis on the relationship between tourism and regional economic development, and finds that there is a mutually reinforcing relationship between the two. Based on the data of China's cities, literature [43] investigates the dynamic relationship between tourism and economic growth, and finds that there is a long-term equilibrium relationship between them. Tourism plays a significant role in promoting economic growth, and its role in urban economic growth shows certain regional differences.

3. Research Methods and Data Sources

This section summarizes the research methods and data sources commonly used in regional economic research based on big data.

3.1. Research Methods

(1) Regression analysis

Regression analysis is a predictive modelling method that can reveal the significant relationships between dependent variables (goal) and independent variables (predictor). At present, the main regression analysis methods used in regional economic research based on big data include support vector regression [22], logical regression [42] and polynomial regression [26].

(2) Stability test

The stationarity test is the most basic method to judge the stationarity of time series. At present, the stationarity test method mainly used in regional economic research based on time series data is the unit root test [16].

(3) Accessibility study

Accessibility [37-38][44] refers to the possibility of interaction between various subsystems in the transportation network [44], which can reflect the convenience of transportation and is mainly used in regional economic research related to transportation.

(4) VECM model

The VECM models transform unstable time series into stable ones by differential methods. The VECM model [35][43] is mainly applied to regional economic research based on panel data of various fields.

(5) Correlation analysis

Correlation analysis [22][28][35] is a method to analyze the degree of closeness between two or more variables, which is mainly applied to the research on the correlation between domain big data and regional economy.

3.2. Data Sources

At present, the main data sources in regional economic research based on domain big data are shown in Table 1. Data mainly comes from the statistical yearbook, and research data in various fields are respectively from the census data of the Ministry of transport, QQ (an instant messaging software service in China) and Sina Weibo (a Chinese microblogging website) data, world bank data and panel data in various fields. Scholars select the corresponding attributes of data in these fields for regional economic research.

Table 1 Summary of major data sources.

Data sources	Internet	Social	Education	Traffic	Tourism
China Statistical Yearbook	√	√	√	√	√
City Statistical Yearbook	√	√			√
China Railway Yearbook				√	
Atlas of the People's Republic of China				√	
Ministry of transportation transportation census data				√	
Panel data	√	√	√	√	√
Sina weibo data		√			
Tencent QQ data set		√			
World bank data set					√
University ranking data			√		
Civil aviation passenger relations				√	

4. Existing problems and future research directions

Although some achievements have been made in China's regional economic research based on big data of the different domains, there are still some problems to be solved:

(1) Existing research work mainly focuses on the research on the correlation between big data of the different domains and regional economy, and less work on causality research;

(2) At present, the publicity of big data is low, and data in some fields are difficult to obtain;

(3) The research data is relatively single, and the existing research work mainly studies regional economy based on the big data of a single field, which leads to the lack of comprehensive regional economic analysis.

In view of the above problems, we think the future research direction can be summarized as follows:

(1) Explore more methods or models to mine the causal relationship between big data and regional economy, and find the dominant factors affecting regional economy, so as to provide more accurate opinions for decision makers;

(2) Explore more open source data in more fields, such as big data of science and technology, and try to apply it to regional economic research;

(3) Integrate data from multiple fields for regional economic research and improve the comprehensiveness of regional economic analysis.

5. Conclusion

This paper reviews Research results, research methods and data sources of domestic regional economy based on big data from the aspects of Internet, social, education, transportation and tourism based on the literature reviewed. At the same time, the problems existing in the current research are put forward, and the future research trends are prospected. It is hoped that this paper can be helpful to researchers in related fields and policy makers.

Acknowledgements

The National Science Foundation of China (No. 61662068, 61662067) provides the support for this research work.

References

[1] Y Zhang, (2003) Western regional economic review. *Contemporary Finance & Economics, (12)*, 45-46.

[2] G B Liu, J Z Song, (2019) Research on high quality development of China's regional economy. *Regional economic review, (02)*, 55-60.

[3] J Gao, T Zhou, (2016) Big data reveal the status of economic development. *Journal of University of Electronic Science and Technology of China, 45(4)*, 625-627.

[4] Einav, L., & Levin, J. (2014). Economics in the age of big data. *Science*, *346*(6210), 1243089.

[5] Y Zhang, (2004) Western regional economic review. *Journal of Theoretical Reference, (07)*, 45-46.

[6] Q Qi, Z H Qiang, H X Deng, (2019) A survey on the quality of regional economic development in China. *The Science Education Article Collects, (01)*, 181-182+189.

[7] Z B Yang, X F Zeng, (2010) Study on China's regional economic disparities summary. *Economic Geography, 30(06)*, 899-905.

[8] A Z Wu, K Z Yang, G P Li, (2011) Review on regional economic disparities in China. *Economic Geography, 31(05)*, 705-711.

[9] S Q Liu, B Zhou, (2008) Review on Chinese regional economic development gap. *On Economic Problems, (01),* 53-55.

[10] J X Lei, (2011) Review on the regional economic and social development of the higher vocational education service. *Chinese Vocational and Technical Education, (36),*54-58.

[11] Q L Chao, X L Li, (2007) The theory summarize of the interactive development of higher education and regional economics. *Open Education Research, (03),* 18-21.

[12] W L Zhou, (2011) A review on impact of tourism on economic growth. *Economic Geography, 31(08),* 1402-1408.

[13] T X Liu, X F Xu, (2015) A summary of big data and macroeconomic analysis research. *Foreign Theoretical Trends, (01),* 57-64.

[14] H J Li, D Shi, L M Ma, (2018) Economic research based on big data method: frontier progress and research overview. *Economist, (06),* 96-104.

[15] Z Su, Q Zhang, (2015) Researches on socio-economic forecasting based on internet big data. *Journal of Intelligence,34(07),* 18-21.

[16] Z J Liu, W J Chen, (2017) An empirical study on the relationship between internet development level and economic growth in China. *Economic Geography, 37 (08),* 108-113+154.

[17] X L Zhang, J C Shi, X D Li, H X Zhang, (2017) The mechanism and effect of the internet on regional innovation in China. *Economic Geography, 37(12),* 129-137.

[18] C Zhang, (2017) Internet development and economic growth: mechanism and empirical study. *Journal of Finance and Economics, (07),*32-36+63.

[19] J Lin, M F Wang, (2014) Research on relationship between internet resources and regional economic development. *Resource Development & Market, 30(03),* 321-325.

[20] Y Liu, (2010) Quantitative analysis of influence of internet on national economy. *Journal of Central University of Finance & Economics, (12),*44-49.

[21] Zhu Z. F. (2018) Research on the development of social network economy and shared economy model. *Modern Business Trade Industry, 39(07),* 90-91.

[22] J H Liu, J Wang, J Shao, & T Zhou. (2016) Online social activity reflects economic status. *Physica A: Statistical Mechanics and its Applications, 457,* 581-589.

[23] Goddard, J. B., & Chatterton, P. (1999). Regional Development Agencies and the knowledge economy: harnessing the potential of universities. *Environment and planning C: Government and Policy*, *17*(6), 685-699.

[24] J Wang, J Gao, J H Liu, D Yang, & T Zhou (2019). Regional economic status inference from information flow and talent mobility. *EPL (Europhysics Letters)*, *125*(6), 68002.

[25] X L Ren, Y Y Zhu, S Y Wang and L Hao, (2015) Online social network analysis and the relation with regional economic development. *Journal of University of Electronic Science and Technology of China, 44(05),* 643-651.

[26] J Sun, X K Jiang, (2014) A study on the difference of the influence of scientific and technological talents on regional economic development: a comparative analysis based on the regional data of east, middle and west. *Social Sciences in Guangdong, (02),*15-21.

[27] D P Hua, Y Y Lu, (2007) Talent and regional economic development. *Science&Technology Ecnony Market, (2),* 76.

[28] W Yu, (2019) A study on the influence of the gathering of scientific and technological talents on the development of regional economy in Ningbo City. *Technology and Economic Guide, 27(18),* 190-191.

[29] C G Zhang, L H Zhuang, (2011) The composition of human capital and economic growth: Evidence from China using dynamic panel data analysis. *China Economic Review, 22(1),* 165-171.

[30] D D Wu, (2016) research on knowledge overflow mechanism and effect of Chinese college research and development activities. *University of Science and Technology of China.*

[31] Z W Sun, X P Li, (2018) Research on the application of educational big data from the perspective of system integration. *China Educational Technology, (11),* 111-117.

[32] X S Yang, Y Zhao, Y Li, (2019) Exploration on the mechanism and method of local universities servicing the regional economy. *Journal of Chengdu University of Traditional Chinese Medicine (Educational Science Edition), 21(01),* 106-109.

[33] X T Wang, (2017) A quantitative study of colleges and universities comprehensive contribution of to the regional economic growth. *Soochow University.*

[34] Q Z Yan, W Y Miao, (2006) An empirical study on the unbalanced development of regional higher education and economy. *Research in Education Development, (23),* 62-65.

[35] Y L Tao, Q Li, (2015) Synergy between financial investment in regional higher education and regional economic development: a spatial causality analysis based on provincial panel data. *Fudan Education Forum, 13(01),* 74-81.

[36] G D Liu, T J Yang, X M Zhang, (2019) Spatial econometric analysis of the impact of regional logistics on regional economy. *Statistics & Decision, 35(20),* 137-140.

[37] J S Wu, (2018) Study on the impact of high-speed rail construction on regional economic development. *Price: Theory& Practice, (05),* 135-138.

[38] X Zhang, (2019) The impact of high-speed railway on regional economic innovation and development. *China Journal of Commerce, (01),*167-168.

[39] M Xu, (2019) Research on the impact of high-speed railway on regional economic development. *Zhengzhou University.*

[40] K L Zha, M J Peng, Y F Liu, Y L Song, L Y Zhang, (2017) Spatio-Temporal evolution and correlation analysis of road network accessibility and economic contact in Wuhan metropolitan area. *Economic Geography, 37(12),* 74-81+210.

[41] Sokhanvar, A., Çiftçioğlu, S., & Javid, E. (2018). Another look at tourism-economic development nexus. *Tourism management perspectives, 26,* 97-106.

[42] Y Li, (2018) Study on the relationship between regional economy and tourism industry. *Journal of Heihe University, (26),* 97-106.

[43] B Wang, Y Chen, (2017) The development of domestic tourism and regional economic growth-an empirical study of 285 local-level cities in China. *Inquiry into Economic Issues, (12),* 62-72.

[44] B Wang, Y Chen, (2017) Domestic tourism development and regional economic growth -- based on empirical research of 285 prefecture-level cities in China. *Inquiry into Economic Issues, (12),* 62-72.

[45] Y Yang, W D Cao, Z Q Yao, D P Zhao, H M Yu, Y W Yu, (2019) Study on influence of Lanzhou-xinjiang high-speed railway on accessibility and economic contact in north west area. *Human Geography, 34(01),*131-139.