Information Systems and Economics (2023)
Clausius Scientific Press, Canada

DOI: 10.23977/infse.2023.040905
ISSN 2523-6407 Vol. 4 Num. 9

Measurement and Evaluation of Urban-Rural Integration Level in Various Provinces of China and Analysis of Comprehensive Situation

Li Zhaoying, Wang Yilin

College of Mathematics and System Science, Shandong University of Science and Technology, Qingdao, Shandong, China

Keywords: Urban-rural integration, rural urbanization, strategic adjustment

Abstract: In the trajectory of China's urban-rural integration process, there exists a significant development disparity between the southeast coast and the northwest region. This urban-rural dichotomy has become a salient factor in the assessment of rural urbanization dynamics. At the national level, there's an imperative to propel technological advancements, foster economic acceleration, and enact policy recalibrations to catalyze financial upliftment of the underserved regions. It is incumbent upon local governing bodies to strategically harness endogenous resources, assimilate contemporary agricultural methodologies, and drive economic innovation. The symbiotic relationship between cities and rural landscapes is quintessential; urban centers ought to augment their service delivery and infrastructural footprint, whilst rural domains should leverage intrinsic industrial competencies to attract and retain talent. This paper provides a comprehensive discourse on integration mechanisms, underscoring the indispensability of collaboration and inventive approaches.

1. Introduction

A predominant source of societal tension in China emanates from escalating aspirations juxtaposed against disparate developmental outcomes. The urban-rural integration framework, delineated in the 14th Five-Year Plan, seeks to redress this by facilitating a bi-directional flux of productive factors and ensuring egalitarian allocation of public resources. China has manifested tangible progress in mitigating urban-rural divides, as demonstrated by a convergence in income ratios and augmented balances in agricultural loans[1]. However, impediments such as asymmetrical distribution of public services and persistent resource variances between urban and rural sectors continue to be obstinate challenges[2]. The endeavor of urban-rural integration is not merely instrumental for reconciling extant societal dichotomies but is also a linchpin for broader modernization aspirations. Metrics such as per capita GDP and educational differentials serve as evaluative barometers for integration efficacy, and employing game theoretic analyses within this milieu can elucidate dynamics of urban-rural synergistic development, proffering implications for both theoretical constructs and pragmatic applications[3].

2. Data system and research design

2.1 Data introduction

Drawing from authoritative sources such as the China Statistical Yearbook, China Urban Construction Statistical Yearbook, China Rural Statistical Yearbook, China Health Statistical Yearbook, and China Environmental Yearbook, we curated a comprehensive dataset elucidating urban-rural integration across 31 provincial entities. This dataset encompasses 31 rows and 60 columns, encapsulating multifaceted dimensions including, but not limited to, per capita GDP, delta in income and expenditure between urban and rural demographies, per capita consumption metrics, and the fraction of agricultural expenditure[4]. During the data assimilation phase, a lacuna was identified pertaining to the count of beds in rural health institutions in Shanghai for the year 2021. To rectify this void, we elected to employ a predictive model using bed count data from Shanghai spanning 2011 to 2018. Subsequent to this analytical intervention, our estimation deduced the bed count in rural health institutions in Shanghai for 2021 to be approximately 3,200.

2.2 Indicator system

The metric system is bifurcated into dual hierarchical strata. The primary-tier indicators primarily serve the purpose of delineating evaluative trajectories pertinent to specific dimensions of urban-rural integration. These primary indicators are segmented into four quintessential facets: economic progression, societal advancement, cultural and educational proliferation, and spatial optimization. Subsequently, the secondary-tier indicators are meticulously crafted to cater to these four domains, providing granular and specific evaluative metrics. For an exhaustive elucidation of the indicator content, computational methodologies, metric units, and directional orientation, one may refer to Table 1.

Table 1: Establishment of urban-rural integration indicators

Level 1 indicators	Secondary indicators	Indicator Interpretation	Index unit	Indicator direction
economic development	GDP per capita	Regional GDP / Regional Population	ten thousand yuan	Forward
	Income gap between urban and rural residents	|Urban Per Capita Income/Rural Per Capita Income -1 |	1	Negative
	The gap between urban and rural residents' expenditure	|Urban per capita expenditure/Rural per capita expenditure -1|	1	Negative
	per capita consumption level	Total retail sales of consumer goods/population	Yuan	Forward
	Agricultural expenditure share	Total Agricultural Expenditure/General Public Budget Expenditure	1	Negative
	The added value of non-agricultural industries as a percentage of the region's total output	Output value of secondary and tertiary industries/Gross regional product	1	Forward

Level 1 indicators	Secondary indicators	Indicator Interpretation	Index unit	Indicator direction
	value			
Social development	urbanization rate	Urban Population/Total Population	1	Forward
	Gap between urban and rural medical beds	|Number of rural medical beds/Number of urban medical beds|	1	Negative
	Share of Social Security and Employment Spending	Social Security and Employment Expenditures/General Public Budget Expenditures	1	Forward
	The gap between urban and rural per capita expenditure on education, culture and entertainment	|Rural per capita education, culture and entertainment expenditure/urban per capita education, culture and entertainment expenditure |	1	Negative
	General public service expenditure as a share of total	General Public Service Expenditure/General Public Budget Expenditure	1	Forward
cultural education	Per capita investment in education	Per capita public education investment	Yuan	Forward
	Compulsory Education Level Gap	Work | Rural students/urban students - 1 |	1	Negative
	Gap in Compulsory Education Conditions	Work | Rural Student-Teacher Ratio® / Urban Student-Teacher Ratio - 1 |	1	Negative
	Proportion of Education Expenditure	Education Expenditure/General Public Budget Expenditure	1	Forward
Use the space	Ratio of change rate of cultivated land to construction land	(t_2 - t_i) ②Cultivated land area / t_i cultivated land area (t_2 - t_i) construction land area / t_i construction land area	1	Forward
	Cultivated land and construction land efficiency Ratio	Primary industry GDP / arable land area Secondary and tertiary industry GDP / urban construction land area	1	Forward
	road network density	National road, provincial road, railway length/total area of the region	1/ km	Forward

① t_i and t_2 are the data of 2017 and 2022 respectively, with a time gap of 5 years.

② The student-teacher ratio is defined as: the ratio of teachers to students/the student-teacher ratio stipulated by the Ministry of Education. Among them, the Ministry of Education stipulates that the student-teacher ratio is 1:19 for primary schools, 1:13.5 for junior high schools, and 1:12.5 for high schools.

3. Measurement and evaluation of regional urban-rural integration level

3.1 Evaluation of regional urban-rural integration level based on simulated annealing optimized projection pursuit model

3.1.1 Evaluation of urban-rural integration level

Comprehensive application of dimensionality reduction technology and projection pursuit method.

3.1.2 Data preprocessing

Prior to the formulation of the evaluation metrics, it is imperative to undertake preliminary data preprocessing, encompassing procedures such as data normalization and standardization. Based on the intrinsic characteristics of the indicators, the methodologies adopted for data normalization and standardization can be categorized into four distinct modalities:

①Positive indicators (the bigger the indicator data, the better):

$$x_i = \frac{x_i - \min x_i}{\max x_i - \min x_i}$$

②Negative indicators (the smaller the indicator data, the better):

$$x_i = \frac{\max x_i - x_i}{\max x_i - \min x_i}$$

③Intermediate indicators (the closer the indicator data is to a certain value, the better):

$$x_i = 1 - \frac{|x_i - x_{\text{best}}|}{\max |x_i - x_{\text{best}}|}$$

④Interval indicators (the closer the indicator is to a certain interval, the better):

$$x_i = \begin{cases} x_i = \dfrac{a - x_i}{M}, x_i < a \\ 1 \quad , x_i \in [a,b] \\ x_i = \dfrac{x_i - b}{M}, x_i > b \end{cases}$$

Among them, $M = \max\left(a - \min\left(x_i\right), \max\left(x_i\right) - b\right)$.

3.1.3 Model establishment

i- th indicator on the one-dimensional sample space is calculated :

$$Z_i = \sum_{j=1}^{m} x_{ij} a_j$$

Among them, m is 18, which refers to 18 indicators. For these 18 indicators, there are a total of n=31 pieces of data. Let the projection value of the data be Z i, and then construct the objective function of projection pursuit as:

$$\max Q(a) = S_a * D_a$$

The standard deviation S a of the projected value Z i is:

$$S_a = \sqrt{\sum_{i=1}^{n} \frac{\left(Z_i - \bar{Z}_i\right)^2}{n-1}}$$

The local density D a of projected values Z i is:

$$D_a = \sum_{i=1}^{n} \sum_{j=1}^{n} \left(R - r_{ij}\right) f\left(R - r_{ij}\right)$$

In the local density is a unit step function, and its function is that when the value is greater than 0, the value is 1, and when the value is less than 0, the value is 0; R is the window aperture for calculating the local density; r is the distance formula:

$$r_{ij} = \left|Z_i - Z_j\right|$$

To procure the most accurate representation of the original dataset, we employ the simulated annealing algorithm to optimize the objective function and ascertain the optimal weighting coefficients. The algorithm commences with predefined initial conditions, specifically an initial solution and a stipulated starting temperature. For each iteration within the process, the variation Δ of the objective function is derived via adjacent solution modifications.

When $\Delta < 0$, signifying that the newly generated solution is superior to the previous one, the solution is updated. Conversely, if $\Delta \geq 0$, an ostensibly suboptimal solution is accepted with a designated probability, a mechanism designed to extricate the algorithm from potential local minima or maxima. As the algorithm progresses and the pseudo-temperature diminishes, the likelihood of accommodating a less favorable solution concomitantly wanes. This iterative refinement allows the algorithm to gravitate towards either a global or a proximate optimal solution. Once predefined termination criteria are met, the iteration sequence is halted, culminating in the output of the final evaluative metric. After optimization according to the above method, the evaluation value of the solution projection can be listed:

$$Z_i^* = \sum_{j=1}^{m} a_j^* \cdot x_{ij}$$

3.1.4 Result solution

Use MATLAB to program equations (1) to (10), set the initial temperature of simulated annealing to 100 degrees, and iterate 100 times. Use the suiji function to generate random initial values to avoid local optima. After the program runs, Table 2 shows the weight values of each indicator:

According to the principle of projection pursuit, the urban-rural integration of each region is calculated by using the weight value of each index evaluated level value, the urban-rural integration development level of each city can be obtained (Table 3), and the urban-rural integration development level of each city can be made.

Table 2: The weight value of each index in the evaluation of urban-rural integration

index	Weights	index	Weights	index	Weights
GDP per capita	0.1803	urbanization rate	0.2395	Compulsory Education Level Gap	0.2761
Income gap between urban and rural residents	0.3431	Gap between urban and rural medical beds	0.0531	Gap in Compulsory Education Conditions	0.0809
The gap between urban and rural residents' expenditure	0.3499	Share of Social Security and Employment Spending	0.1327	Proportion of Education Expenditure	0.2868
per capita consumption level	0.1961	The gap between urban and rural per capita expenditure on education, culture and entertainment	0.1499	Ratio of change rate of cultivated land to construction land	0.0279
Agricultural expenditure share	0.3721	Proportion of general public service expenditure	0.1723	Ratio of cultivated land to construction land efficiency	0.3641
The added value of non-agricultural industries as a percentage of the region's total output value	0.1460	Per capita investment in education	0.3102	road network density	0.0681

Table 3: Projection value of urban-rural integration level

area	projected value	area	projected value	area	projected value	area	projected value
Beijing	2.1606	Shanghai	2.1215	Hubei	2.0401	Yunnan	1.391
Tianjin	2.1691	Jiangsu	2.2634	Hunan	1.8918	Tibet	1.0627
Hebei	1.8136	zhejiang	2.3027	Guangdon	2.3221	Shaanxi	1.7589
Shanxi	1.7018	Anhui	1.789	Guangxi	1.7962	Gansu	1.2329
Inner Mongolia	1.6579	Fujian	2.2639	Hainan	2.027	Qinghai	1.4813
liaoning	1.923	Jiangxi	1.8648	Chongqing	2.048	Ningxia	1.5501
Jilin	1.7458	Shandong	2.0589	sichuan	1.7991	Xinjiang	1.3724
Heilongjiang	1.6436	Henan	1.8216	Guizhou	1.5511		

3.2 Identification of dominant factors of regional urban-rural integration indicators based on random forest

When analyzing the regional urban-rural integration level, we aim to determine the most critical urban-rural integration evaluation index, namely the dominant factor. Considering that there are 18 data dimensions, the random forest method was chosen because it can effectively handle high-dimensional, non-linear, collinear and interactive data and provide importance scores for each variable. For the variables X 1 to X 18, we use the Gini index to evaluate their importance. By constructing the statistic $VIM_j^{(\text{Gini})}$, which represents the average change in impurity of the jth variable in all trees of the random forest , the Gini importance of this indicator in the random forest is defined.

$$VIM_j^{(\text{Gini})} = \frac{1}{n}\sum_{i=1}^{n} VIM_{ij}^{(\text{Gini})}$$

Using Python, we import the dataset and split it by 70% training and 30% testing to train a random forest model and evaluate feature importance. Next, we use the RandomForestClassifier function in Python's sklearn library, set the parameters n_estimators to 10000, random_state to 0 and n_jobs to -1 for model fitting. Finally, the importance ranking of features is obtained through feature_importances, and the detailed results are shown in Table 4.

Table 4: Index importance ranking

Index	Importance	Index	Importance	Index	Importance
Cultivated land versus built-up land Efficiency ratio	0.0605	Share of social security and employment expenditures	0.0578	The added value of non-agricultural industries accounts for the total regional output value	0.0527
Difference between urban and rural residents' expenditures	0.0595	Income gap between urban and rural residents	0.0572	Gap between urban and rural per capita expenditure on education, culture and entertainment	0.0525
Poor level of compulsory education distance	0.0593	Gap between urban and rural medical beds	0.0572	Road network density	0.0522
Poor conditions of compulsory education distance	0.0588	Proportion of education expenditure	0.0565	urbanization rate	0.0507
Cropland and built-up land Rate of change ratio	0.0586	Education investment per capita	0.0537	GDP per capita	0.0507
General public service expenditures Expenditure as a percentage of	0.0585	per capita consumption level	0.0528	Agricultural expenditure share	0.0506

The computational outputs demonstrate that the significance of each variable in evaluating the degree of urban-rural integration is relatively congruent, underscoring the relevance of each selected metric in the assessment process. Within the hierarchy of attributes, metrics such as the efficiency quotient of arable to developed land, the fiscal discrepancy between urban and rural inhabitants, and the delta in compulsory education provision emerge as paramount. This accentuates the salience of the urban-rural divide in the context of contemporary urbanization and integrative endeavors. In contrast, variables like road network density, urbanization coefficients, and per capita GDP manifest reduced prominence within the evaluative framework. This suggests that their immediate influence in bridging the urban-rural chasm is comparatively subdued.

3.3 Based on K-Means cluster analysis of development differences in urban-rural integration areas

To elucidate disparities in urban-rural integration, this research leveraged the gravel diagram implemented in the R programming environment to cluster provinces based on similar developmental attributes. By adopting the Euclidean distance metric within the K-Means clustering algorithm, the resultant heat map vividly illuminated pronounced developmental variances, notably between regions such as Qinghai and Tibet. Notwithstanding the gravel diagram's absence of a pronounced elbow or inflection point, three salient clusters were discernible: avant-garde cities, intermediate urban centers, and laggard cities. Provinces like Guangdong and Fujian were demarcated as vanguards, while regions such as Guizhou and Tibet were identified as trailing entities. Conversely, provinces like Shandong and Chongqing occupied an intermediary position. This stratification is congruent with assessments postulating that prominent provinces like Beijing, Shanghai, and Guangzhou are at the zenith of urban-rural integration trajectories. In juxtaposition, third-tier provinces grapple with formidable developmental conundrums, whereas second-tier provinces manifest a combination of robust attributes and potential domains warranting enhancement.

4. Comprehensive analysis of the status quo of regional urban-rural integration development

4.1 Urban-rural evolution game under the background of urban-rural integration

Table 5: Symbols used in evolutionary games and their corresponding meanings

symbol	Meaning
C_p	The cost of rural areas choosing to integrate into cities
M	Benefits brought by the development of the integration of rural areas into urban areas
T	The large cities acquired by the rural areas rely on infrastructure construction, policy subsidies,
	Support income from inputs such as tax incentives
P	Penalties received for non- integration of rural areas into cities
C_s	The cost of urban investment driving rural development
S	Amount of Subsidies from Cities to Actively Integrating Rural Areas
L	The City's Costs of Non- Integrated Rural Development
E	The benefits that cities gain from rural integration into megatrends
x	Probability of cities driving rural development
$1-x$	Probability that cities do not drive rural development
y	Probability of integrating rural areas into urban-rural integration construction
$1-y$	rural areas not being integrated into urban-rural integration construction

Within the urban-rural integration schema articulated by the 17th National Congress of the Communist Party of China, the linchpin of rural revitalization hinges on integration. However, owing to divergent stakeholder interests, large- and medium-sized urban centers, juxtaposed with rural locales, often recalibrate their strategic initiatives, engendering potential mutual suboptimal outcomes. To engender harmonious progression, this research delves into an evolutionary game theoretic analysis between these entities, with an objective of discerning a dynamic Nash equilibrium. The decision matrices of both urban and rural entities are punctuated with inherent costs and concomitant benefits, encapsulating strategic imperatives such as support versus integration. This manuscript delineates pivotal game theoretic parameters, including integration overheads and prospective penalty structures, with their corresponding symbolic representations elucidated in Table 5.

4.2 Model establishment

In elucidating the strategic imperatives of large- and medium-sized urban centers vis-à-vis rural locales concerning urban-rural integration, we formulated a payoff matrix. This matrix delineates the anticipated utility outcomes for urban and rural domains predicated on their strategic decisions: to lead or abstain, and to integrate or refrain. The mean expected payoff for large and medium-sized urban centers is contingent upon their decision to champion rural development. Simultaneously, the analogous payoff for rural sectors is contingent upon their decision to assimilate within the urban-rural integration paradigm. To explicate the dynamics underpinning these strategic modalities, we leverage the Malthusian replication dynamics axiom to extrapolate the replication dynamic equations pertinent to large- and medium-sized urban centers and rural areas. Through the resolution of these equations, we discern five equilibria, each emblematic of distinct strategic conjunction outcomes.

4.3 Dynamic analysis of unilateral decision-making

By analyzing the dynamic equations of strategy replication in large and medium-sized cities and rural areas, we determine their evolutionary stability. The stability of strategy evolution in large and medium-sized cities depends on parameter values $y(P-C+S+P)$, and the result may be a stable strategy x=0, x=1, or x ranges between 0 and 1. At the same time, the stability of strategy evolution in rural areas is related to parameter values $x(S+P-C-M+T)$, and its stable strategy may be y = 0, y = 1 or y ranges between 0 and 1. Generally speaking, there are three possible evolutionary stability strategies in both large and medium-sized cities and rural areas.

4.4 Stability analysis of equilibrium point

Utilizing Friedman's methodology, we conducted an assessment of the Jacobian matrix J to discern system stability, predicated on its determinant and trace. The J corresponding to equilibrium points O, A, B, and C exhibits negative values, categorizing them as saddle-node bifurcations. Contrarily, the equilibrium point D, manifesting characteristic roots with purely imaginary components, signifies a center equilibrium, which is non-asymptotically stable. Consequently, within the framework of urban-rural integration, the strategic interactions between large enterprises and rural mechanisms unfold dynamically, eschewing a monolithic strategy for either entity.

4.5 Numerical Simulation

In this part, the above calculation results are mainly simulated. In the simulation, the Python language is mainly used for simulation, and the initial value assignment is carried out for the relevant parameters. The main assignment parameters and assignment results are shown in Table 6:

Table 6: Assignment table of related parameters

parameter	initial value	parameter	initial value
Cp	0.6	S	0.2
m	0.3	P	0.3
T	0.2	C s	0.1

Considering that each element is greater than 0 and satisfies a specific relationship, the initial value of the parameter refers to the "China Statistical Yearbook" and related materials. Using Python simulation, we get five equilibrium points between large and medium-sized cities and rural areas. Taking (0.6, 0.4) as the initial value, the situation where the probability of implementing the driving strategy is 0.6 and the probability of integrating the strategy is 0.4 is analyzed, see Figure1 for details.

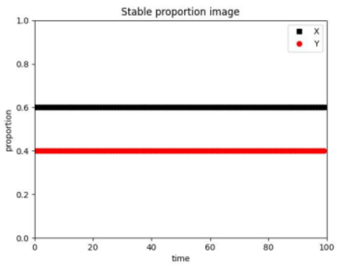

Figure 1: The evolutionary game diagram when the implementation drives the strategy with probability 0.6 and the implementation integration strategy probability is 0.4

Based on the graphical analysis, with an initial implementation probability of the driving strategy in large- and medium-sized cities set at 0.6 and the onset probability of the integration strategy in rural domains fixed at 0.4, the evolution of both entities appears to stabilize. This equilibrated state is denoted as the "stable evolutionary probability." A marginal adjustment to the initial parameters, wherein the driving strategy is recalibrated to 0.64 and the integration strategy to 0.37, was subsequently made. The resultant evolution, as illustrated, can be further scrutinized in Figure 2.

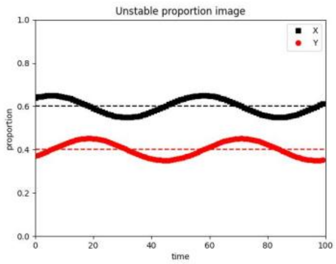

Figure 2: The evolutionary game diagram when the probability of implementing the strategy is 0.64 and the probability of implementing the integrated strategy is 0.37

Upon meticulous examination of the graphical representation, it's discernible that subtle perturbations in the initial conditions engender oscillations in the strategy implementation probabilities for both the urban and rural sectors. This observation proffers potential implications for policy recalibration. Although the probabilities gravitate towards an equilibrial state, they exhibit dynamic tendencies rather than static equilibria, alluding to an incessant, cyclical progression of urban-rural integration dynamics. Despite these fluctuations, the intensity of the strategic game remains relatively tempered. The sustained prevalence of the rural integration strategy, juxtaposed with the urban strategy attenuating to 0.2, is expounded in Figure 3.

The graphical representation elucidates that an augmented differential between strategy implementation probabilities and the stable evolutionary probabilities amplifies the game's inherent volatility, consequentially influencing strategic outcomes. The strategic trajectory is predominantly orchestrated by large and medium-sized urban entities. When these urban nuclei unequivocally embrace driving strategies, the rural sectors exhibit a propensity to align, indicative of an inclination towards comprehensive urban-rural integration. Conversely, when rural integration achieves its zenith, urban conglomerates might recalibrate their strategic focus, which could imply a governmental pivot towards the augmentation of urban clusters.

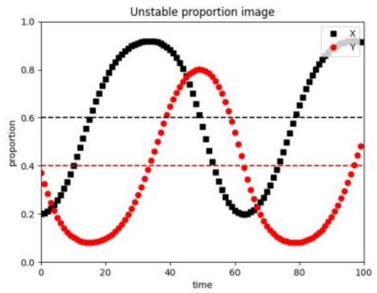

Figure 3: The evolutionary game diagram when the probability of implementing the driving strategy is 0.2 and the probability of implementing the integrated strategy is 0.37

5. Conclusion

The spectrum of urban-rural integration in China manifests pronounced regional heterogeneities. Coastal metropolises such as Guangzhou and Beijing exemplify developmental vanguards, whereas northwestern territories like Tibet and Xinjiang exhibit developmental lags, attributable to a plethora of unique challenges. The paradigmatic shift from an exclusive emphasis on rural revitalization towards a nuanced assessment of urban-rural disparities has accentuated the instrumental role of urban centers in catalyzing growth juxtaposed with the imperative of rural sectors for proactive amalgamation. To ameliorate these disparities, it is incumbent upon the central administration to champion technological proliferation in marginalized regions, recalibrate macroeconomic policies, and augment fiscal interventions. At a regional level, entities should strategically leverage their endogenous assets, encompassing niches like cultural tourism and avant-garde agricultural practices. Urban epicenters bear the onus of extending infrastructural tendrils into their rural peripheries, while rural territories are tasked with enhancing their industrial landscape, amplifying their cultural narratives, and magnetizing diaspora talents.

References

[1] Wang Qin, Li Jing, Ding Keke, etc. Digital infrastructure, factor allocation efficiency and urban-rural income gap [J]. Statistics and Decision Making, 2023, 39(09):29-34.

[2] Xie Guoxiong. On Optimizing Strategies of Guangxi Urban-Rural Integration Development System and Mechanism [J]. Coastal Enterprises and Technology, 2022, No. 211(06):61-67.

[3] Huang Hua. The logic and path construction of urban-rural integration development under the background of rural revitalization [J]. Journal of China University of Petroleum (Social Science Edition), 2022, 38(06):97-104.

[4] Gong Yanhong, Yin Yanan. My country's Urban-Rural Integration Benefits and Regional Development Differences —Based on Three-Dimensional Mean Clustering [J]. Future and Development, 2022, 46(10):71-78.