Statistical Analysis of Factors Influencing the Per Capita Disposable Income of Urban Residents in Anhui Province

Jiangshuang Huang(✉)

North University of China, No. 3, Xueyuan Road, Taiyuan 03005, Shangxi, China
hjs000528@163.com

Abstract. With rapid social and economic development, the living standard of residents is also rising. The disposable income per urban resident, which is one of the indicators of the living standard of the residents, is also attracting more and more attention. Therefore, it is of great significance to China's economic development to find out the main influencing factors affecting the per capita disposable income of urban residents.

Firstly, the data of urban residents in Anhui Province from 2011to 2019 were intercepted, and five influencing factors, namely per capita gross regional product, urban unemployment rate, number of students in general colleges and universities, wages of employed persons in urban units, and the proportion of tertiary industry, were selected as independent variables through literature review. Secondly, multiple cointegration was found in the multiple regression equation for each region, and the cointegration was removed by ridge regression. The GM(1,1) model was chosen for the small sample data. By comparing the predicted values with the real values, it was found that the model fitted well and the per capita disposable income of urban residents in Anhui Province was increasing year by year and the increase rate was increasing.

Keywords: Disposable income per capita · GM(1,1) · return of the Ridge

1 Introduction

The per capita disposable income of urban residents is that part of the total income of a household that can be used to arrange the daily life of the household, and it is the total income of the household after deducting the income tax paid, the social security contributions paid by individuals and the bookkeeping allowance of the surveyed households [1]. Since the reform and opening up, the economy has developed rapidly and the standard of living of residents has been rising. The per capita disposable income of urban residents, which is one of the indicators of the living standard of residents, has also attracted increasing attention. The per capita disposable income of urban residents is an important factor affecting the living standard of urban residents, and only by increasing the per capita disposable income can we expand our own material and spiritual consumption expenditure. Therefore, it is of great significance to China's economic development to find out the main influencing factors affecting the per capita disposable income of urban residents.

J. Yen et al. (Eds.): ICBIS 2023, AHCS 14, pp. 743–749, 2023.
https://doi.org/10.2991/978-94-6463-198-2_77

2 Analysis and Forecasting of Influencing Factors

2.1 Factor Analysis of Disposable Income of Urban Residents in Anhui Region

In this paper, the regional gross domestic product per capita is selected as the independent variable, the urban registered unemployment rate as the independent variable, the number of students enrolled in general higher education schools as the independent variable, the average wage of employed persons in urban units as variable, the share of the tertiary industry as the independent variable, and the disposable income per urban resident as the dependent variable y [2]. Regression analysis is used to quantify the influence of each factor on the dependent variable The degree of influence of each factor on the dependent variable was quantified using regression analysis. By consulting the statistical yearbooks of the National Bureau of Statistics. Data extracted from 2011–2019 were standardised as shown in Table 1.

The multiple linear regression equation on the dependent variable is now done using SPSS with a simultaneous covariance diagnosis. From Tables 2–3, it can be found that although the regression equation is highly significant, the coefficients of the independent variables are not significant, while the degrees of freedom adjusted complex coefficient of determination reaches 1.000 and this phenomenon can be explained by Table 4 the variance expansion factors of four of the independent variables are all greater than 10, where the variance expansion factors are well over 100 respectively, indicating that the regression equation for urban disposable income per capita has serious The regression equation of urban disposable income per capita has serious multicollinearity.

In order to resolve the multicollinearity, after trying various processing methods for the Anhui Province data, the results of Ridge regression were finally found to be more

Table 1. Anhui Province Data

2011	-1.39877	-1.32813	1.45834	-1.49396	-1.33196	-1.34556
2012	-1.01572	-1.00720	1.29893	-1.07958	-.94184	-.99376
2013	-.73613	-.66684	.58157	-.69920	-.70363	-.91258
2014	-.41138	-.38792	.05019	-.32796	-.47412	-.57431
2015	-.07919	-.21655	-.13580	.32825	-.15862	.14282
2016	.27249	.13191	.02362	.51517	.13592	1.49590
2017	.66598	.61107	-.82659	.54655	.58542	1.02233
2018	1.10209	1.22139	-.95943	.43805	1.27128	.75171
2019	1.60062	1.64225	-1.49081	1.77268	1.61755	.41344

Table 2. Model Summary

Model		Adjusted R-squared	Error in standard estimation
	1.000	1.000	.02019158

Table 3. Analysis of variance

Model	square and	Freedom	Mean Square	F	Significance
Return	7.999	5	1.600	3923.854	.000b
Residuals	.001	3	.000		
Total	8.000	8			

Table 4. Regression Coefficient

Model	Unstandardised factor		Standardisation factor			Covariance statistics	
	B	Standard error	Beta	t	Significance	tolerance	VIF
Constant	-4.297E	.007		.000	1.000		
x1	.339	.181	.339	1.871	.158	.002	645.583
x2	-.059	.037	-.059	-1.629	.202	.038	26.177
x3	.145	.024	.145	6.089	.009	.089	11.189
x4	.411	.166	.411	2.478	.089	.002	539.983
x5	.070	.012	.070	5.820	.010	.348	2.876

in line with the actual situation, and the process of Ridge regression solution is given here as follows.

The ridge traces obtained after the Ridge regression (Ridge regression) analysis using SPSSAU are shown in Fig. 1. It can be seen that when the k value is 0.03, the standardised regression coefficients of the independent variables tend to be stable at this time, and thus SPSSAU recommends setting the optimal k value to be taken as 0.03.

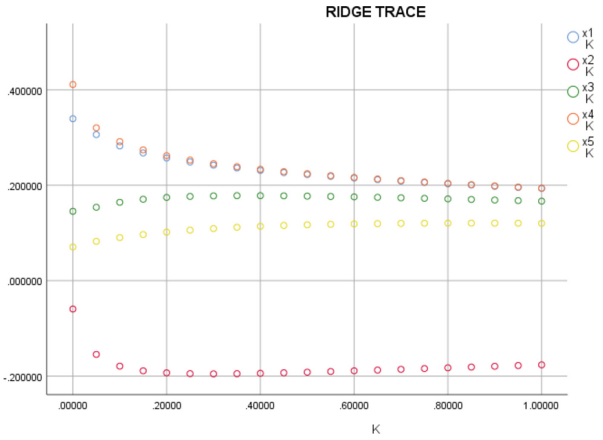

Fig. 1. Map of the Ridge

Table 5. Ridge regression equation coefficients

	Unstandardised factor		Standardisation factor	t	p		Adjusted R-squared	F
	B	Standard error	Beta					
Constant	900.449	4442.991	–	0.203	0.852	0.999	0.998	F(5,3) = 819.303 p = 0.000
x1	0.193	0.013	0.321	14.927	0.001**			
x2	−2233.701	648.033	−0.133	−3.447	0.041*			
x3	122.485	29.76	0.148	4.116	0.026*			
x4	0.159	0.012	0.34	13.624	0.001**			
x5	67.037	19.743	0.078	3.396	0.043*			

The value of k is taken as 0.030, and from Table 5, it can be seen that the model adjustment = 0.998, implying that the per capita gross regional product (RMB), the registered urban unemployment rate (%), the number of students in general higher education schools (10,000), the average wage of employed persons in urban units (RMB), and the proportion of tertiary industry (%) can explain 99.93% of the variation in the per capita disposable income of urban residents (RMB), indicating that The equation fits well.

The regression coefficients of the five independent variables and the regression equation all passed the significance test, which means that at least one of them would have an impact relationship on urban residents' disposable income per capita (RMB). The standardised ridge regression equation is:

$$y = 0.321x_1 - 0.133x_2 + 0.148x_3 + 0.34x_4 + 0.078x_5 \quad (1)$$

The coefficient of 0.34 indicates that the average wage of employed persons in urban units has the greatest influence on the per capita disposable income of urban residents in Anhui Province, followed by the regional per capita GDP, the number of students in higher education institutions, the urban registered unemployment rate and the share of the tertiary sector respectively.

3 Forecast of Per Capita Disposable Income of Urban Residents in Anhui Province

Due to the small sample data, grey forecasting and exponential smoothing are considered in this paper. After a later comparison of the prediction effect of the exponential smoothing method and the grey prediction model (the lower the root mean square error RMSE, the better the model fit) [3]. As can be seen from Table 6, the regional data of Anhui Province are more suitable for grey forecasting. Here we have selected urban disposable income per capita data for the period 2011–2020 for forecasting.

Let variable be the original data column of urban residents' per capita disposable income.

Table 6. RMSE Comparison

Model	RMSE
GM(1,1)	892.163
Exponential Smoothing	1164.198

Firstly, a rank-ratio test was conducted to determine the suitability of the data series for model construction, and a rank-ratio between and indicated that the data were suitable for model construction. As can be seen from Table 7, the values of the rank-ratio test are all within the standard range [0.834, 1.199], implying that the data are suitable for the GM(1,1) model.

Grey forecasting using the R language yielded a development coefficient of a = − 0.0797, a grey action of u = 18857.6185, and a calculated posterior difference ratio C value = 0.0018, where the posterior difference ratio C value of 0.00180.35 implies a very good model accuracy rating., (expresses its interest in after a single accumulation of (1 - AGO).

$$\hat{y}^{(1)}(k) = \left(y^{(0)}(1) - \frac{u}{a}\right)e^{-a(k-1)} + \frac{u}{a} = 255213.50941028e^{0.0797(k-1)} - 236607.50941028 \tag{2}$$

The predicted values and prediction accuracy are shown in Table 8. The maximum value of the relative error of the model, $0.015 < 0.1$, implies that the model fit is at a high level; for the grade deviation value, a value less than 0.2 indicates that the requirement is met, and if it is less than 0.1, it indicates that the higher requirement is met. The maximum value of the grade deviation of the model, $0.042 < 0.1$, also implies that the model fit is at a high level.

Table 7. GM(1,1) model level ratios

Serial number	Original value	Level Ratio	shift value(shift = 0)	Converted step ratio values
1	18606	–	18606	–
2	21024	0.885	21024	0.885
3	22789	0.923	22789	0.923
4	24839	0.917	24839	0.917
5	26936	0.922	26936	0.922
6	29156	0.924	29156	0.924
7	31640	0.921	31640	0.921
8	34393	0.92	34393	0.92
9	37540	0.916	37540	0.916
10	39442	0.952	39442	0.952

Table 8. GM (1,1) model level ratios

Serial number	Original value	Predicted value	Residuals	Relative error	Grade ratio deviation
1	18606	18606	0	0.000%	–
2	21024	21172.05	−148.054	0.704%	0.042
3	22789	22927.77	−138.77	0.609%	0.001
4	24839	24829.08	9.92	0.040%	0.006
5	26936	26888.06	47.942	0.178%	0.001
6	29156	29117.78	38.221	0.131%	−0.001
7	31640	31532.4	107.599	0.340%	0.002
8	34393	34147.26	245.74	0.715%	0.004
9	37540	36978.96	561.042	1.495%	0.008
10	39442	40045.48	−603.477	1.530%	−0.031

4 Conclusion

This article analyses and predicts the factors affecting disposable income per capita based on the yearbook data of the Bureau of Statistics. Among the factors influencing disposable income per capita, the average wage of employed persons in urban units has the greatest influence [4]. This indicates that wages account for the majority of disposable income of urban residents in Anhui Province, and when the level of wages is relatively high, it can be expected that the level of disposable income in that area is also relatively high [5]. Secondly, the unemployment rate is negatively correlated with disposable income. An increase in the unemployment rate represents a slowdown in economic development and is not conducive to currency appreciation.

References

1. Han Haiyan. Research on income structure, instability and consumption of urban residents in China [D]. Northwestern University, 2007.
2. Dai L, Xu HK, Chen T et al. A MapReduce-based multiple linear regression prediction model[J]. Computer Applications, 2014, 07: 1862-1866.
3. Ye Feng. Application of multiple linear regression in economic and technical yield forecasting[J]. China and Foreign Energy, 2015, 02:45-48
4. Hoerl A, Kennard R. Ridge Regression: Biased Estimation for Nonorthogonal Problems[]. Technometrics, 2000, 12(1):13.
5. Yang Shuang. Research on Influencing Factors of urban residents' Consumption [J]. Finance Theory and Teaching, 2012, (01): 67-70.
6. Bai Xin-jun, Cao Xin, He Shuang. Empirical analysis of influencing factors of urban residents' consumption in China. Journal of Shandong Agricultural Engineering University, 2018, 35(04): 3–4+6.

7. Liu Xiaopeng. Integration Analysis and Error Correction Model -- An Empirical Study of Chinese Foreign Trade and Economic Growth [J]. Nankai Economic Research, 2011, 5(4): 55-57.
8. Cao Jian, Wang Minghan, Zhang Jingyi. Analysis on the consumption level of Chinese urban residents and its influencing factors [J]. Modern Business, 2016, (22): 75-76.
9. High iron plum. Econometric Analysis Method and Modeling [M]. 2nd Ed. Beijing: Tsinghua University Press, 2006
10. LI S J. Demand situation and countermeasures research on the sports service products of rural residents[J]. Cross-Cultural Communication, 2013, 9(1): 26-30.
11. Qi Xin, Wang Xinhua. The trend and influencing factors of income gap in China [J]. Economic Research, 2016, 10(3): 43-44