Prediction and Analysis of Chinese Rural Households' Consumption Level Based on the ARIMA Model

YAN Jian-biao, LI Qiang *

College of Economics & Management, Beijing Forestry University, Beijing 100083, China

Abstract By using the software SAS9.2 and the relevant data of consumption level of rural residents in China from 1952 to 2008, the ARIMA model is established. The model is used to analyze and forecast the time series of the consumption level of Chinese rural residents. The results show that in the near future, the consumption level of Chinese rural residents will be further raised. In 2012, the level will break through per capita 5 000 yuan, almost 100 times more than that in the primary time period. But consumption level does not equal to living standard. To let farmers lead a good life, the government should follow the objective laws; take the overall situation into consideration; coordinate the relations among farmers' consumption level, national subsidies and farmers' production enthusiasm. Therefore, The paper suggests that the historical and objective factors should be attached more importance to. Besides, raising farmers' income and allaying farmers' fear were effective measures in developing the consumptive potential of rural market and promoting the economic sustainable development.

Key words ARIMA Model, Rural households consumption, Economic growth, China

China is a developing country, with an overwhelming majority of rural population. The insufficient consumption of rural residents is accountable for the inadequate consumption of aggregate Chinese residents. Laying stress on and continuously promoting the demand of rural consumption are effective way of enlarging domestic demand. Chinese scholars have studied the consumption level of Chinese rural and urban residents and put forward the relevant opinions. CAI Yue-zhou *et al*., conducted the research from the aspects of economic stimulus package, deepening rural finance, social security expenses and income sources[1 – 4]. CHEN Liang *et al*., analyzed the promotion role played by the consumption of rural residents and the effects of urbanization[5 – 6]; HAN Sheng-juan conducted the relevant research from the perspective of farmers' consumption behaviors, the change of consumption structure *et al*.[7 – 8]. Through trying to conduct the time series analysis and forecast on the consumption level data of Chinese rural and urban residents, ZHANG Wen-bing pointed out that the increase of Chinese rural and urban residents' consumption level is the engine to directly drive the economic development; besides that, the forecasting results could be applied to the upgrade of consumption structure and industrial structure[9]. However, the above mentioned researches are all based on the urban and rural perspectives, few of them particularly concerns the forecast and analysis of rural residents' consumption level. So on the basis of the relevant researches, the paper forecasts and analyzes the consumption level of Chinese rural residents according to the data of Chinese rural residents' consumption level and by using the SAS9.2 ARIMA model.

Received: February 11, 2011 Accepted: March 27, 2011

Supported by National Natural Science Foundation of China (70803005) and Special fund of the Baisc Sicentific Research Fund for Central Colleges and Universisties (RW2010-6).

* Corresponding author. E-mail: qiangli@bjfu.edu.cn

1 Data source and ARIMA model

1.1 Data source The relevant data of rural residents' consumption level from 1952 to 2008 come from *Statistical Data of Sixty Years of New China.*

1.2 Basic principles of ARIMA model The ARIMA(p,d,q) model is short for autoregressive integrated moving average model. Among the model, AR(p) is the auto-regression model; MA(q) is moving-average model; p and q are the corresponding orders respectively; *I* means the combination of the two models; d refers to the frequency of differential process on non-stationary time series, which has the features of long term trend, seasonal change and cycling change. The equation of ARIMA model is as follows:

$$\begin{cases}\Phi(B)\nabla^d x_t = \Theta(B)\varepsilon_t \\ E(\varepsilon_t)=0, Var(\varepsilon_t)=\sigma_\varepsilon^2,\ E(\varepsilon_t\varepsilon_s)=0, s\neq t \\ Ex_s\varepsilon_t=0,\ \forall s<t\end{cases}$$

In the equation, $\nabla^d=(1-B)^d$ and $\Phi(B)=1-\varphi_1 B-\cdots-\varphi_p B^p$ are the auto-regressive coefficient polynomial expressiones of smooth reversible ARMA(p,q) model; $\theta(B)=1-\theta_1 B-\cdots-\theta_q B^q$ is the polynomial expression of moving smooth coefficient; $\{\varepsilon_t\}$ is the white nose sequence with zero-mean[10]. The essence of ARIMA model is the combination of differential operation and ARMA model. Any non-stationary time series can be smoothened after differential process through proper order differential, and the ARMA model fitting can be conducted on the sequences.

1.3 Steps to establish ARIMA model The establishment of ARIMA can be divided into six steps[11]. The specific steps can be seen on following Fig.1.

2 Results and analysis

2.1 Test and process of the smoothness of sequence In the first place, establishing the time series $\{r_t\}$ by using the

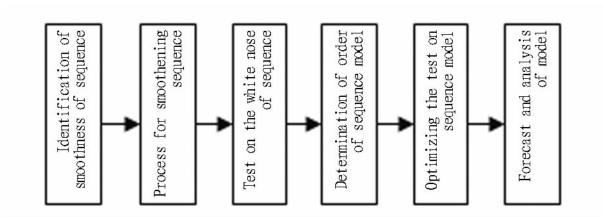

Fig. 1 Steps to make ARIMA model for sequence

data of rural households' consumption level (variable: rhc) from *Statistical Data of Sixty Years of New China* from 1952 *to* 2008. In the second place, the time series diagram of these data can be established.

It can be seen from Fig. 2 that, with the pass of time, the sequence shows obvious non-linear increase trend, so it can be regarded as non-stationary sequence. And then logarithmic transformation and differential process are conducted on the sequence. Supposing the sequence after logarithmic transformation is $\{lr_t\}$ and the heteroscedasticity is reduced; after conducting the first-order difference and the second-order difference, the $\{d1lr_t\}$ and $\{d2lr_t\}$ can be obtained, and then the long term trend in $\{lr_t\}$ can be extracted.

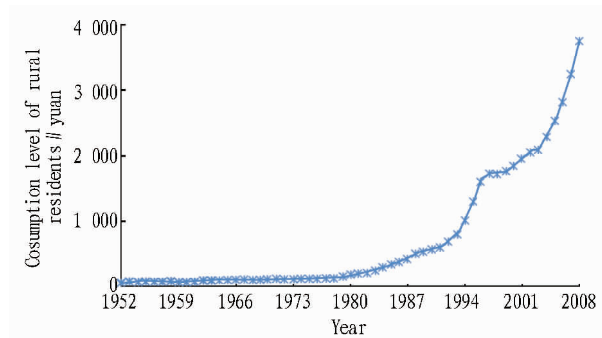

Fig. 2 Time series diagram of $\{r_t\}$ sequence

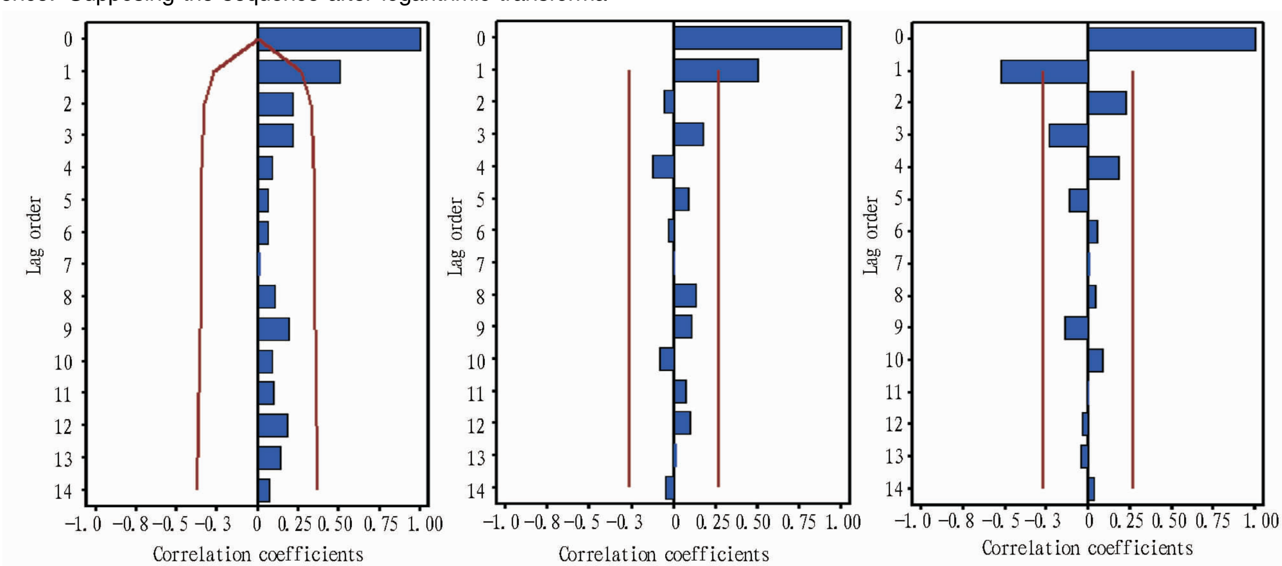

Fig. 3 Autocorrelation, partial correlation and inverse correlation diagrams of $\{d1lr_t\}$ sequence

It can be seen from Fig. 3 that the autocorrelation coefficient and partial correlation coefficient of $\{d1lr_t\}$ are obviously the first-order truncate and show strong short-term relativity and stability; the first-order relevant coefficient of $\{d2lrt\}$ is rather small and very smooth, but the inverse correlation coefficient has many order twice more than the standard difference, that is the over differential exists, which leads to the waste of valid information and reduces the accuracy of evaluation. So $\{d1lr_t\}$ is the better choice.

In the white noise test on $\{d1lr_t\}$, the p value of test statistics of QLB of each order is very small (<0.01), so it can be proved that (confidence level $>99\%$) that $\{d1lr_t\}$ is non-white noise series, that is to say, it still has relevant information that is worthy to be extracted. Thus, it can be known that $\{d1lr_t\}$ is stable non-white noise series.

2.2 Identification and optimization of ARIMA model The identification of model mainly depends on the autocorrelation function and partial autocorrelation function. The delay auto

Table 1 The result of white noise test of $\{d1lr_t\}$

Lag order	ChiSq. Value	Degree of free	Pr > ChiSq. Value	Autocorrelation coefficient		
6	21.42	6	0.001 5	0.502	0.212	0.218
12	28.34	12	0.004 9	0.011	0.104	0.193

correlation coefficient of first order of $\{d1lr_t\}$ is significant and does not equal to zero, after that all the coefficient is included in the two times standard scope, and stays on the side of zero axial and trailing wave; the partial auto correlation coefficient of delay first order is significantly larger than two times of the standard deviation, the delay third order is close to the two times of standard deviation, and the other orders fluctuate around zero axial and become weak suddenly, so the partial auto correlation coefficient has the obvious truncation. Therefore, the value of p in AR(p) model can be 1 or 3. Combining the analysis of first order difference and auto correlation coefficient, the paper analyzes the following ARIMA(1,1,0) and ARIMA(3,1,0) models.

The test is divided into two parts: model test and parameter test. Model test is the white noise test of residual error after the fitting of model. The P value of ARIMA(1,1,0) and ARIMA (3,1,0) under QLB statistics of each delay order is larger than 0.05, so it can be fixed that the residual sequences of the two fitting models belong to white noise series, that is to say, the model is significant. In parameter test, it can be seen from the P value of t statistics that the typical value of ARIMA(1,1,0) and coefficient 1 are all significant, but the coefficient 2 and 3 of ARIMA(3,1,0) are not so significant, so the model should be eliminated. After being optimized, the model is ARIMA(1,1,0).

Table 2 Comparison of ARIMA(1,1,0) & ARIMA(3,1,0)

	ARIMA(1,1,0)	ARIMA(3,1,0)
Constant Estimate	0.036 615	0.064 733
Variance Estimate	0.004 313	0.004 337
Standard Error Estimate	0.065 675	0.065 858
AIC	−144.096	−142.831
SBC	−140.045	−136.755

In the end, comparing the value of AIC and SBC, AIC and SBC fit by ARIMA(1,1,0) model is smaller than ARIMA(3,1,0) model. According to the rule of AIC and SBC, ARIMA(1,1,0) model is more effective. In summary, the $\{d1lr_t\}$ sequence is fit to ARIMA(1,1,0) model.

In order to simplify expression, the sequence variable logrhc is represented by x_t, and according to the prediction value of parameter, the diameter of the model is as follows:

$$\nabla x_t = 0.074\,7 + \frac{1}{1-0.509\,8B}\quad \varepsilon_t$$

2.3 The prediction and test of ARIMA model The prediction is based on the historic data of rural residents' consumption level and the model above. Through model calculation and exponent calculation, the predictive value of Chinese rural residents' consumption level from 2009 to 2010 can be predicted (Table 3). Comparing with the actual value of the consumption level of Chinese rural residents from 2009 to 2010, the difference of the prediction result of the model is small.

Table 3 The predictive value of the rural households' consumption level in China from 2009 to 2013

Year	Predictive value	Upper bound	Lower bound
2009	4 184	3 679	4 760
2010	4 586	3 633	5 791
2011	4 985	3 599	6 906
2012	5 395	3 589	8 111
2013	5 827	3 605	9 418

3 Conclusions and discussions

Rural households' consumption level is affected by multiple factors , but the ARIMA(1,1,0) model only takes historical factors into consideration to find its rules. It can be seen from the prediction outcome that in the future decades of years; Chinese rural households' consumption level will be further improved rapidly. In 2012, the consumption volume will break through 5 000 yuan per capita, nearly 100 times more than that in the primary days of Chinese foundation. The growth reflects on the development of Chinese economy and social progress, which is indispensable to Chinese government's attention to "Three Agricultural Problems". The complete exemption of agricultural taxes in particular has greatly simulated the consumption potential of farmers. But the consumption level does not equal to living standard. In order to ensure the comfortable life of farmers, the objective laws should be obeyed, the overall situation should be considered and the relations among farmers' consumption level, national subsidy and farmers' enthusiasm should be coordinated. Researches show that farmers' income is the crucial element that affects farmers' consumption. Farmers' income concerns the social problems of the standardization of farmers' training institutions, rural land occupation and wage arrears for migrant rural workers, so increasing farmers' income needs the efforts from various aspects. The government should clearly know the historic rules; strengthen the macro-control and support on agriculture; lighten the medical and education burden of farmers; adjust agricultural structure; develop high-quality and high efficiency agriculture according to local conditions; develop rural township enterprises; and accelerate the transfer of rural surplus labors. The enterprises should fully respect rural labors; take people as the priority and promote mutual profits. Farmers should continue to improve their quality and service ability. Only by jointly increasing farmers' income and solving farmers' rear fear and worries, can farmers' living standard be improved. The consumption potential of Chinese rural households is really the mainstay of future economic growth, but due to the differences of location, culture and history, the consumption potential should be treated discriminately. The forecast and analysis of the time series of rural residents' consumption level should be launched according to different types and should combine the time and spatial comprehensive contrastive analysis, so as to provide useful information for making policies concerning three agricultural problems; improving farmers' consumption level and facilitating economic growth.

References

[1] CAI YZ. Economic stimulus package and start-up of rural consumption in China: empirical study based on the decomposition of rural household income[J]. Journal of Finance And Economics, 2009, 35 (7): 4−12. (in Chinese).

[2] LIU CB, SANG TZ. Rural financial deepening and the growth of rural residents consumption: hypothesis and empirical analysis[J]. Journal of Jiangxi University of Finance and Economics, 2010(2): 62−66. (in Chinese).

[3] XIE W, WU QT. The influence of the social security on the consumption expenditure of rural residents[J]. The Theory and Practice of Finance and Economics, 2009, 30(5): 27−32. (in Chinese).

[4] LI JQ, XU HQ. Income source and consumption of rural residents the conclusion and enlightenment based on panel data[J]. Economic Survey, 2009(6): 107−110. (in Chinese).

[5] CHEN L, ZHU C. An empirical analysis on the effect of the consumption of China rural residents on economy growth and its countermeasures[J]. Economic Review, 2010(2): 50−54. (in Chinese).

[6] ZHANG XY, ZHENG Q. Analysis on the consumption of China rural residents and cointegration level of urbanization[J]. Commercial Times, 2010(4): 12, 89. (in Chinese).

(To page 88)

$$R'_1=\begin{Bmatrix}0.5333 & 0.2000 & 0.0667 & 0.2000 & 0.0000\\ 0.4667 & 0.2000 & 0.2000 & 0.1333 & 0.0000\\ 0.5333 & 0.2000 & 0.2667 & 0.0000 & 0.0000\end{Bmatrix}$$

$$R'_2=\begin{Bmatrix}0.6000 & 0.1333 & 0.0667 & 0.2000 & 0.0000\\ 0.4667 & 0.2667 & 0.2666 & 0.0000 & 0.0000\\ 0.4667 & 0.3333 & 0.1333 & 0.0667 & 0.0000\end{Bmatrix}$$

$$R'_3=\begin{Bmatrix}0.4667 & 0.3333 & 0.2000 & 0.0000 & 0.0000\\ 0.5333 & 0.2667 & 0.2000 & 0.0000 & 0.0000\end{Bmatrix}$$

3.3 Evaluation of hook benefit

3.3.1 Calculation of the fuzzy comprehensive rating model. Weight matrix A and index membership degree matrix R determined by Table 1 are brought into equation (2). Hence, the evaluation models of ecological, social and economic benefits before and after the hook are obtained: Economic benefit evaluation models before and after hook are $B_1=[0.2243, 0.3349, 0.3023, 0.1385, 0.000]$ and $B'_1=[0.5091, 0.2000, 0.1524, 0.1385, 0.0000]$; social benefit evaluation models before and after hook are $B_2=[0.2804, 0.21449, 0.4125, 0.0927, 0.0000]$ and $B'_2=[0.5196, 0.2270, 0.1607, 0.0927, 0.0000]$; and ecological benefit evaluation models before and after hook are $B_3=[0.1593, 0.4000, 0.4148, 0.0259, 0.0000]$ and $B'_3=[0.4926, 0.3074, 0.2000, 0.0000, 0.0000]$.

Table 2 The weight of hook benefits evaluation index

Target layer	First-grade index	Weight	Second-grade index	Weight
Comprehensive benefit	Ecological benefit	0.430 4	Added value of per capita GDP in research region	0.449 7
			Increasing rate of per capita annual net income of farmers	0.364 5
			Increasing rate of land unit yield	0.185 8
	Social benefit	0.363 0	Increasing rate of agricultural labor productivity	0.397 0
			Increasing rate of per capita traffic area	0.403 6
			Increasing rate of per capita cultivated area	0.199 4
	Ecological benefit	0.206 6	Added value of forest coverage rate	0.611 3
			Added value of land reclamation rate	0.388 7

3.3.2 Calculation of the score of hook benefit evaluation. In order to compare the benefits in research region before and after the hook and to reflect the benefits actually, equidifferent method is introduced to set the evaluation grade standard according the results of fuzzy calculation. The evaluation criteria score function is established:

$$F=(f_1,f_2,f_3,f_4,f_5)^T=(100,80,60,40,20)^T.$$

Calculation result of equation (2) and the evaluation criteria score function F are brought into equation (3). Then, the economic benefits evaluation scores before and after the hook are 72.90 and 81.59; the social benefits evaluation scores before and after the hook are 73.65 and 83.47; and the ecological benefits evaluation scores before and after the hook are 73.85 and 85.85.

4 Conclusion

Benefits evaluation scores before the hook are all lower than 80 in the research region, while those after the hook are all greater than 80, indicating that the implementation of hook has brought relatively great economic, social and ecological benefits. Thus, it can be concluded that the implementation of hook is successful as a whole.

Experimental work of hook is a complex system. Its benefit evaluation involves economy, society, ecology and other benefit goal factors. In order to fully and correctly evaluate the comprehensive benefit, indices of various aspects should be considered by combining with the actual situation of research region. At present, Fuzzy Comprehensive Evaluation Method is an effective method of fuzzy mathematics which is widely applied. Considering all the influencing factors, the merits of a given subject are scientifically evaluated. Since there are few researches on the hook benefit evaluation, standards for reference are lacked. Therefore, benefit evaluation method in this research can provide references for the hook benefit evaluation in the suburbs of Chongqing City, China.

References

[1] QIU DT, XIAO G, LIAO HP, *et al*. Comprehensive evaluation for environmental benefit of land use in towns: example for Beibei Area, Chongqing[J]. Journal of Southwest China Normal University: Natural Science Edition, 2008, 26(6): 729 –732. (in Chinese).

[2] YANG J, YANG B. Beneficial evaluation system construction of land development and consolidation projects in Sichuan[J]. Journal of Anhui Agricultural Sciences, 2009, 37(25): 12213 –1221, 12231. (in Chinese).

[3] HU YM, XIAO L, JIANG H, *et al*. Application of analytic hierarchy process in study on landuse measures[J]. Journal of Northwest A&F University: Natural Science Edition, 2001, 29(3): 57 – 60. (in Chinese).

[4] ZHAN W. Performance evaluation for land consolation project using fuzzy comprehensive assessment method[J]. Journal of Nanjing Forestry University: Natural Science Edition, 2009, 33(2): 145 – 148. (in Chinese).

[5] REN DF. Analysis of the early warning methods of personnel loss crisis in the township enterprises based on fuzzy integrated evaluation[J]. Journal of Anhui Agricultural Sciences, 2009, 37(13): 6216 –6217. (in Chinese).

(From page 85)

[7] HAN SJ, YU YF. Analysis on the consumption behavior characteristics of rural residents in Jiangxi Province based on clustering method[J]. Journal of Anhui Agricultural Sciences, 2009(7): 3287 –3288. (in Chinese).

[8] WANG YX, XIAO HF. Study on the division of change stages of the consumption structure of rural residents in China since the nineties of the last century[J]. Market Modernization, 2008(15): 168 – 169. (in Chinese).

[9] ZHANG WB, FAN L. The application of time series analysis in prediction of Chinese rural and urban residents consumption level[J]. Science & Technology Information, 2010(8): 120 –121. (in Chinese).

[10] WANG Y. Application of time series[M]. Beijing: Renmin University of China Press, 2005. (in Chinese).

[11] CAO F. Application of ARIMA model in the prediction of rural residents consumption per capita in Yunnan Province[J]. Journal of Anhui Agricultural Sciences, 2009, 37(30): 14923 – 14925. (in Chinese).