Fitting and Prediction of Total Retail Sales of Consumer Goods Based on Consumption Indicators

Wenting Zhang
School of Statistics and Mathematics
Zhongnan University of Economics and Law
Wuhan, China 430073
Project Management Center
Central and Southern China Municipal Engineering Design
& Research Institute Co., Ltd.
Wuhan, China 430010

Yingmei Xu
School of Statistics and Mathematics
Zhongnan University of Economics and Law
Wuhan, China 430073

Abstract—**This paper uses eight indicators of consumption as explanatory variables to construct a fitting and prediction model for the monthly total retail sales of consumer goods. This essay used data from January 2005 to June 2018 for empirical study. The result is good with 1.99% fitting MAPE and 2.39% prediction MAPE. This method is a checksum supplement to the measurement error of the actual total retail sales of consumer goods. At the same time, the model not only contains physical consumption, but also includes the measurement of service consumption. Compared with the index of the total retail sales of consumer goods, it is more comprehensive and real time to reflect the fluctuation of general consumption level. From a macro perspective, this method is a better reference for supply and demand analysis, economic heat, policy effects analysis, and reflection of people's livelihood.**

Keywords—total retail sales of consumer goods; consumption; prediction

I. INTRODUCTION

The total retail sales of consumer goods index is one of the classical indicator in China. This indicator was created during the founding of the People's Republic of China and has been used for more than 60 years. After years of development and improvement, its measurement method has formed a set of its own statistical survey methods. The total retail sales of consumer goods is published monthly, mainly using the data directly reported by enterprises above designated size, and data sampling surveys of enterprises below the quota. The statistics range includes wholesale and retail sales, and accommodation and catering sales. From a statistical perspective, the main component of total retail sales of consumer goods is physical goods sales, and the secondary part is accommodation and catering sales. Based on the maturity of its statistical system and means and the rapid and timely publish frequency, the total retail sales index of consumer goods has become a barometer of China's measurement of macroeconomic trends, domestic trade circulation, purchasing power of consumer goods, and people's living conditions, and has played a pivotal role in real-time monitoring of China's economic situation and measurement.

Although the total retail sales of consumer goods have an important position in China's macroeconomic operation index system, its limitations have gradually emerged. According to the data published by the National Bureau of Statistics, the GDP of 2017 was 82,712.17 billion yuan, the added value of the primary industry was 6,546.76 billion yuan (accounting for 7.92%), and the added value of the secondary industry was 33,462.26 billion yuan (40.46%), the tertiary industry. The added value is 42,703.15 billion yuan (accounting for 51.63%). It can be seen that the service industry accounts for a relatively large proportion of China's GDP. The total retail sales of consumer goods are proposed in the era of planned economy. Its main measurement target is the physical goods produced in China, and does not include the measurement of service commodities. At that time, China's service industry was not developed, and the relevant statistics for the service industry were relatively weak. As a result, there is no mature high-frequency accounting index for service goods, which is also an important part.

Other consumer indicators published by the National Bureau of Statistics include final consumption expenditure, per capita consumption expenditure of residents, consumer price index (CPI) and retail price index (RPI). The final consumption expenditure is a relatively comprehensive consumption indicator. Its connotation includes not only physical commodity consumption, but also service commodity consumption and virtual consumption. This indicator is still an international metric, unlike the total retail sales of consumer goods, which is only used in China. The final consumption expenditure indicator also has its own shortcomings. The biggest problem is that its data is published very slowly. The indicator is published on an annual frequency and generally two years delay. Another problem with it is that data is not collected directly from companies and individuals, but is derived from other economic indicators. The per capita consumption expenditure of residents is divided into per capita

consumption expenditure of urban residents and per capita consumption expenditure of rural residents, including the total final consumption of goods and services by resident residents. These two indicators are quarterly frequency publish data, with more real-time update and comprehensive coverage. The downside is that it measures the per capita level rather than the overall measure. The consumer price index (CPI) and the retail price index (RPI) are also relatively mature indicators in China. They are in line with international indicators and are published on a monthly frequency. The sample survey method is used to achieve direct data collection, which is an effective tool for measuring the price fluctuations and inflation. The limitation is that it cannot reflect the level of consumption from total amount. As for the consumer index, whether total retail sales of consumer goods can be replaced by other relevant indicators or not, is the question that the author will discuss later.

At present, some scholars use time series model to carry out the fitting and prediction of the total retail sales of consumer goods, and they believe that only the time characteristics and autocorrelation can explain the total retail sales of consumer goods. The representative approach is to use the ARIMA model to fit the total retail sales of consumer goods from the time dimension [1] [2]. Other methods include ARMA criteria for support vector regression feature extraction [3], cubic exponential smoothing method [4], and GMDH autoregressive algorithm [5]. The author believes that this is only based on the time factor. At most, the method of fitting and predicting the autocorrelation of the total retail sales of the consumer goods in the current period and the lag period is a big problem. Can China's total retail sales of consumer goods grow unrestricted over time? Is there no bottleneck and "ceiling" effect? The author believes that the short-term forecast data of the total retail sales of consumer goods should be calculated according to the relevant consumer economic indicators, so as to assist the correction and reference control of the actual data measurement error.

II. RESEARCH DESIGN

The research design section is divided into three parts: research ideas and research hypotheses, variable selection, and model design. This paper expounds the basic ideas, premise assumptions, data sources, substitution variables, auxiliary variables and design ideas of model calculation for the replacement of fitting and prediction of total retail sales of consumer goods.

A. Research Ideas and Research Hypotheses

The author believes that the monthly total retail sales of consumer goods can be replaced by their relevant economic indicators and carry on the fitting and predictions. On the one hand, the current price of the price index can be multiplied by the total freight volume to make a rough estimate, and on the other hand, the per capita consumption expenditure of residents is multiplied by the total population to estimate and the volatility model is constructed for the corresponding error portion. Although the final consumption

expenditure index is the most complete in terms of statistical caliber and connotation, due to its low frequency hysteresis and the characteristics of indirect data collection, it is not taken into consideration here. The CPI, RPI, and freight volume data used in this study are statistical data published on a monthly frequency and can be used directly; per capita consumption expenditure of urban residents and per capita consumption expenditure of rural residents are quarterly data, taking into account short-term stickiness of prices and consumer demand, the monthly data of these two variables will follow the data values published in their quarters; the population data is published annually. Due to the fact that there have been no major changes in natural disasters and man-made disasters in recent years, which can make the population suddenly, the author uses the population data at the beginning and end of the year to make an equidistance difference for each month and then divide it into months.

Based on the above, the author puts forward four basic hypotheses:

Hypothesis 1: The product of the mixed price index and the total freight volume converted to the current period can reflect the trend of the total retail sales of consumer goods.

Hypothesis 2: Multiplying the per capita consumption expenditure of residents with the current estimated population can reflect the changing trend of total retail sales of consumer goods.

Hypothesis 3: In the past decade or so, the total population of China has changed steadily.

Hypothesis 4: The consumption level of urban and rural residents has a short-term viscosity within three months.

B. Variable Selection

The data collected by the author are all from the official website of the National Bureau of Statistics (www.stats.gov.cn), and the monthly data from January 2005 to December 2017 are used for training fitting, and the monthly data from January 2018 to June 2018 are used to test the predicted effect.

The total retail sales of consumer goods (SLE) data is used as a "big circulation" indicator for real-time observation of changes in the domestic retail market, directly reflecting changes in domestic economic dynamics and demand. Its macro role is to estimate GDP in the short term instead of final consumption expenditure. The new indicator substitutions are based on the above factors. This indicator is the only explanatory variable for this study, and other indicators are explanatory variables or control variables.

The consumer price index (CPI) and the retail price index (RPI) reflect price trends and domestic inflation. Fluctuations in prices directly affect residents' short-term consumption decisions, which in turn affect overall commodity sales. These two indicators can be converted into the Multi-period consumer price index (JCPI) and the Multi-period retail price index (JRPI). The JCPI and JRPI calculation method is to set the initial year of data collection as the base period, set it as

the standard price unit, and convert it to the current price level at each time point according to the price index.

The freight volume (HY) can reflect the overall heat of China's commodity circulation. In the rapid development of e-commerce, the freight volume reflects the popularity and scale of retail sales to a certain extent. For traditional businesses, the turnover rate and sales volume of goods can also be reflected indirectly by the freight volume. The indicator adopted a new accounting method in 2014. There was a structural drop in the data values from December 2013 to January 2014. As shown in "Fig. 1", a dummy variable (FT) was introduced to mark this change.
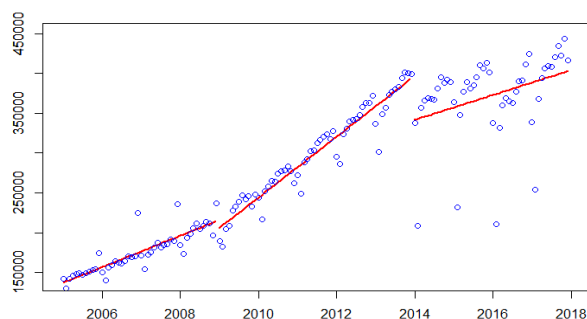

Fig. 1. Monthly data of freight volume.

The per capita consumption expenditure of urban residents (CC) and the per capita consumption expenditure of rural residents (RC) mainly reflect the average level of personal consumption. This indicator reflects social consumption and overall demand from the perspective of people's life. The urban population (CP) and the rural population (RP) are annual indicators. The author considers the equidistance difference to monthly, multiplied by the per capita consumption expenditure level to estimate the overall consumption expenditure.

The final consumption expenditure indicator is of comprehensive connotation and high level internationalization. The biggest problem is the slow frequency of data publication. This indicator is published on an annual frequency and generally only obtains data that is lagging two years behind, so it is not considered as an explanatory variable.

The information of each variable is shown in "Table I".

TABLE I.   VARIABLE SELECTION AND ITS DESCRIPTION

Variable name	Variable symbol	Variable description
The total retail sales of consumer goods	SLE	Wholesale and retail trade, accommodation and catering industry and other industries directly sold to urban and rural residents and social groups of consumer goods retail sales, monthly frequency data.
Consumer Price Index	CPI	The change in the price level of consumer goods and service items generally purchased by households, monthly frequency data.
Multi-period consumer price index	JCPI	Multi-period consumer price index converted to each time point
Commodity Retail Price Index	RPI	The price index of retail goods paid in cash or credit card, monthly frequency data.
Multi-period retail price index	JRPI	Multi-period retail price index converted to each time point
Freight Volume	HY	The quantity of goods actually transported by the transport company during a certain period of time. Monthly frequency data.
Freight volume adjustment variable	FT	This variable is a dummy variable that marks the data hopping caused by changes in accounting methods in January 2014 due to changes in accounting methods.
Urban population	CP	The annual data, divided into monthly according to the growth trend.
Rural population	RP	The annual data, divided into monthly according to the growth trend.
Per capita consumption expenditure of urban residents	CC	The total expenditure of urban residents' individuals and households for living consumption and collective use for personal consumption. The quarterly data is divided into monthly frequencies.
Per capita consumption expenditure of rural residents	RC	The total expenditure of rural residents' individuals and households for living consumption and collective use for personal consumption. The quarterly data is divided into monthly frequencies.

C. Model Design

The model design part mainly starts from the missing data imputation, and constructs the fitting and prediction model from the aspects of seasonal effect processing, static trend regression and trend fitting, aiming at the basic characteristics of monthly data on the total retail sales of consumer goods.

1) Price index conversion: Assuming that the acquired data have Q years, 12 months per year, the consumer price index of the n month of the m year can be expressed as CPI_{mn}, where $m = 1, 2, \cdots Q$, $n = 1, 2, \cdots, 12$. Recording the price of the base year as unit 1, the current consumption index of the n month of the m year can be expressed as formula (1).

$$JCPI_{mn} = CPI_{1n} \times CPI_{2n} \times \cdots CPI_{mn} \quad (1)$$

Similarly, the retail price conversion index of the commodity with the n month of the m year is the formula (2).

$$JRPI_{mn} = RPI_{1n} \times RPI_{2n} \times \cdots RPI_{mn} \quad (2)$$

The economic meaning of JCPI and JRPI is the current price level converted to each time point.

2) Data interpolation and seasonal effect processing: The monthly data of the total retail sales of consumer goods has data missing, and has obvious seasonality as shown in "Fig. 2". It is also necessary to consider the influence of the Spring Festival factor. The author marks the construction of dummy variables in the January and February of the Spring Festival, and uses the seasonal index method to interpolate the missing data.
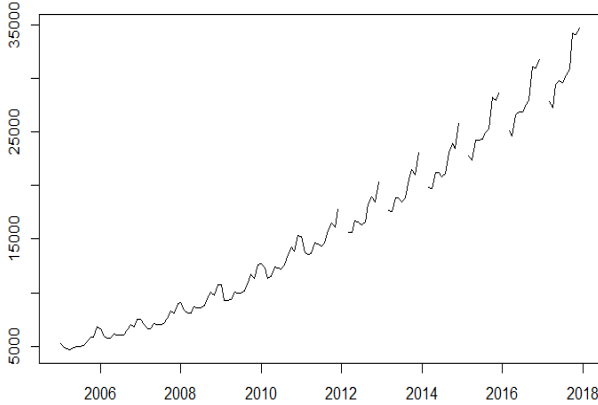

Fig. 2. Monthly total data of total retail sales of consumer goods.

For data with exponential trend in the model, it is necessary to perform logarithmic transformation and then extract seasonal factors. The role of the seasonal index method is to extract the seasonal factors in the data, and use the simple average method to calculate the relative value of seasonal effects in each period of the cycle. Assuming that the structure of Y of the original data is a period of N, and there are periods of M, the seasonal index S_k of k period can be expressed as formula (3).

$$S_k = M \frac{\sum_{i=1}^{N} y_{ik}}{\sum_{k=1}^{M} \sum_{i=1}^{N} y_{ik}} \quad (3)$$

Therefore, the monthly frequency data seasonal factor can be expressed as $S_1, S_2, S_3 \cdots, S_{12}$.

3) Static trend regression: Since the total retail sales of consumer goods have a clear index trend, it is taken logarithmically and the seasonal factors $LSLEM = \ln(SLE) - S(\ln(SLE))_k$ are removed for further analysis. $LSLEM$ is used as the explanatory variable. The explanatory variables include: the current price level converted to the current time point multiplied by the total freight volume, that is, $JRPI \times HY$ and $JCPI \times HY$. The author estimates the total consumption $CP \times CC + RP \times RC$ by multiplying the per capita consumption expenditure of the residents by the total population, and the dummy variables FT that mark the change in the statistical method of freight volume. The static trend fitting model of the total retail sales of consumer goods is shown in formula (4), where Z is a random error term, which satisfies the zero mean and the same variance hypothesis.

$$LSLEM = \beta_0 + \beta_1 JCPI \times HY + \beta_2 JRPI \times HY + \beta_3 (CP \times CC + RP \times RC) + \beta_4 FT + Z \quad (4)$$

4) Change trend fitting: In the previous step, the determined trend part of the model can be obtained, and the real data often has stochastic volatility, that is, the random error term Z in the above formula, and the second law of its fluctuation law needs to be fitted. Here, the SARIMA method is used to fit the prediction. The commonly used SARIMA model $ARIMA(p,d,q) \times (P,D,Q)_s$ structure is expressed as formula (5).

$$\begin{cases} \nabla^d \nabla_s^D z_t = \dfrac{\Theta(B)\Theta_s(B)}{\Phi(B)\Phi_s(B)} e_t \\ \Phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^P \\ \Phi_s(B) = 1 - \phi_1 B^s - \cdots - \phi_p B^{Ps} \\ \Theta(B) = 1 - \theta_1 B - \cdots - \theta_q B^q \\ \Theta_s(B) = 1 - \Theta_1 B^s - \cdots - \Theta_Q B^{Qs} \end{cases} \quad (5)$$

Among them, $\nabla^d \nabla_s^D z_t = (1-B)^d (1-B^S)^D z_t$, s is the period of the sequence, e_t is white noise sequence, $|\theta| < 1$, $|\Theta| < 1$. $\Phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^P$ and $\Phi_s(B) = 1 - \phi_1 B^s - \cdots - \phi_p B^{Ps}$ are autoregressive coefficient polynomials of the stationary reversible $ARMA(p,q) \times (P,Q)_s$ model. $\Theta(B) = 1 - \theta_1 B - \cdots - \theta_q B^q$ and $\Theta_s(B) = 1 - \Theta_1 B^s - \cdots - \Theta_Q B^{Qs}$ are the moving average coefficient polynomials of the stationary reversible model $ARMA(p,q) \times (P,Q)_s$. Therefore, the monthly fitting expression of total retail sales of consumer goods is $SLE = e^{LSLEM} \times e^{S_k} \times e^{Z_t}$, which can be decomposed into

trend factors $LSLEM$, seasonal factors S_k and irregular factors Z_t .

III. EMPIRICAL TEST

A. Data Interpolation and Seasonal Effect Processing

The monthly data of the total retail sales of consumer goods has obvious seasonal characteristics, which is increasing year by year and exhibits an exponential growth. At the same time, the monthly total data of total retail sales of consumer goods has a lack of data from 2012 to 2018. The official data gives the accumulated values in January and February, and the data needs to be completed by interpolation. The per capita consumption expenditure of urban residents and the per capita consumption expenditure of rural residents are quarterly data. Taking into account the short-term stickiness of prices and consumer demand, the monthly data of these two variables will follow the data values published in the quarter. The population data is published annually. Due to the stable changes in population data in recent years in China, the author uses the population data at the beginning and end of the year to make an equidistance difference for each month and then divide them into months. The CPI and RPI are monthly frequency data, which are processed into indicators JCPI and JRPI according to formulas (1) and (2).

It can be easily found from "Fig. 1" and "Fig. 2" that there are obvious seasonal characteristics of the total retail sales volume and freight volume monthly data of consumer goods. The seasonal factors for extracting these two indicators according to formula (3) are shown in "Table II" and "Fig. 3".

TABLE II. TOTAL RETAIL SALES OF CONSUMER GOODS AND SEASONAL FACTORS OF FREIGHT VOLUME

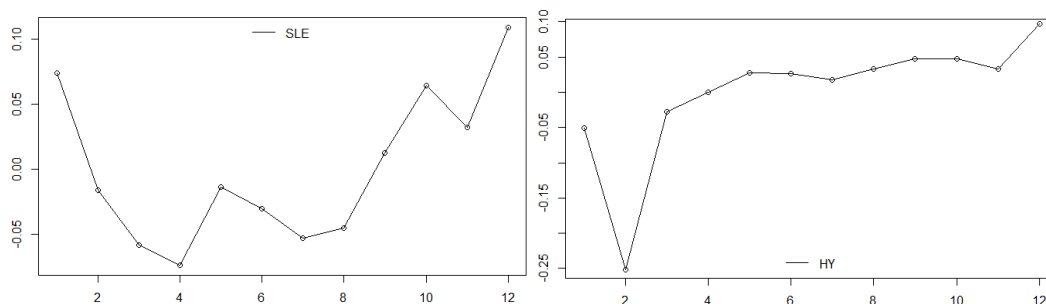|	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	S_{11}	S_{12}
SLE	0.0735	-0.0161	-0.0583	-0.0742	-0.0136	-0.0306	-0.0529	-0.0453	0.0121	0.0641	0.0323	0.1092
HY	-0.0510	-0.2518	-0.0276	0.0000	0.0278	0.0262	0.0181	0.0333	0.0475	0.0474	0.0329	0.0974

Fig. 3. Total retail sales of consumer goods, seasonal factors of freight volume.

It is easy to see that the seasonal characteristics of total retail sales of consumer goods are mainly represented by higher values at the beginning and end of the year, while the values in the second quarter are significantly falling. The freight volume has a trough in February due to the Spring Festival effect, and the values of other months are relatively stable.

B. Static Trend Regression

According to formula (4), the static trend of the total retail sales of consumer goods is analyzed by regression analysis. The results obtained are shown in "Table III". Model 1 and Model 2 respectively use the household consumption price conversion index multiplied by the freight volume, the commodity retail price conversion index multiplied by the freight volume as the explanatory variable to fit and replace, and the overall interpretation degree reaches 0.9787 and 0.9758 respectively. Model 3 uses the per capita consumption expenditure of urban residents multiplied by the urban population, the per capita consumption expenditure of rural residents multiplied by the sum of rural populations as explanatory variables, and the overall interpretation degree reaches 0.9948. Model 4 and Model 5 consider the mixture price multiplied by the freight volume, the per capita consumption expenditure of the residents multiplied by the population as the explanatory variable. The two models have the best fitting effect, the parameters are significant and the overall interpretation reaches 0.9961 and 0.996 respectively. Considering that the retail price index of goods can better reflect the price changes of physical goods, and the degree of compliance with freight volume indicators is higher, the author finally chooses model 5 as the optimal model.

TABLE III.   STATIC TREND FITTING ANALYSIS OF MONTHLY TOTAL RETAIL SALES OF CONSUMER GOODS

	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**
ln(JCPI*HY)	1.0910*** (0.0199)			0.2303*** (0.0342)	
ln(JRPI*HY)		1.1241*** (0.0219)			0.2173*** (0.0339)
ln(CC*CP+RC*CP)			1.3487*** (0.0078)	1.0636*** (0.0409)	1.0878*** (0.0392)
FT	0.22831** (0.0188)	0.2664*** (0.0197)		0.0561*** (0.0105)	0.0594*** (0.0109)
C	-9.4078*** (0.3411)	-9.9404*** (0.3750)	-17.0336*** (0.1541)	-15.4188*** (0.2734)	-15.6617*** (0.2569)
Residual s.e.	0.0818	0.0871	0.0403	0.0351	0.0355
Adjusted R^2	0.9787	0.9758	0.9948	0.9961	0.996
F-statistic	3560	3128	2.97e+04	1.31e+04	1.28e+04
p-value	< 2.2e-16	< 2.2e-16	< 2.2e-16	< 2.2e-16	< 2.2e-16

The overall interpretation of Model 5 reached 0.996, and the estimation of each parameter was significant. It is shown that the retail price conversion index of goods is multiplied by the freight volume, and the per capita consumption expenditure of the residents multiplied by the population. Generally speaking, it has a good explanatory for the total retail sales of consumer goods, and the obtained static trend fitting expression is as shown in equation (6).

$$\hat{LSLEM} = -15.6617 + 0.2173 ln(JRPI \times HY) + 1.0878 ln(CC \times CP + RC \times CP) + 0.0594 FT \tag{6}$$

C. Analysis of Change Trend Fitting

On the basis of the previous step, the random error term is extracted, and the parameter selection calculation is performed by using the auto.arima() function in the R language. The fitting results are shown in "Table IV".

TABLE IV.   FITTING RESULTS OF MONTHLY DATA CHANGE TRENDS OF TOTAL RETAIL SALES OF CONSUMER GOODS

ARIMA(1,0,2) with zero mean		
AR(1)	MA(1)	MA(2)
0.7442	-0.2993	0.2009
s.e. 0.0790	0.1045	0.0864

sigma^2 estimated as 0.0006639: log likelihood = 349.04, aic = -688.09

The stochastic fluctuations of the model are optimally fitted by ARIMA(1,0,2), and the fitting expression is constructed as $z_t = 0.7442 z_{t-1} + \varepsilon_t + 0.2993\varepsilon_{t-1} - 0.2009\varepsilon_{t-2}$ according to formula (5). Substituting the fitting value obtained in this step into $SLE = e^{LSLEM} \times e^{S_k} \times e^{Z_t}$ of the model is compared with the original data. The average relative error is -0.00012, the average absolute error is 0.017554, and the root mean square error is 342.27. The fitting effect is shown in "Table V" and "Fig. 4".

TABLE V.   MONTHLY DATA FITTING ERROR OF TOTAL RETAIL SALES OF CONSUMER GOODS

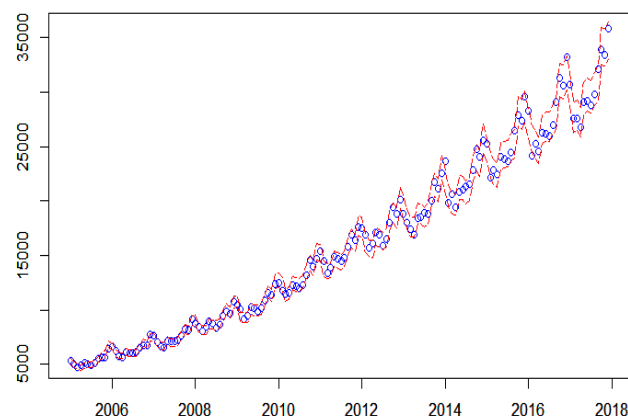| Mean relative error MPE	Mean absolute error MAPE	Mean square error MSE	Root mean square error RMSE
2.3773e-04	0.0199	224718.41	474.04

Fig. 4.   Monthly data fitting of total retail sales of consumer goods.

"Fig. 4" is a comparison of the monthly fitting data of the total retail sales of consumer goods from 2005 to 2017 and the confidence interval of the original data. The circle represents the fitting data, and the upper and lower curves represent the 95% confidence interval of the total retail sales of consumer goods. It can be seen that through the static trend of multiple indicators and dynamic trend fitting, the fitting data can better reflect the exponential trend, seasonal trend and change level of the original data, and the fitting

data basically fall within the 95% confidence interval. The model fitting effect is better.

D. Model Prediction

From January 2018 to June 2018, the retail price conversion index, freight volume, freight volume adjustment variable, urban population, rural population, per capita consumption expenditure of urban residents, and per capita consumption expenditure of rural residents will be substituted into the expression $SLE = e^{LSLEM} \times e^{S_k} \times e^{Z_t}$. The value of the predicted value is compared with the actual retail sales of consumer goods, as shown in "Table VI" and "Table VII".

TABLE VI. COMPARISON OF MONTHLY FORECAST DATA OF TOTAL RETAIL SALES OF CONSUMER GOODS

Time	Static trend	Seasonal factor	Trend	Predictive value	Actual value	95% confidence interval
January, 2018	10.3309	0.0735	-0.0026	32918.48	31371.20	[29802.63，32939.74]
February, 2018	10.3048	-0.0161	-0.0080	29161.52	29710.60	[28225.08，31196.15]
March, 2018	10.3624	-0.0583	-0.0059	29676.12	29193.60	[27733.92，30653.28]
April, 2018	10.3569	-0.0742	-0.0044	29092.37	28541.90	[27114.81，29967.00]
May, 2018	10.3692	-0.0136	-0.0033	31327.62	30359.10	[28841.15，31877.06]
June, 2018	10.3776	-0.0306	-0.0025	31085.00	30841.60	[29299.52，32383.68]

TABLE VII. MONTHLY DATA PREDICTION ERROR OF TOTAL RETAIL SALES OF CONSUMER GOODS

Mean relative error MPE	Mean absolute error MAPE	Mean square error MSE	Root mean square error RMSE
0.017742	0.023902	704778.88	839.51

According to the composition of the predicted values, the static trend portion accounts for more than 98% of the predicted value, indicating the retail price conversion index, freight volume, freight volume adjustment variable, urban population, rural population, per capita consumption expenditure of urban residents, and per capita consumption expenditure of rural residents has played a leading role in the substitution of monthly indicators of total retail sales of consumer goods. Seasonal factors and trending factors account for a small proportion of the predicted values, and only serve as a supplementary adjustment mechanism, and as an explanatory variable for the regular changes in the total retail sales of consumer goods and stochastic volatility. The average relative error of the model prediction value is 0.017742, the average absolute error is 0.023902, and the root mean square error is 839.51. The prediction effect is better.
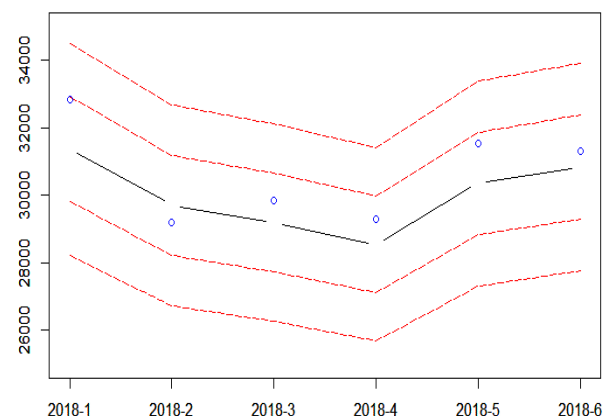

Fig. 5. Monthly prediction of total retail sales of consumer goods.

"Fig. 5" is a comparison of the deviation between the predicted and actual value of the monthly total retail sales of consumer goods, and its 95% and 90% confidence intervals. It can be seen that one month of the 6-month forecast value falls on the 95% confidence interval boundaries, and all within the 90% confidence interval. In general, the predicted values are within a reasonable range, and the predicted values for most of the months are slightly higher than the actual values, due to the slowdown of the economic situation, structural adjustments, and the economic effects caused by the "three to one and one subsidy" policy. As a result, the growth of business and trade has become smaller than in previous years.

Based on the empirical analysis, it is found that from the perspective of data fitting and trend forecasting, total retail sales of consumer goods can be replaced by commodity retail price conversion index, freight volume, freight volume adjustment variable, urban population, rural population, per capita consumption expenditure of urban residents, and per capita consumption expenditure of rural residents. And the

static trend accounts for 98%, while the seasonal factor and random fluctuation factor account for a small proportion.

IV. CONCLUSION

The article uses eight indicators of consumption as the explanatory variables to construct a fitting and prediction model for the monthly total retail sales of consumer goods and use actual data from January 2005 to June 2018 to conduct an empirical study, indicating a close relationship between the total retail sales of consumer goods and other related categories of consumption. On the one hand, the current price converted from the price index is multiplied by the total freight volume, which is the use of "mixed price" and "mixture amount" to estimate the total retail sales of consumer goods. This part focuses on the measurement of "physical consumption". On the other hand, the per capita consumption expenditure of residents is multiplied by the total population, that is, "per capita consumption" and "population" to estimate the total retail sales of consumer goods. This part contains the full-caliber consumption types, including physical consumption, service consumption and virtual consumption. The result is good with 1.99% fitting MAPE and 2.39% prediction MAPE.

Although the total retail sales of consumer goods has realized a set of standard system for direct collection of monthly data, considering the error of data collection and measurement, the author proposes a fitting and prediction model based on consumer indicators, and uses relevant indicators to verify and supplement the actual measurement errors. At the same time, the model built by the author not only contains physical consumption, but also contains the measurement of service consumption. Compared with the index connotation of the total retail sales of consumer goods, it can reflect the overall consumption level of China's real-time changes more comprehensively. The analysis of supply and demand has better guiding significance in the macroeconomic dimension, economic heat, policy effects, and people's lives.

REFERENCES

[1] Huang Yan, An Empirical Analysis of the Total Retail Sales of Consumer Goods in China Based on ARIMA Model [J]. Economic Tribune, 2011, (11): 31 - 37. (in Chinese)

[2] Liu Lingpo, Time Series Model and Forecast of Total Retail Sales of Consumer Goods in China [J]. Economic Tribune. 2011 (6): 5 - 8. (in Chinese)

[3] Jiang Hui, Zhang Po, ARMA Criterion of Feature Extraction in Support Vector Regression: Empirical Research of Our National Social Total Retail Sales of Consumer Goods Prediction [J]. Statistics & Information Tribune, 2012, 27 (07): 3 - 7. (in Chinese)

[4] Luo Zhongde, Lai Meiyan, Forecast Analysis of Total Retail Sales of Consumer Goods in China [J]. Statistics and Decision, 2013 (2): 143 - 145. (in Chinese)

[5] Deng Weiru, He Rui, Pu Yanxi, Research on Forecast Model of Total Retail Sales of Social Consumption Goods [J]. Statistics and Decision, 2014 (4): 24 - 27. (in Chinese)