The Influence and Model Analysis of China's Strategic Emerging Industries on CPI and PPI

---- Empirical Analysis based on Chinese Data from January 2021 to August 2024

Qianyu Li

Business School, University of Birmingham, Birmingham B15 2TT, United Kingdom

2646104208@qq.com

Abstract. Strategic emerging industries are the new pillar and new track leading China's future development, and are the main fields to cultivate and develop new quality productivity. The consumer Price index (CPI) and the producer price index (PPI) are important economic indicators. This paper selects Zhanxin industry, CPI and PPI data released by authoritative channels in China from January 2021 to August 2024 for analysis. Innovatively, 20 industries in new industries were compared with CPI and PPI data to carry out trend, principal component and correlation empirical analysis, conduct stationary root analysis, and establish Almon prediction model, which is cutting-edge, targeted and practical. The study found that Zhanxin industry has a high correlation with CPI and PPI data, and most of the data have a positive interaction relationship, and a few have a negative impact relationship. The research specifically proposed that the development of new industries should do a good job in the investment and guarantee of electricity consumption in the secondary industry and residential electricity consumption, strengthen the supply and price guarantee of liquefied natural gas and natural gas, highlight the development of non-fuel oil products, pay attention to the security of raw coal and coal-bed methane, promote the green and low-carbon energy transformation, and improve the security guarantee ability of the industrial chain and supply chain. The study is of great significance to the implementation of the Third Plenary Session of the 20th Central Committee to improve the strategic industry development policy and governance system, and to cultivate and develop new quality productivity.

Keywords: Strategic emerging industries; The energy industry; Impact analysis.

1 Introduction

1.1 Research Background

At present, China is vigorously developing strategic emerging industries (referred to as "new industries"), which represent the frontier direction and growth point of future

K. Zhang et al. (eds.), *Proceedings of the 5th International Conference on Economic Management and Big Data Application (ICEMBDA 2024)*, Advances in Economics, Business and Management Research 313,
https://doi.org/10.2991/978-94-6463-638-3_42

development, constitute a new source of power and competition track in China, and are the key areas to shape and promote the growth of new productivity. The Third Plenary Session of the 20th Central Committee of the Communist Party of China decided to improve the system and mechanism for developing new quality productive forces in accordance with local conditions, vigorously develop and strengthen the institutional supply of new areas and new track, and establish a mechanism for future industrial input growth. We will improve policies and governance systems for the development of strategic industries such as next-generation information technology, artificial intelligence, aerospace, new energy, new materials, high-end equipment, biomedicine, and quantum science and technology, and guide the healthy and orderly development of emerging industries. The Third Plenary Session of the 20th CPC Central Committee also made it clear that it should improve the statistical index accounting system supporting high-quality development and strengthen the integration coverage of new economy and new fields [1]. In the 14th Five-Year Plan and the outline of the 2035 vision goals, the Chinese government clearly proposes to develop and strengthen strategic emerging industries to achieve high-quality economic development and the country's long-term development goals. In the future, China will focus on the development of a new generation of information technology industries, such as 5G communications, artificial intelligence, big data, cloud computing, etc. Biotechnology industry: including biomedicine, biological agriculture, biological manufacturing, etc. New energy industry: such as solar energy, wind energy, new energy vehicles, etc. New materials industry: involving advanced structural materials, functional materials, composite materials, etc. High-end equipment manufacturing industry: including intelligent manufacturing equipment, aerospace equipment, rail transit equipment, etc. Green environmental protection industry: such as energy saving and environmental protection technology, resource recycling, environmental governance, etc. Digital creative industry: Combining digital technology and creative content, such as digital media, animation games and other industries.

The Consumer Price Index (CPI) and the Producer Price Index (PPI) are two important economic indicators. CPI is a measure of the relative number of changes in the price level of a basket of consumer goods and services purchased by residents over time, reflecting the trend and degree of change in the price level of consumer goods and services, and is usually used to measure the degree of inflation or deflation. It is an important indicator for macroeconomic analysis and decision-making, monitoring and regulation of the general price level and national economic accounting[2]. The compilation method of CPI is to select representative specifications among many measurement objects, collect their price data, calculate the average price and individual price index of these representative specifications, and then use the geometric average to calculate the basic price index of categories, and then use the weighted average method to calculate the price index of small, middle and large categories step by step, until the total price index. CPI summary calculation method adopts chain pull formula to compile month-on-month, month-on-year and fixed base price index. The CPI is equal to 100, indicating that the comprehensive price has not changed compared with the base period. The consumer price index is greater than 100, indicating that compared with the base period, the overall price has increased. The higher the price index, the more the price has increased. When the consumer price index is less than 100, it indicates that the overall price has declined in the reporting period compared with the base period [3]. According to the International Monetary Fund (IMF) "Producer Price Index Manual", PPI should theoretically cover the producer price index of all industries, that is, including agricultural producer price index, industrial producer price index and service producer price index. The industrial producer price index is an important part of it. Due to historical and technical reasons, the PPI compiled and released by many countries only refers to the industrial producer price index [4]. China's PPI only refers to the industrial producer price index, the survey catalogue contains 41 industrial categories, 207 industrial categories, 666 industrial sub-categories, 1,638 basic categories, more than 20,000 representative products. The price survey is reported monthly, and the survey date is the 5th and 20th of the survey month. The calculation method of PPI and the calculation method of consumer price index (CPI) and the conversion method of various price indexes are basically the same. There is a certain conduction relationship between PPI and CPI. If the increase in PPI can be transmitted to CPI, it indicates that the pressure of rising prices has spread from the production field to the consumption field. If PPI rises while CPI remains stable or falls, it could indicate that the upward pressure on prices has not yet fully passed through to consumers and businesses may be absorbing higher cost pressures. It is generally accepted that moderate inflation is beneficial. When prices go up, manufacturers are encouraged to produce more, and people are willing to consume more because they are worried about future price increases, so demand and supply in the economy will increase, and the economy will be better. If deflation occurs, commodity prices fall, manufacturers are not highly motivated to produce, people tend to save money rather than consume, which is not conducive to the development of economic cycle.

1.2 Literature review

From the perspective of literature and data research, there are not many articles on the impact of strategic emerging industries on CPI and PPI and their models, mainly focusing on the research of development mode and operation mechanism. Sun Tianyang and Yang Danhui (2022) systematically reviewed the latest important literature on emerging industries at home and abroad, and found that the current research mainly focused on the definition and statistics of emerging industries, the effect assessment of emerging industry policies, and the influencing factors of the development of emerging industries. Relevant achievements have enriched the understanding of the general law and operation mechanism of the development of emerging industries, but more in-depth discussions are still needed in terms of data accumulation, research paradigm and theoretical innovation [5]. Dai Peichao, Zhang Rongjia and Zhou Yuan (2023) used CNKI database to coidentify 7,659 sample literatures and analyze the research trends, time period characteristics, author distribution, institution distribution and literature periodical distribution of strategic emerging industries respectively. Current research focuses on six aspects of strategic emerging industries: model selection, development evaluation, spatial layout, innovation efficiency, collaborative development, and the relationship with financial, fiscal and tax policies. It is found that the study of strategic emerging industries has gradually evolved from management to a comprehensive study integrating statistics, finance, geography, ecology and other disciplines. In the future, it is necessary to focus on research directions such as dynamic collaborative innovation ecosystem, application of new evaluation methods, action mechanism of policy tools, and industrial ecological governance system [6].

Research on the impact of CPI and PPI mainly focuses on the impact of traditional products. Deng Haiyun, Zhang Xi, Yang Junbo, Liu Qing and Yin Jiabin (2023) used multiple linear regression (MLR) and grey prediction model GM(1,N) to fit the CPI of urban and rural residents in China with all kinds of CPI data, and compared and analyzed the effectiveness and difference of the two models. At the same time, the time series prediction model of urban and rural CPI is established by using Gaussian model. The research results show that MLR and GM(1,N) models both have high prediction accuracy, and the Gaussian prediction model of urban and rural CPI about time also shows a good goodness of fit coefficient of determination. The main factors that influence CPI discussed in the literature include food, tobacco and alcohol, housing, transportation and communication, etc. These factors have a strong correlation with the total CPI in urban and rural areas, in which housing, food, tobacco and alcohol have a particularly significant impact on the total CPI in rural areas [7]. The International Monetary Fund (IMF) researches on CPI and PPI mainly focus on commodity price forecast, world consumer price forecast, energy price forecast, service industry inflation forecast, etc. [8].

1.3 Research Value

The research is highly targeted, and the analysis content is the strategic emerging industries that the Chinese Party and government are most concerned about and pay attention to, CPI which is closely related to the people, and PPI which is closely related to enterprise management. The analysis sample size covers a wide range, and 20 types of data are innovatively selected for research in strategic industries. The sample data is new, using the most recent data from January 2021 to August 2024. The research conclusions, recommendations and prospects are highly targeted and of great significance for further implementing the spirit of the Third Plenary Session of the 20th Central Committee of the Communist Party of China, further promoting the development of strategic emerging industries, improving the CPI and PPI statistical system covering strategic emerging industries, vigorously developing and strengthening the supply of new fields and new tracks, and accelerating the cultivation of new quality productivity. In terms of research methods, this study uses directivity analysis, principal component analysis, stationary root analysis, Almon analysis and other methods to carry out large order of magnitude analysis of 22 variables, which also has strong academic reference value for better improving the analysis method of large variable data.

2 Research Data

2.1 Strategic Emerging Industry Data

According to the "New Industry Standardization Pilot Project Implementation Plan (2023-2035)" issued by the Ministry of Industry and Information Technology and other four departments, strategic emerging industries mainly include "8 emerging industries +9 future industries" [9]. Among them, the eight emerging industries refer to the new generation of information technology, new energy, new materials, high-end equipment, new energy vehicles, green environmental protection, civil aviation, shipping and Marine engineering equipment and other fields. The nine major future industries cover the meta-universe, brain-computer interface, quantum information, humanoid robots, generative artificial intelligence, biological manufacturing, future display, future network, and new energy storage. These industries are characterized by active innovation, intensive technology, and broad development prospects, and are of great significance to national economic and social development and industrial structure optimization and upgrading.

To ensure the accuracy and compliance of the study, this paper collected data from January 2021 to August 2024, The National Development and Reform Commission of the People's Republic of China (referred to as the National Development and Reform Commission), the Ministry of Industry and Information Technology of the People's Republic of China (referred to as the Ministry of Industry and Information Technology), the National Bureau of Statistics of the People's Republic of China (referred to as the National Bureau of Statistics), the China Association of Automobile Manufacturers (referred to as the China Association of Automobile Manufacturers) and other units announced in the Internet open channels of new energy vehicles, solar cells (photovoltaic cells) Representative products of 20 industrial enterprises above designated size, such as lithium-ion batteries and integrated circuits, are used as research sample data of strategic emerging industries.

Due to the Spring Festival and other factors, part of the data of the National Bureau of Statistics will generally only publish the cumulative number of January and February. In this paper, the output data of January and February in such data are processed by interpolation method according to equations (1) and (2).

$$X_1 = \frac{C_2}{\frac{X_3 - X_{-12}}{\mathrm{X3}} + 2} \tag{1}$$

$$X_2 = C_2 - X_1 \tag{2}$$

In formula (1) and (2), X1 is the output of January, X2 is the output of February, C2 is the cumulative output of February, X3 is the output of March this year, and X-12 is the output of December last year.

In order to facilitate data analysis, the data of strategic emerging industries in this paper, the output of new energy vehicles (10,000 units), is coded as X1; Solar cell (photovoltaic cell) production current value (million kilowatts), coded as X2; The current value of lithium-ion battery production (10,000 (naturally only)), coded as X3; Current value of integrated circuit output (billion pieces), coded as X4; The current output value of the whole computer (10,000 units), the code is X5; The current output value of electrical instruments and meters (10,000 units), coded X6; Current value of EMU output (units), coded as X7; Industrial robot output current value (set), coded as X8; Current output value of service robot (set), coded as X9; The current value of the output of optoelectronic devices (hundreds of millions (pieces, sets)), coded as X10; Optical cable output current value (million core kilometers), code X11; Current value of mobile communication handset (mobile phone) production (10,000 units), coded as X12; The current output value of mobile communication base station equipment (ten thousand RF modules), the code is X13; Smartwatch production current value (10,000), coded as X14; The current value of the production of chemical agents (10,000 tons), coded as X15; The current value of the production of proprietary Chinese medicine (10,000 tons), the code is X16; Current value of solar power generation (billion KWH), coded as X17; Current value of nuclear power generation (billion KWH), coded X18; Current value of wind power generation (billion KWH), coded X19. Relevant data are shown in Table 1 and Table 2. Due to the large amount of data, only data from January, June and December 2021, 2022 and 2023 and data from January to August 2024 are retained in the table. Relevant data can be downloaded from relevant websites of the Chinese government or obtained from corresponding authors. In this paper, all 44 months of data from January 2021 to August 2024 are used for calculation.

Table 1. Strategic emerging Industries X1-X9 data sheet from January 2021 to August 2024.

Month	X1	X2	X3	X4	X5	X6	X7	X8	X9
2021-1	19.4	1461	156121	259	3336	1861	212.0	21497	493058
2021-6	24.8	1741	203061	308	4300	2613	245.0	36383	817251
2021-12	51.8	2269	231166	299	4664	2975	5.0	35175	902551
2022-1	45.2	1960	181370	294	3238	1790	1.9	33796	610309
2022-6	59	2702	200299	288	4334	2402	68.0	46144	476557
2022-12	79.5	3562	212966	284	3806	2540	164.0	40457	490992
2023-1	42.5	2846	163976	218	2493	1552	49.5	29758	311580
2023-6	78.4	4698	206955	317	3278	2661	360.0	39974	662658
2023-12	117.2	5083	246002	362	3256	3124	144.0	41980	738284
2024-1	78.7	3355	184174	352	2316	1758	150.0	34450	433290
2024-2	46.4	3638	184396	352	2229	1705	50.0	41542	562657
2024-3	86.3	5511	246299	362	3133	3030	48.0	50623	958713
2024-4	87	4620	236910	376	2815	2651	25.0	50380	973048
2024-5	94	4982	240717	355	2883	2776	16.0	51496	912013
2024-6	100.3	4914	243897	362	3119	2878	459.0	53088	846704
2024-7	98.4	4597	229173	375	2920	2546	17.0	45528	879091
2024-8	109.2	4515	228912	373	3251	2357	72.0	47947	989470

[a] Sources: National Development and Reform Commission, Ministry of Industry and Information Technology, National Bureau of Statistics, China Automobile Association.

Table 2. Strategic emerging Industries X1-X9 data sheet from January 2021 to August 2024

Month	X10	X11	X12	X13	X14	X15	X16	X17	X18	X19
2021-1	1990	11386	36.4	404	20.3	18.2	105	296	424	1990
2021-6	2673	14322	47.1	504	25.2	20.8	161	359	386	2673
2021-12	3111	17461	61.5	748	31.7	24.6	142	373	571	3111
2022-1	2364	11663	48.7	556	26.9	25.7	127	330	433	2364
2022-6	3153	13612	69.2	584	30.6	21.2	206	327	524	3153
2022-12	2430	14310	55.8	627	36.9	26.4	162	397	723	2430
2023-1	1703	10219	49.0	359	29.3	21.7	148	348	663	1703
2023-6	3001	12604	61.2	791	27.2	18.4	259	372	558	3001
2023-12	2659	16935	76.7	604	40.1	21.4	211	381	814	2659
2024-1	1923	12813	39.2	499	30.3	16.6	198	361	709	1923
2024-2	1621	10593	22.6	543	24.3	14.9	292	330	788	1621
2024-3	2242	14001	44.1	657	32.2	19.1	311	348	905	2242
2024-4	2069	12586	71.8	520	30.2	17.5	314	366	808	2069
2024-5	2419	12703	40.4	649	30.7	16.3	359	360	772	2419
2024-6	2273	13291	72.4	759	31.5	16.8	352	357	669	2273
2024-7	2280	12830	29.0	707	27.2	13.4	360	397	661	2280
2024-8	2285	13648	28.0	823	25.2	12.2	386	402	502	2285

[a] Sources: National Development and Reform Commission, Ministry of Industry and Information Technology, National Bureau of Statistics, China Automobile Association.

2.2 CPI and PPI Data

The statistics of CPI and PPI are relatively sound, and the National Bureau of Statistics regularly publishes the national CPI and PPI statistics every month. To facilitate data analysis, CPI code is Y1 and PPI code is Y1. Relevant data are shown in Table 3.

Table 3. CPI and PPI Data from January 2021 to August 2024

Month	CPI	PPI	Month	CPI	PPI	Month	CPI	PPI
2021-1	99.7	100.3	2022-4	102.1	108	2023-7	99.7	95.6
2021-2	99.8	101.7	2022-5	102.1	106.4	2023-8	100.1	97
2021-3	100.4	104.4	2022-6	102.5	106.1	2023-9	100	97.5
2021-4	100.9	106.8	2022-7	102.7	104.2	2023-10	99.8	97.4
2021-5	101.3	109	2022-8	102.5	102.3	2023-11	99.5	97
2021-6	101.1	108.8	2022-9	102.8	100.9	2023-12	99.7	97.3
2021-7	101	109	2022-10	102.1	98.7	2024-1	99.2	97.5
2021-8	100.8	109.5	2022-11	101.6	98.7	2024-2	100.7	97.3
2021-9	100.7	110.7	2022-12	101.8	99.3	2024-3	100.1	97.2
2021-10	101.5	113.5	2023-1	102.1	99.2	2024-4	100.3	97.5
2021-11	102.3	112.9	2023-2	101	98.6	2024-5	100.3	98.6
2021-12	101.5	110.3	2023-3	100.7	97.5	2024-6	100.2	99.2
2022-1	100.9	109.1	2023-4	100.1	96.4	2024-7	100.5	99.2
2022-2	100.9	108.8	2023-5	100.2	95.4	2024-8	100.6	98.2
2022-3	101.5	108.3	2023-6	100	94.6	--	--	--

[a] Sources: National Development and Reform Commission, Ministry of Industry and Information Technology, National Bureau of Statistics, China Automobile Association.

3 Research Method

3.1 Linear Trend

Simple linear trend analysis is a statistical method used to assess the tendency of one variable (dependent variable) to change with another variable (independent variable). In simple linear trend analysis, we usually assume that the relationship between two variables can be described by a straight line. For a simple linear model, see formula (3).

$$y = \beta_1 X + \varepsilon \tag{3}$$

In equation (1), y is the dependent variable, β1 is the slope, and Ɛ is the intercept.

In linear analysis, R-squared (also known as the coefficient of determination) is a statistical measure of how well a model fits the data. See formula (4) for the calculation of R-squared value.

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} \tag{4}$$

In formula (4), SSres are the sum of squares of residuals, that is, the sum of squares of the difference between the actual observed value and the predicted value of the model. SStot is the total sum of squares, that is, the sum of squares of the difference between the actual observed value and the mean of the observed value. The r-squared values range from 0 to 1 and can be interpreted as the percentage of the variance explained by the model as a percentage of the total variance. An R-squared value of 0 indicates that the model does not explain the variance of any data, that is, all data points fall on the mean, and the model has no predictive power. An R-squared value of 1 indicates that the model perfectly fits all data points, that is, all data points are on the regression line, and the model has full predictive power.

3.2 Spearman Rank Correlation Coefficient

Spearman's Rank Correlation Coefficient, usually denoted by the Greek letter ρ (rho), is a non-parametric statistical method used to measure the monotonic relationship between two variables. It does not require that the data be normally distributed, nor does it require that the relationship between variables be linear. The calculation formula of Spearman's rank correlation coefficient is formula (5).

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2-1)} \tag{5}$$

Where n is the logarithm of the observations and di is the rank difference of each pair of observations. Spearman's rank correlation coefficient ranges from -1 to +1. +1 represents a completely positive monotonic relationship, that is, when one variable increases, the other variable always increases; -1 represents a completely negative monotone relationship, that is, when one variable increases, the other variable always decreases; 0 means there is no monotonic relationship. The advantage of Spearman's rank correlation coefficient is that it is applicable in a wide range and does not require data to follow a specific distribution, which is very useful for data that is not normally distributed or where there are outliers. However, if the data is actually normally distributed, using the Spearman rank correlation coefficient may reduce accuracy because it does not take into account the actual value of the data, only the rank.

3.3 Principal Component Analysis

Principal Component Analysis (PCA) is a statistical method that transforms a set of possibly correlated variables into a set of linearly uncorrelated variables, called principal components, by orthogonal transformation. This method is often used in dimensionality reduction, data compression, and pattern recognition. The steps of PCA usually include, standardizing the data, standardizing the data so that the mean of each feature is 0 and the standard deviation is 1. Calculate the covariance matrix to provide a measure of the linear relationship between variables. Eigenvalues and eigenvectors are calculated, and the covariance matrix is decomposed to obtain eigenvalues and eigenvectors. The eigenvalues represent the variance contribution of each principal component, while the eigenvectors define the direction of the principal component. The principal components are selected, based on the size of the eigenvalues, and the eigenvectors corresponding to the first few largest eigenvalues are selected as the principal components, which can explain most of the variability in the data set. Transform to the new space and project the original data onto these principal components, resulting in a new data set with a smaller dimension than the original. The mathematical expression of the principal component (PC) is shown in equation (6).

$$\begin{cases} \mathrm{PC}_1 = u_{11}X_1 + U_{21}X_2 + \ldots + u_{p1}X_p \\ \mathrm{PC}_2 = u_{11}X_1 + U_{22}X_2 + \ldots + u_{p2}X_p \\ \ldots \\ \mathrm{PC}_p = u_{1p}X_1 + U_{2_p}X_2 + \ldots + u_{\mathrm{pp}}X_p \end{cases} \tag{6}$$

In formula (6), X1, X2,... Xp is P primitive variables, PC1, PC2,... PCP is the p principal component, and upi represents the weight of the p principal component. The above model can be solved by correlation coefficient matrix or covariance matrix of p -dimensional raw data. It is common to consider the contribution of each principal component to the variance of the data set, and to select the number of principal components that make the cumulative contribution reach a particular value, such as 85% or 90%.

3.4 Data Stationarity Test

CPI and PPI forecasting is of great significance to national macro-control and economic development. Currently, there are many models to forecast CPI and PPI, and the forecasting method based on time series model is mainly adopted at present. Data stationarity analysis (ADF) is usually required before modeling to determine whether the time series data has a statistical test of unit root, that is, it tests whether the data is non-stationary. The unit root means that the time series has a lasting effect, that is, past changes have a long-term effect on the current value, which causes the fluctuation of the data not to revolve around a constant mean, but to drift over time. The test formula is shown in formula (7).

$$\Delta y_t = a + \beta_t + \gamma y_{t-1} + \sum_{i=1}^{p} \varphi_i \Delta y_{t-i} + \varepsilon_t \tag{7}$$

In Formula 7, Δyt is the first-order difference, γ is the autoregressive coefficient, yt-1 is the lag term, φi is the coefficient of the lag difference term, alpha is the constant term, β is the time trend term, and p is the order of the lag. If the P-value is less than a certain significance level (e.g. 0.05), the null hypothesis is rejected and the time series is considered stationary. If the p-value is large, the null hypothesis cannot be rejected and the time series may have a unit root, that is, non-stationary. Stationary series, statistical properties do not change with time, easy to predict. Non-stationary series, whose statistical properties change over time and are difficult to predict. If a sequence is determined to be non-stationary, it usually needs to be differenced or some other transformation to make it a stationary sequence before further analysis and modeling.

3.5 Almon's Estimation Method

The Almon estimation method is an econometric method for dealing with distributed lag models. In economics and econometrics, distributed lag models are used to analyze the dynamic process of the effect of one variable on another, especially when the effects of policy changes, investments, or other actions unfold over time. Almon's estimation method approximates the Lag structure by Polynomial Distributed Lag (PDL), so as to estimate the cumulative and individual lag effects more effectively. Almon's estimation method is shown in formula (8).

$$y_t = \beta_0 + \sum_{j=1}^{m} \beta_j P_j(L) + \varepsilon_t \tag{8}$$

In Formula 8, β0 is the constant term and βj is the polynomial coefficient, representing the influence of the j-order polynomial on the dependent variable. Pj (L) is a polynomial of order j about the lag operator L, where L represents the lag operator. εt is the error term. Almon estimation method is widely used in economic policy analysis, investment effect evaluation and other fields.

4 Research process

4.1 Trend Analysis and Research

According to equation (3), EXCEL software was used for linear trend analysis of X1-X19, Y1 and Y2 data, and linear trend equation of each dependent variable and determination coefficient R2 value were generated. Specific values were shown in Table 4.

Table 4. Linear trend equations and determination coefficients for X1-X19, Y1, Y2

Dependent variable	Linear trend equation	Coefficient of determination	Dependent variable	Linear trend equation	Coefficient of determination
X1	y=1.9638x+17.466	0.8023	X12	y=10.521x+13060	0.0054
X2	y=94.248x+1130.1	0.8529	X13	y=-0.0447x+55.759	0.0013
X3	y=1119.4x+177814	0.3428	X14	y=4.4385x+471.99	0.1759
X4	y=1.7591x+261.51	0.3125	X15	y=0.0989x+25.897	0.1016
X5	y=30.554x+4098.4	0.3939	X16	y=-0.1005x+21.313	0.1751
X6	y=7.623x+2160.8	0.0592	X17	y=4.7403x+105.91	0.7631
X7	y=0.8295x+69.017	0.0117	X18	y=1.0479x+328.81	0.2753
X8	y=388.49x+28890	0.4732	X19	y=7.8177x+410.18	0.4725
X9	y=4798.6x+563800	0.1180	Y1	y=-0.0275x+101.51	0.1353
X10	y=14.762x+794.25	0.5081	Y2	y=-0.3288x+109.58	0.5804
X11	y=-6.8355x+2773	0.0372	——	——	——

[a] Sources: National Development and Reform Commission, Ministry of Industry and Information Technology, National Bureau of Statistics, China Automobile Association.

As can be seen from Table 4, during the 44-month period from January 2021 to August 2024, the determining value R2 of X1, X2 and X17 is greater than 0.7, and the fitting of the linear trend equation is relatively good, while the fitting of other variables is average. In particular, the determination value R2 of X7, X9, X11, X12, X13, X14, X15, X16 and Y1 variables is less than 0.2, and the jump is large and not linear.

4.2 Spearman Grade Correlation Coefficient Analysis

From the perspective of trend analysis, during January 2021 to August 2024, there is less linear data between CPI and PPI of strategic emerging industries. Instead of using Pearson correlation coefficient analysis, this paper adopts a wide range of application, which does not require data to follow a specific correlation coefficient of Subbuspilman level. The data in Table 1, Table 2 and Table 3 are analyzed, and the calculation results are shown in Table 5.

Table 5. X1-X19, Y1, Y2 Spearman grade correlation coefficients

Dependent variable	category	X1	X2	X3	X4	X5	X6	X7
Y1	Correlation coefficient	-.355*	-.445**	-0.229	-.541**	.501**	-0.13	-.299*
	Sig. (Double tail)	0.018	0.002	0.135	0	0.001	0.399	0.049
Y2	Correlation coefficient	-.688**	-.811**	-0.256	-0.256	.678**	-0.104	-0.22
	Sig. (Double tail)	0	0	0.094	0.093	0	0.503	0.151
Dependent variable	category	X8	X9	X10	X11	X12	X13	X14

Y1	Correlation coefficient	0.012	-.333*	-.578**	.436**	0.096	.426**	-0.1
	Sig. (Double tail)	0.937	0.027	0	0.003	0.534	0.004	0.517
Y2	Correlation coefficient	-.388**	-0.108	-.537**	0.153	0.127	0.046	-0.274
	Sig. (Double tail)	0.009	0.487	0	0.322	0.412	0.766	0.072
Dependent variable	category	X15	X16	X17	X18	X19	Y1	Y2
Y1	Correlation coefficient	0.167	.484**	-.360*	-0.156	-0.269	1	.621**
	Sig. (Double tail)	0.278	0.001	0.016	0.311	0.078	.	0
Y2	Correlation coefficient	-0.119	.447**	-.722**	-.370*	-.621**	.621**	1
	Sig. (Double tail)	0.441	0.002	0	0.013	0	0	.

[a] Note: ** At 0.01 level (double-tailed), the correlation was significant. * At level 0.05 (two-tailed), the correlation was significant.

As can be seen from Table 5, during the 44 months from January 2021 to August 2024, Y1 is negatively correlated with X2, X4 and X10, and positively correlated with X5, X11, X13, X16 and Y2. At the level of 0.01 (double-tail), the correlation is significant. There was a certain negative correlation with X7, X9 and X17, and the correlation was significant at 0.05 level (double-tailed). Low correlation with X3, X6, X8, X12, X14, X15, X18, X19.

As can be seen from Table 5, during the 44 months from January 2021 to August 2024, Y2 is negatively correlated with X1, X2, X8, X10, X17 and X19, and positively correlated with X5, X16 and Y1. At 0.01 level (double-tail), the correlation is significant. There was a certain negative correlation with X18, and the correlation was significant at 0.05 level (two-tailed). Low correlation with X3, X4, X6, X7, X9, X11, X12, X13, X14, X15.

4.3 Principal Component Analysis

4.3.1 Principal Component Analysis with CPI.

Using EViews12 software to run the Principal Components Analysis table analysis program on X1-X19, Y1 data, and generate the principal components analysis Table 6 and Table 7. Because there are many industries involved in this paper, there are more calculation outputs when conducting principal component analysis. The table in this paper retains only the main values of Eigenvalues, Eigenvectors (loadings) and Ordinary correlations listed in the Principal Components Analysis table.

Table 6. X1-X19, Y1 Eigenvalues: ((Sum = 20, Average = 1)

Number	Value	Difference	Proportion	Cumulative Value	Cumulative Proportion
1	7.794914	3.75442	0.3897	7.794914	0.3897
2	4.040494	1.717646	0.202	11.83541	0.5918
3	2.322848	1.004704	0.1161	14.15826	0.7079

Number	Value	Difference	Proportion	Cumulative Value	Cumulative Proportion
4	1.318144	0.231966	0.0659	15.4764	0.7738
5	1.086178	0.255617	0.0543	16.56258	0.8281
6	0.830561	0.199971	0.0415	17.39314	0.8697

Table 7. X1-X19, Y1 Eigenvectors (loadings)

Variable	PC1	PC2	PC3	PC4	PC5	PC6
X1	0.303114	-0.028247	0.191235	0.128219	0.053984	-0.321993
X2	0.306935	-0.121168	0.164286	0.044935	0.078677	-0.264327
X3	0.320558	0.150977	-0.0122	-0.117402	-0.120042	0.096707
X4	0.273846	-0.134591	-0.218965	-0.143159	-0.083662	0.210885
X5	-0.058634	0.403361	-0.322017	-0.036669	-0.019057	0.168714
X6	0.258772	0.247293	-0.174647	-0.054985	0.033528	0.103828
X7	0.082348	0.01172	-0.11498	-0.137163	0.87277	0.010562
X8	0.262889	0.037046	0.237824	0.141613	0.014835	0.474998
X9	0.256801	0.04812	-0.341708	-0.092479	-0.120246	0.191771
X10	0.32679	-0.110061	-0.114083	-0.02857	-0.016734	0.155482
X11	0.019558	0.386498	0.039861	0.377173	0.052616	-0.054081
X12	0.167484	0.351583	-0.162361	-0.136178	-0.018816	-0.270708
X13	-0.001063	0.295512	0.370689	0.192085	0.301518	0.125541
X14	0.240282	0.158027	-0.192101	0.174391	0.053019	-0.199623
X15	0.154813	0.235678	0.359354	-0.380882	-0.03266	-0.011466
X16	-0.096751	0.339983	0.155473	-0.44302	-0.095556	-0.044729
X17	0.290037	-0.193042	0.079872	0.289029	0.008219	0.220677
X18	0.23333	0.080507	0.02056	0.226761	-0.215651	-0.431371
X19	0.199708	-0.108886	0.406045	-0.296061	-0.09272	0.06685
Y1	-0.138514	0.302973	0.165357	0.320309	-0.170581	0.271106

According to Table 6, Table 7 and equation (6), PC1 with Cumulative Proportion of 38.97% was selected, and analysis formula (9) was established for analysis. For PC2, PC3, PC4 and PC5, whose Cumulative Proportion is above 85%, the cumulative proportion was only observed and analyzed due to limited space.

$$
\begin{aligned}
PC1 &= +\ 0.3031\,X1 + 0.3069\,X2\,X3\,X4\ 0.05863 + 0.2739 \\
&+ 0.3206\,X5\,X6\,X7 + 0.2629 + 0.0823 + 0.2588)\ by\ 8\,X9\,X10 \\
&+ 0.3268 + 0.0196 + 0.2568 + 0.1675\,X12 - 0.0\,X11\ 011X13 \\
&+0.2402X14 + 0.1548X15 - 0.0968X16 + 0.2900X17 + 0.2333X18 \\
&+0.1997X19 - 0.1385Y1.
\end{aligned} \tag{9}
$$

From equation (9), we can see that during the period from January 2021 to August 2024, X10, X3, X1 and X2 have the greatest negative impact on Y1 in the war industry.

According to Table 7, from the PC2 values, we can see that X5 has the greatest positive impact on Y1. From the PC3 value, we can see that X19 has the greatest positive effect on Y1. From the PC4 value, we can see that X16 has the greatest negative effect on Y1. From the PC5 value, we can see that X7 has the greatest negative effect on Y1. From the PC6 value, we can see that X8 has the greatest positive effect on Y1.

4.3.2 Principal Component Analysis of PPI.

Use EViews12 software to run the Principal Components Analysis table analysis program on X1-X19 and Y2 data, generating Principal Component Analysis Tables 8 and 9. Due to the wide range of industries involved in this article, there is a significant amount of content to be calculated during principal component analysis. Therefore, the table in this article only lists the main values for Eigenvalues, Eigenvectors (loads), and Ordinary correlations analyzed and calculated in the Principal Components Analysis table.

Table 8. X1-X19, Y2 Eigenvalues: ((Sum = 20, Average = 1)

Number	Value	Difference	Proportion	Cumulative Value	Cumulative Proportion
1	7.938619	3.926745	0.3969	7.938619	0.3969
2	4.011874	1.612529	0.2006	11.95049	0.5975
3	2.399345	1.126811	0.12	14.34984	0.7175
4	1.272534	0.193782	0.0636	15.62237	0.7811
5	1.078752	0.202087	0.0539	16.70112	0.8351
6	0.876666	0.320601	0.0438	17.57779	0.8789

Table 9. X1-X19, Y2 Eigenvectors (loadings)

Variable	PC1	PC2	PC3	PC4	PC5	PC6
X1	0.308801	-0.057249	0.159541	-0.154975	-0.043885	-0.221237
X2	0.312202	-0.140835	0.120736	-0.093092	0.020439	-0.187206
X3	0.312029	0.172134	-0.01391	0.17682	-0.054589	0.06125
X4	0.259742	-0.069175	-0.289295	0.237707	0.03087	0.154769

X5	-0.071669	0.444635	-0.199121	-0.008629	0.036267	0.076319
X6	0.24781	0.281969	-0.118452	0.011267	0.055131	0.08802
X7	0.080939	0.025763	-0.05225	-0.248791	0.8696	0.062285
X8	0.26689	0.00423	0.16654	0.015495	-0.057756	0.476471
X9	0.241503	0.115992	-0.343644	0.140726	0.003132	0.061342
X10	0.317929	-0.069166	-0.176328	0.0864	0.003719	0.160548
X11	0.020571	0.359697	0.144903	-0.433763	-0.240946	0.139148
X12	0.158297	0.376082	-0.049597	0.013729	0.055583	-0.313795
X13	0.008157	0.228136	0.4366	-0.289649	0.027528	0.369356
X14	0.235239	0.174862	-0.166979	-0.202363	0.044082	-0.255017
X15	0.158586	0.194339	0.400273	0.347661	0.113538	-0.065313
X16	-0.098624	0.324681	0.262339	0.373281	0.107818	-0.159786
X17	0.293194	-0.203197	-0.037025	-0.14263	-0.115312	0.26975
X18	0.233532	0.072632	0.015567	-0.168505	-0.321789	-0.336304
X19	0.20785	-0.150779	0.342849	0.350145	0.022348	0.030274
Y2	-0.197868	0.283583	-0.229343	0.21262	-0.155242	0.280042

According to Table 8, Table 9 and equation (6), PC1 with Cumulative Proportion of 39.69% was selected, and analysis formula (10) was established for analysis. For PC2, PC3, PC4 and PC5, whose Cumulative Proportion is above 85%, the cumulative proportion was only observed and analyzed due to limited space.

$$
\begin{aligned}
PC1 &= + \ 0.3088\,X1 \ + \ 0.3122\,X2\,X3\,X4 \ 0.0717 \ + \ 0.2597 \\
&+ 0.3120\,X5\,X6\,X7 \ + \ 0.2669 \ + \ 0.0809 \ + \ 0.2478)\ by\ 8\,X9\,X10 \\
&+ 0.3179 \ + \ 0.0206 \ + \ 0.2415\,X11\,X12 \ + \ 0.1583 \ + \ 0.00\ 82X13 \\
&+0.2352X14+0.1586X15-0.0986X16+0.2932X17+0.2335X18 \\
&+0.2079X19-0.1979Y2
\end{aligned} \tag{10}
$$

From equation (10), we can see that during the period from January 2021 to August 2024, X10, X2 and X3 have the greatest negative impact on Y2 in the war industry.

According to Table 9, from the PC2 values, we can see that X5 has the greatest positive impact on Y2. From the PC3 value, we can see that X13 has the greatest negative effect on Y2. From the PC4 value, we can see that X11 has the greatest negative effect on Y2. From the PC5 value, we can see that X7 has the greatest negative effect on Y2. From the PC6 value, we can see that X8 has the greatest positive effect on Y2.

4.4 Selection of Influencing Factors of Strategic Emerging Industries on CPI and PPI

Based on the Spearman grade correlation coefficient and principal component analysis results, the intersection factors of strategic emerging industries' influence on CPI and PPI (referring to the set composed of elements belonging to both the first set and the second set) and the combination factors (referring to the set composed of all elements belonging to the first set or the second set) are screened in Table 10 and Table 11.

Table 10. Screening of the main influencing factors of strategic emerging industries on CPI

variable	Variable name	A：Spearman correlation coefficient analysis		B：Principal component analysis		A∩B	A∪B
X1	New energy vehicle output	--	--	Negative influence	Principal component	--	√
X2	Solar cell output	Negative correlation	** Significant correlation	Negative influence	Principal component	√	√
X3	Lithium-ion battery production	--	--	Negative influence	Principal component	--	√
X4	Integrated circuit output	Negative correlation	**Significant correlation	--	--	--	√
X5	Computer output	Positive correlation	** Significant correlation	Positive influence	Principal component	√	√
X6	Output of electrical instruments	--	--	--	--	--	--
X7	Output of EMU units	Negative correlation	* Significant correlation	Negative influence	Principal component	√	√
X8	Industrial robot output	--	--	Positive influence	Principal component	--	√
X9	Service robot output	Negative correlation	* Significant correlation	--	--	--	√
X10	Output of optoelectronic devices	Negative correlation	** Significant correlation	Negative influence	Principal component	√	√
X11	Optical cable output	Positive correlation	** Significant correlation	--	--	--	√
X12	Mobile phone	--	--	--	--	--	--
X13	Output of mobile communication base stations	Positive correlation	** Significant correlation			--	√
X14	Smart watch production	--	--	--	--	--	--
X15	The original chemical	--	--	--	--	--	--
X16	Production of proprietary Chinese medicine	Positive correlation	** Significant correlation	Negative influence	Principal component	√	√
X17	Solar power generation	Negative correlation	* Significant correlation	--	--	--	√
X18	Nuclear power generation	--	--	--	--	--	--
X19	Wind power generation	--	--	Positive influence	Principal component	--	√

Table 11. Screening the main influencing factors of strategic emerging industries on PPI

Variable	Variable name	A：Spearman correlation coefficient analysis		B：Principal component analysis		A∩B	A∪B
X1	New energy vehicle output	Negative correlation	** Significant correlation	--	--	--	√

X2	Solar cell output	Negative correlation	** Significant correlation	Negative influence	Principal component	√	√
X3	Lithium-ion battery production	--	--	Negative influence	Principal component	--	√
X4	Integrated circuit output	--	--	--	--	--	√
X5	Computer output	Positive correlation	** Significant correlation	Positive influence	Principal component	√	√
X6	Output of electrical instruments	--	--	--	--	--	--
X7	Output of EMU units	--	--	Negative influence	Principal component	--	√
X8	Industrial robot output	Negative correlation	** Significant correlation	Positive influence	Principal component	√	√
X9	Service robot output	--	--	--	--	--	√
X10	Output of optoelectronic devices	Negative correlation	** Significant correlation	Negative influence	Principal component	√	√
X11	Optical cable output	--	--	Negative influence	Principal component	--	√
X12	Mobile phone	--	--	--	--	--	--
X13	Output of mobile communication base stations	--	--	Negative influence	Principal component	--	√
X14	Smart watch production	--	--	--	--	--	--
X15	The original chemical	--	--	Negative influence	Principal component	--	√
X16	Production of proprietary Chinese medicine	Positive correlation	** Significant correlation	--	--	--	√
X17	Solar power generation	Negative correlation	** Significant correlation	--	--	--	√
X18	Nuclear power generation	Negative correlation	* Significant correlation	--	--	--	√
X19	Wind power generation	Negative correlation	** Significant correlation	--	--	--	√

From Table 10, we can see that during the period from January 2021 to August 2024, the output of electrical instruments X6, mobile communication handsets (mobile phones) X12, smart watches X14, chemical drugs X15, and nuclear power generation X18 in the Zhanxin industry have almost no impact on the change of CPI. The most significant impact on CPI is solar cell production X2, electronic computer output X5, EMU output X7, optoelectronic device output X10, and Chinese patent medicine output X16.

According to Table 11, we can see that from January 2021 to August 2024, the output of integrated circuits X4, electrical instruments X6, service robots X9, mobile communication handsets X12, and smart watches X14 in the war industry have little impact on the change of PPI. The most significant impact on PPI is the output of solar cells X2, the output of electronic computers X5, the output of industrial robots X8, and the output of optoelectronic devices X10.

4.5 Data Stationarity Test and Differential Transformation Analysis

4.5.1 Data Stationarity Test.

In order to facilitate modeling, it is necessary to judge the predictability of the data, so it is necessary to conduct ADF test on the data to judge the stationarity of the data. By using EViews12 software, ADF analysis was conducted on the main factors affecting CPI (X2, X5, X7, X10, X16) and PPI (X2, X5, X8, X10) during January 2021 to August 2024, and the analysis results were shown in FIG. 1 and FIG. 2

Null Hypothesis: Unit root (individual unit root process)
Series: X2, X5, X7, X10, X16, Y1
Date: 10/01/24 Time: 07:45
Sample: 2021M01 2024M08
Exogenous variables: Individual effects
Automatic selection of maximum lags
Automatic lag length selection based on SIC: 0 to 4
Total number of observations: 252
Cross-sections included: 6

Method	Statistic	Prob.**
ADF - Fisher Chi-square	42.0269	0.0000
ADF - Choi Z-stat	-2.52131	0.0058

** Probabilities for Fisher tests are computed using an asymptotic Chi -square distribution. All other tests assume asymptotic normality.

Intermediate ADF test results UNTITLED

Series	Prob.	Lag	Max Lag	Obs
X2	0.7898	2	9	41
X5	0.6145	4	9	39
X7	0.0000	0	9	43
X10	0.3956	0	9	43
X16	0.0774	0	9	43
Y1	0.3187	0	9	43

Fig. 1. ADF analysis was the main reason for the intersection of CPI

Null Hypothesis: Unit root (individual unit root process)
Series: X2, X5, X8, X10, Y2
Date: 10/01/24 Time: 07:46
Sample: 2021M01 2024M08
Exogenous variables: Individual effects
Automatic selection of maximum lags
Automatic lag length selection based on SIC: 0 to 4
Total number of observations: 205
Cross-sections included: 5

Method	Statistic	Prob.**
ADF - Fisher Chi-square	6.09034	0.8076
ADF - Choi Z-stat	0.40046	0.6556

** Probabilities for Fisher tests are computed using an asymptotic Chi -square distribution. All other tests assume asymptotic normality.

Intermediate ADF test results UNTITLED

Series	Prob.	Lag	Max Lag	Obs
X2	0.7898	2	9	41
X5	0.6145	4	9	39
X8	0.3922	3	9	40
X10	0.3956	0	9	43
Y2	0.6319	1	9	42

Fig. 2. ADF analysis was the main reason for the intersection of PPI

As can be seen from Figure 1, the main factors affecting the intersection of CPI and the results of ADF-Fisher Chi-square and ADF-Choi Z-stat show that the P-values are both very small, 0 and 0.0058 respectively, which indicates that the null hypothesis is rejected, that is, the sequence is likely to be stable in general and has good predictability.

As can be seen from Figure 2, the main factors affecting the intersection of PPI, from the results of ADF-Fisher Chi-square and ADF-Choi Z-stat, showed that the P values were both large, 0.8076 and 0.6556 respectively, far greater than 0.05, indicating that the null hypothesis was not rejected, that is, the sequence as a whole is likely to be unstable. Sequences are less predictive and need to be transformed differentially or logarithmically to make them stationary before further analysis and modeling.

4.5.2 Logarithmic Conversion and Judgment.

Since the ADF test of the main factors affecting PPI (Y2) intersection (X2, X5, X8, X10) data is not stable during January 2021 to August 2024, differential conversion is required. Logarithmic conversion of X2, X5, X8 and X10 (genr lny=log(y)) was performed using EViews12 software. The logarithmic conversion data corresponding to X2, X5, X8 and X10 were lnX2, lnX5, lnX8 and lnX10 were ADF tested again on the converted data, and the output results were shown in Figure 3.

Null Hypothesis: Unit root (individual unit root process)
Series: LNX2, LNX10, LNX8, LNX5, Y2
Date: 10/01/24 Time: 21:40
Sample: 2021M01 2024M08
Exogenous variables: Individual effects
Automatic selection of maximum lags
Automatic lag length selection based on SIC: 0 to 3
Total number of observations: 206
Cross-sections included: 5

Method	Statistic	Prob.**
ADF - Fisher Chi-square	8.95072	0.5368
ADF - Choi Z-stat	-0.33137	0.3702

** Probabilities for Fisher tests are computed using an asymptotic Chi -square distribution. All other tests assume asymptotic normality.

Intermediate ADF test results UNTITLED

Series	Prob.	Lag	Max Lag	Obs
LNX2	0.6885	2	9	41
LNX10	0.2234	0	9	43
LNX8	0.3711	3	9	40
LNX5	0.3157	3	9	40
Y2	0.6319	1	9	42

Fig. 3. The intersection of main factors affecting PPI was analyzed by ADF after first-order difference.

As shown in Figure 3, the test results of ADF-Fisher Chi-square and ADF-Choi Z-stat showed that the P-value decreased and the predictability of the sequence improved.

4.6 Strategic New Industries and CPI and PPI Forecasting Modeling

CPI and PPI forecasts are important for understanding economic conditions and guiding policy making and business decisions. From January 2021 to August 2024, this paper finds that among the 19 industrial categories of strategic emerging industries, except the output of electrical instruments and mobile communication handsets, which have almost no impact on the changes of CPI and PPI, the other 17 industries all have different degrees of correlation, positive or negative impact. Therefore, we can establish the relevant forecasting model of Zhanxin industry output, CPI and PPI, and carry out the relevant statistical forecasting work.

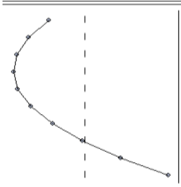
4.6.1 Forecasting Model with CPI.

By using EViews12 software, solar cell output X2, electronic computer output X5, EMU output X7, optoelectronic device output X10, and Chinese patent medicine output X16, which have the most significant impact on consumer price index Y1, were selected for ARIMA modeling, and the maximum lag coefficient (Max lag) was selected as 9 months. Run LS Y1 C PDL(X2,9,2,0), LS Y1 C PDL(X5,9,2,0), LS Y1 C PDL(X7,9,2,0), LS Y1 C PDL(X10,9,2,0), LS Y1 C PDL(X16,9,2,0) "command is modeled successively, where 9 is the lag length,2 is the polynomial frequency,0 is the parameter that controls the distribution lag characteristics (without any restriction on the parameter distribution), and FIG. 4, FIG. 5, FIG. 6, FIG. 7, FIG. 8 can be output respectively.

Dependent Variable: Y1
Method: Least Squares
Date: 10/01/24 Time: 12:47
Sample (adjusted): 2021M10 2024M08
Included observations: 35 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	103.6362	0.341087	303.8406	0.0000
PDL01	-0.000187	7.47E-05	-2.507875	0.0176
PDL02	2.54E-05	1.76E-05	1.445096	0.1585
PDL03	1.18E-05	9.24E-06	1.281714	0.2094

R-squared	0.699200	Mean dependent var	100.9600
Adjusted R-squared	0.670090	S.D. dependent var	1.035602
S.E. of regression	0.594827	Akaike info criterion	1.906117
Sum squared resid	10.96838	Schwarz criterion	2.083872
Log likelihood	-29.35706	Hannan-Quinn criter.	1.967478
F-statistic	24.01946	Durbin-Watson stat	0.739720
Prob(F-statistic)	0.000000		



Lag Distribution o...	i	Coefficient	Std. Error	t-Statistic
	0	-9.9E-05	0.00013	-0.77094
	1	-0.00016	6.6E-05	-2.39269
	2	-0.00019	4.5E-05	-4.22033
	3	-0.00020	6.2E-05	-3.25767
	4	-0.00019	7.5E-05	-2.50788
	5	-0.00015	7.4E-05	-2.03825
	6	-8.9E-05	5.9E-05	-1.52283
	7	-4.6E-06	4.2E-05	-0.11000
	8	0.00010	6.8E-05	1.53426
	9	0.00024	0.00013	1.74991
Sum of Lags		-0.00074	9.1E-05	-8.12846

Fig. 4. And solar cell yield model

Dependent Variable: Y1
Method: Least Squares
Date: 10/01/24 Time: 12:49
Sample (adjusted): 2021M10 2024M08
Included observations: 35 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	94.22627	0.892084	105.6249	0.0000
PDL01	7.87E-05	7.24E-05	1.086883	0.2855
PDL02	3.67E-06	2.31E-05	0.158665	0.8750
PDL03	1.35E-05	9.09E-06	1.484393	0.1478

R-squared	0.651107	Mean dependent var	100.9600
Adjusted R-squared	0.617343	S.D. dependent var	1.035602
S.E. of regression	0.640615	Akaike info criterion	2.054435
Sum squared resid	12.72203	Schwarz criterion	2.232189
Log likelihood	-31.95262	Hannan-Quinn criter.	2.115796
F-statistic	19.28417	Durbin-Watson stat	0.785718
Prob(F-statistic)	0.000000		

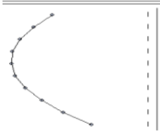
Lag Distribution o...	i	Coefficient	Std. Error	t-Statistic
	0	0.00028	0.00015	1.85917
	1	0.00019	8.8E-05	2.15896
	2	0.00013	5.8E-05	2.14334
	3	8.8E-05	6.3E-05	1.41081
	4	7.9E-05	7.2E-05	1.08688
	5	9.6E-05	7.3E-05	1.31229
	6	0.00014	6.5E-05	2.16024
	7	0.00021	6.2E-05	3.42382
	8	0.00031	9.0E-05	3.43785
	9	0.00043	0.00015	2.86410
	Sum of Lags	0.00195	0.00026	7.59920

Fig. 5. CPI and computer output model

Dependent Variable: Y1
Method: Least Squares
Date: 10/01/24 Time: 12:52
Sample (adjusted): 2021M10 2024M08
Included observations: 35 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	102.9897	0.356863	288.5973	0.0000
PDL01	-0.003166	0.000722	-4.384795	0.0001
PDL02	3.17E-05	0.000172	0.184588	0.8548
PDL03	6.82E-05	6.71E-05	1.015870	0.3176

R-squared	0.548322	Mean dependent var	100.9600
Adjusted R-squared	0.504611	S.D. dependent var	1.035602
S.E. of regression	0.728896	Akaike info criterion	2.312641
Sum squared resid	16.46999	Schwarz criterion	2.490395
Log likelihood	-36.47121	Hannan-Quinn criter.	2.374001
F-statistic	12.54431	Durbin-Watson stat	0.683017
Prob(F-statistic)	0.000015		



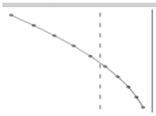
Lag Distribution o...	i	Coefficient	Std. Error	t-Statistic
	0	-0.00220	0.00110	-2.00732
	1	-0.00265	0.00073	-3.65013
	2	-0.00296	0.00061	-4.80956
	3	-0.00313	0.00067	-4.69162
	4	-0.00317	0.00072	-4.38479
	5	-0.00307	0.00071	-4.29614
	6	-0.00283	0.00065	-4.37009
	7	-0.00246	0.00061	-4.04156
	8	-0.00195	0.00077	-2.54340
	9	-0.00130	0.00119	-1.09931
	Sum of Lags	-0.02571	0.00423	-6.08220

Fig. 6. CPI and EMU output model

Dependent Variable: Y1
Method: Least Squares
Date: 10/01/24 Time: 12:54
Sample (adjusted): 2021M10 2024M08
Included observations: 35 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	103.0145	0.813499	126.6314	0.0000
PDL01	-0.000143	0.000262	-0.545112	0.5896
PDL02	0.000250	6.38E-05	3.919757	0.0005
PDL03	-1.70E-05	3.22E-05	-0.529368	0.6003

R-squared	0.547272	Mean dependent var	100.9600
Adjusted R-squared	0.503459	S.D. dependent var	1.035602
S.E. of regression	0.729743	Akaike info criterion	2.314963
Sum squared resid	16.50828	Schwarz criterion	2.492717
Log likelihood	-36.51184	Hannan-Quinn criter.	2.376323
F-statistic	12.49125	Durbin-Watson stat	0.585790
Prob(F-statistic)	0.000016		



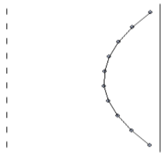
Lag Distribution o...	i	Coefficient	Std. Error	t-Statistic
	0	-0.00142	0.00044	-3.22244
	1	-0.00105	0.00021	-4.86633
	2	-0.00071	0.00015	-4.89711
	3	-0.00041	0.00021	-1.94065
	4	-0.00014	0.00026	-0.54511
	5	9.0E-05	0.00027	0.33781
	6	0.00029	0.00023	1.26865
	7	0.00045	0.00020	2.30423
	8	0.00058	0.00029	2.02684
	9	0.00068	0.00051	1.33203
	Sum of Lags	-0.00163	0.00076	-2.13825

Fig. 7. CPI and output model of optoelectronic devices

Dependent Variable: Y1
Method: Least Squares
Date: 10/01/24 Time: 12:56
Sample (adjusted): 2021M10 2024M08
Included observations: 35 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	92.57620	1.850619	50.02445	0.0000
PDL01	0.036108	0.015830	2.280972	0.0296
PDL02	-0.000862	0.003885	-0.221784	0.8259
PDL03	0.000838	0.001789	0.468356	0.6428

R-squared	0.408933	Mean dependent var	100.9600
Adjusted R-squared	0.351733	S.D. dependent var	1.035602
S.E. of regression	0.833815	Akaike info criterion	2.581599
Sum squared resid	21.55265	Schwarz criterion	2.759353
Log likelihood	-41.17799	Hannan-Quinn criter.	2.642960
F-statistic	7.149185	Durbin-Watson stat	0.480004
Prob(F-statistic)	0.000875		



Lag Distribution o...	i	Coefficient	Std. Error	t-Statistic
	0	0.05296	0.02789	1.89853
	1	0.04623	0.01625	2.84550
	2	0.04118	0.01195	3.44612
	3	0.03781	0.01372	2.75587
	4	0.03611	0.01583	2.28097
	5	0.03608	0.01604	2.25018
	6	0.03774	0.01456	2.59255
	7	0.04106	0.01398	2.93731
	8	0.04606	0.01911	2.40991
	9	0.05274	0.03092	1.70583
	Sum of Lags	0.42797	0.09405	4.55053

Fig. 8. CPI and Chinese patent medicine yield model

According to formula (8), according to figures 4, 5, 6, 7 and 8, CPI and solar cell output X2, electronic computer output X5, EMU output X7, optoelectronic device output X10, and Chinese patent medicine output X16 can be predicted into equation (11), equation (12), equation (13), equation (14) and equation (15).

$$\begin{aligned}Y1-2_t &= 103.6362-0.000099X2_{t-}-0.000160X2_{t-1}-0.000190X2_{t-2}\\ &-0.000200X2_{t-2}-0.000190X2_{t-4}-0.000150X2_{t-5}-0.000089X2_{t-6}\\ &-0.000005X2_{t-7}+0.000100X2_{t-8}+0.000240X2_{t-9}\end{aligned} \tag{11}$$

$$\begin{aligned}Y1-5_t &= 94.22627+-0.77094X5_t-2.39269X5_{t-1}-4.22033X5_{t-2}\\ &-3.25767X5_{t-3}-2.50788X5_{t-4}-2.03825X5_{t-5}-1.52283X5_{t-6}\\ &-0.11X5_{t-7}+1.53426X5_{t-8}+1.74991X5_{t-9}\end{aligned} \tag{12}$$

$$\begin{aligned}Y1-7_t &= 102.9897-0.0022X7_t-0.002657_{t-1}-0.00296X7_{t-2}\\ &-0.00313X7_{t-3}-0.00317X7_{t-4}-0.00307X7_{t-5}-0.00283X7_{t-6}\\ &-0.00246X7_{t-7}-0.00195X7_{t-8}-0.0013X7_{t-9}\end{aligned} \tag{13}$$

$$\begin{aligned}Y1-10_t &= 103.0145-0.00142X10_t-0.0010510_{t-1}-0.00071X10_{t-2}\\ &-0.00041X10_{t-3}-0.00014X10_{t-4}0.00009X10_{t-5}0.00029X10_{t-6}\\ &+0.00045X10_{t-7}+0.00058X10_{t-8}+0.00068X10_{t-9}\end{aligned} \tag{14}$$

$$
\begin{aligned}
Y1-16_t &= 92.5762+0.05296X16_t+0.04623X16_{t-1}+0.04118X16_{t-2}\\
&+0.03781X16_{t-3}+0.03611X16_{t-4}+0.03608X16_{t-5}+0.03774X16_{t-6}\\
&+0.04106X16_{t-7}+0.04606X16_{t-8}+0.05274X16_{t-9}
\end{aligned}
\tag{15}
$$

According to formula (8), the polynomial equation is synthesized by means of average value to improve the prediction accuracy of the model.

$$
Y1_t=(Y1-2_t+Y1-5_t+Y1-7_t+Y1-10_t+Y1-16_t)/5 \tag{16}
$$

4.6.2 Predictive Models with PPI.

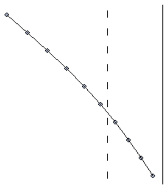
Using EViews12 software, the first-order difference data data of solar cell output X2, electronic computer output X5, industrial robot output X8 and photoelectronic device output X10, which have the most significant impact on the ex-factory price index Y2, are selected for ARIMA modeling, and the lag coefficients are selected as the maximum lag coefficient (Max lag) of 9 months. Run the commands "LS Y2 C PDL(lnX2,9,2,0)", "LS Y2 C PDL(lnX5,9,2,0)", "LS Y2 C PDL(lnX8,9,2,0)" and "LS Y2 C PDL(lnX10,9,2,0)" to model successively. 9 is the lag length, 2 is the polynomial degree, 0 is the parameter that controls the distribution lag characteristics (without any restriction on the parameter distribution), and FIG. 9, FIG. 10, FIG. 11 and FIG. 12 can be output respectively.

Dependent Variable: Y2
Method: Least Squares
Date: 10/01/24 Time: 21:56
Sample (adjusted): 2021M10 2024M08
Included observations: 35 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	219.2045	9.457824	23.17706	0.0000
PDL01	-1.645938	1.109366	-1.483675	0.1480
PDL02	1.193616	0.263049	4.537615	0.0001
PDL03	-0.045975	0.136731	-0.336247	0.7390

R-squared	0.834737	Mean dependent var	101.0200
Adjusted R-squared	0.818744	S.D. dependent var	5.371920
S.E. of regression	2.287052	Akaike info criterion	4.599615
Sum squared resid	162.1488	Schwarz criterion	4.777369
Log likelihood	-76.49326	Hannan-Quinn criter.	4.660976
F-statistic	52.19326	Durbin-Watson stat	0.305585
Prob(F-statistic)	0.000000		

Lag Distribution o...



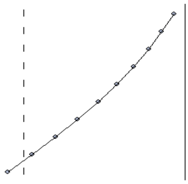
i	Coefficient	Std. Error	t-Statistic
0	-7.15601	1.94111	-3.68655
1	-5.64056	0.99566	-5.66517
2	-4.21707	0.67830	-6.21711
3	-2.88553	0.91449	-3.15536
4	-1.64594	1.10937	-1.48367
5	-0.49830	1.09672	-0.45435
6	0.55739	0.87219	0.63907
7	1.52113	0.60113	2.53046
8	2.39292	0.95405	2.50816
9	3.17276	1.93947	1.63589
Sum of Lags	-14.3992	1.15490	-12.4680

Fig. 9. PPI and solar cell yield (logarithmic transformation) model

Dependent Variable: Y2
Method: Least Squares
Date: 10/01/24 Time: 21:58
Sample (adjusted): 2021M10 2024M08
Included observations: 35 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-173.6258	31.83053	-5.454692	0.0000
PDL01	4.101029	1.036838	3.955323	0.0004
PDL02	-0.776135	0.327395	-2.370637	0.0242
PDL03	-0.038159	0.129747	-0.294103	0.7706

R-squared	0.746980	Mean dependent var	101.0200
Adjusted R-squared	0.722494	S.D. dependent var	5.371920
S.E. of regression	2.829864	Akaike info criterion	5.025545
Sum squared resid	248.2520	Schwarz criterion	5.203299
Log likelihood	-83.94704	Hannan-Quinn criter.	5.086906
F-statistic	30.50666	Durbin-Watson stat	0.266170
Prob(F-statistic)	0.000000		



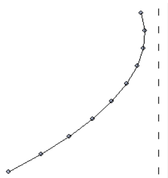
Lag Distribution o...	i	Coefficient	Std. Error	t-Statistic
	0	6.59503	2.11543	3.11758
	1	6.08600	1.24168	4.90144
	2	5.50066	0.85139	6.46078
	3	4.83901	0.91217	5.30493
	4	4.10103	1.03684	3.95532
	5	3.28673	1.03418	3.17811
	6	2.39612	0.91394	2.62175
	7	1.42919	0.89224	1.60181
	8	0.38594	1.33407	0.28930
	9	-0.73362	2.24010	-0.32750
	Sum of Lags	33.8861	3.91561	8.65411

Fig. 10. PPI and computer output (logarithmic transformation) model

Dependent Variable: Y2
Method: Least Squares
Date: 10/01/24 Time: 22:00
Sample (adjusted): 2021M10 2024M08
Included observations: 35 after adjustments

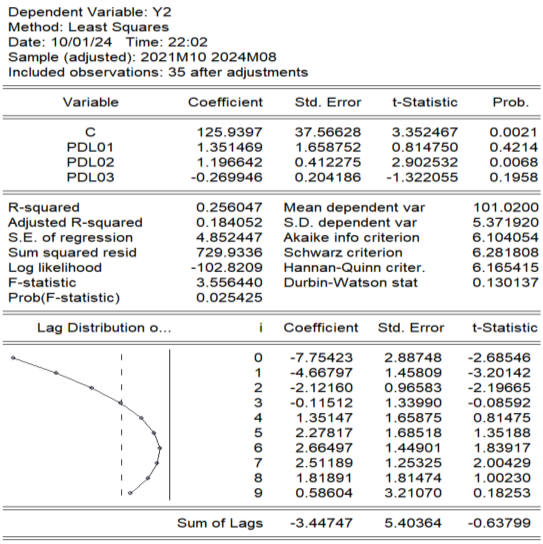
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	602.5649	47.72654	12.62536	0.0000
PDL01	-2.630182	1.045596	-2.515485	0.0173
PDL02	-1.069480	0.264917	-4.037034	0.0003
PDL03	-0.190030	0.113491	-1.674407	0.1041

R-squared	0.811799	Mean dependent var	101.0200
Adjusted R-squared	0.793586	S.D. dependent var	5.371920
S.E. of regression	2.440613	Akaike info criterion	4.729586
Sum squared resid	184.6543	Schwarz criterion	4.907340
Log likelihood	-78.76776	Hannan-Quinn criter.	4.790947
F-statistic	44.57256	Durbin-Watson stat	0.522058
Prob(F-statistic)	0.000000		



Lag Distribution o...	i	Coefficient	Std. Error	t-Statistic
	0	-1.39275	1.74091	-0.80001
	1	-1.13201	0.99655	-1.13593
	2	-1.25134	0.74850	-1.67180
	3	-1.75073	0.89499	-1.95615
	4	-2.63018	1.04560	-2.51548
	5	-3.88969	1.06024	-3.66869
	6	-5.52926	0.94323	-5.86204
	7	-7.54889	0.83385	-9.05310
	8	-9.94859	1.07488	-9.25552
	9	-12.7283	1.78702	-7.12266
	Sum of Lags	-47.8018	4.53969	-10.5297

Fig. 11. Modulus with industrial robot output (logarithmic transformation)



Dependent Variable: Y2
Method: Least Squares
Date: 10/01/24 Time: 22:02
Sample (adjusted): 2021M10 2024M08
Included observations: 35 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	125.9397	37.56628	3.352467	0.0021
PDL01	1.351469	1.658752	0.814750	0.4214
PDL02	1.196642	0.412275	2.902532	0.0068
PDL03	-0.269946	0.204186	-1.322055	0.1958

R-squared	0.256047	Mean dependent var	101.0200
Adjusted R-squared	0.184052	S.D. dependent var	5.371920
S.E. of regression	4.852447	Akaike info criterion	6.104054
Sum squared resid	729.9336	Schwarz criterion	6.281808
Log likelihood	-102.8209	Hannan-Quinn criter.	6.165415
F-statistic	3.556440	Durbin-Watson stat	0.130137
Prob(F-statistic)	0.025425		

Lag Distribution o...	i	Coefficient	Std. Error	t-Statistic
	0	-7.75423	2.88748	-2.68546
	1	-4.66797	1.45809	-3.20142
	2	-2.12160	0.96583	-2.19665
	3	-0.11512	1.33990	-0.08592
	4	1.35147	1.65875	0.81475
	5	2.27817	1.68518	1.35188
	6	2.66497	1.44901	1.83917
	7	2.51189	1.25325	2.00429
	8	1.81891	1.81474	1.00230
	9	0.58604	3.21070	0.18253
	Sum of Lags	-3.44747	5.40364	-0.63799

Fig. 12. PPI and optoelectronic device yield (logarithmic transformation) model

According to formula (8), prediction equations (17), (18), (19), (20) and (21) are formed according to the first-order difference data of PPI and solar cell output X2, electronic computer output X5, industrial robot output X8, and photoelectronic device output X10 in FIG. 9, 10, 11, and 12.

$$\begin{aligned} Y2-2_t &= 219.2045 - 7.15601CX2_t - 5.64056CX2_{t-1} - 4.21707CX2_{t-2} \\ &-2.88553CX2_{t-3} - 1.64594CX2_{t-4} - 0.49830CX2_{t-5} + 0.55739CX2_{t-6} \\ &+1.52113CX2_{t-7} + 2.39292CX2_{t-8} + 3.17276CX2_{t-9} \end{aligned} \tag{17}$$

$$\begin{aligned} Y2-5_t &= -173.6258 + 6.59503CX5_t + 6.08600CX5_{t-1} + 5.50066CX5_{t-2} \\ &+4.83901CX5_{t-3} + 4.10103CX5_{t-4} + 3.28673CX5_{t-5} + 2.39612CX5_{t-6} \\ &+1.42919CX5_{t-7} + 0.38594CX5_{t-8} - 0.73362CX5_{t-9} \end{aligned} \tag{18}$$

$$\begin{aligned} Y2-8_t &= 602.5649 - 1.39275CX8_t - 1.13201CX8_{t-1} - 1.25134CX8_{t-2} \\ &-1.75073CX8_{t-3} - 2.63018CX8_{t-4} - 3.88969CX8_{t-5} - 5.52926CX8_{t-6} \\ &-7.54889CX8_{t-7} - 9.94859CX8_{t-8} - 12.7283CX8_{t-9} \end{aligned} \tag{19}$$

$$\begin{aligned} Y2-10_t &= 125.9397 - 7.75423CX10_t - 4.66797CX10_{t-1} - 2.12160CX10_{t-2} \\ &-0.11512CX10_{t-3} + 1.35147CX10_{t-4} + 2.27817CX10_{t-5} + 2.66497CX10_{t-6} \\ &+2.51189CX10_{t-7} + 1.81891CX10_{t-8} + 0.58604CX10t - 9 \end{aligned} \tag{20}$$

According to formula (8), the polynomial equation is synthesized by means of average value to improve the prediction accuracy of the model.

$$Y2t = (Y2-2_t + Y2-5_t + Y2-8_t + Y2-10_t)/4 \tag{21}$$

4.6.3 Model Accuracy Check.

To further verify the accuracy of the model, this paper selects the latest data published by the National Development and Reform Commission, the Ministry of Industry and Information Technology, the National Bureau of Statistics and the China Automobile Association from April to August 2024, and calculates the actual relative errors of the published and calculated values according to formula (22).

$$\Delta_t = (Yi_t - Ci_t) / Ci_t * 100 \tag{22}$$

In the formula, Δt is the error between the actual value and the calculated value of the model, where Yit is the calculated value of the pre-model, Cit is the actual value and t is the actual month.

The formulas (11), (12), (13), (14), (15) and (16) of the CPI model are verified, and the data are shown in Table 12. PPI model formulas (17), (18), (19), (20) and (21) were verified, and the data are shown in Table 13.

Table 12. Accuracy of CPI actual value and model value from April to August 2024 (n=4)

Category	August 2024		July 2024		June 2024		May 2024	
	C1	Δt	C1	Δt	C1	Δt	C1	Δt
Actual value	100.6	——	100.5	——	100.2	——	100.3	——
Formula 11	100.26	-0.34%	100.39	-0.11%	100.46	0.26%	100.42	0.12%
Formula 12	100.03	-0.57%	99.98	-0.52%	100.18	-0.02%	100.16	-0.14%
Formula 13	100.17	-0.43%	100.24	-0.26%	100.06	-0.14%	100.74	0.44%
Formula 14	99.92	-0.67%	99.80	-0.70%	99.69	-0.51%	99.95	-0.35%
Formula 15	99.71	-0.89%	99.94	-0.56%	100.12	-0.08%	99.98	-0.32%
Formula 16	100.02	-0.58%	100.07	-0.43%	100.10	-0.10%	100.25	-0.05%

Table 13. Accuracy of actual and model PPI values from April to August 2024 (n=4)

Category	May 2024		June 2024		July 2024		August 2024	
	C1	Δt	C1	Δt	C1	Δt	C1	Δt
Actual value	98.20	——	99.2	——	99.2	——	98.6	——
Formula 17	97.38	-0.83%	97.59	-1.62%	97.78	-1.44%	98.28	-0.32%
Formula 18	96.86	-1.36%	95.93	-3.29%	95.65	-3.58%	95.13	-3.52%
Formula 19	94.96	-3.30%	97.87	-1.34%	98.78	-0.42%	100.62	2.05%
Formula 20	99.20	1.02%	98.47	-0.73%	97.80	-1.41%	98.28	-0.32%
Formula 21	97.10	-1.12%	97.47	-1.75%	97.50	-1.71%	98.08	-0.53%

As can be seen from Table 12 and Table 13, in the CPI model, formulas (11), (12), (13), (14), (15) and (16) all have good prediction accuracy, and the fitting formula (16) has the highest accuracy. In the PPI model, the prediction accuracy of formula (17), formula (20) and fitting formula (21) is relatively high, and the overall forecast trend is consistent with the actual value. This shows that a relationship model can be established between CPI and solar cell output X2, computer output X5, EMU output X7, photoelectronic device output X10, Chinese patent medicine output X16, PPI and solar cell output X2, computer output X5, industrial robot output X8, and photoelectronic device output X10. It also has high model fitting accuracy. You can use the data of the current month and the previous month to predict the data of this month, or you can use the model formula, and use the time series prediction function provided by EViews, SPSS and other software to predict the data of the next few months.

5 Research Conclusions, Suggestions and Prospects

5.1 Research Conclusions

This paper adopts trend analysis, principal component analysis, Pearson correlation coefficient analysis and other methods to study the relationship between 17 industrial data of strategic emerging industries and 21 data of energy industry from January 2021 to June 2024, and finds the following six conclusions:

(1) In recent years, strategic emerging industries have developed rapidly, but some industries have lagged behind. Trend and growth rate analysis shows that from January 2021 to June 2024, new energy vehicles, solar cells, industrial robots, optoelectronic devices, solar power generation growth rates are more than 100%. The growth rate of lithium-ion batteries, integrated circuits, electrical instruments, chemical agents, and wind power generation also exceeded 50%, but the production of electronic computer complete machines and proprietary Chinese medicines developed slowly and lagged behind.

(2) In recent years, the overall growth rate of the development of the energy industry has slowed down, and some industries have long-term low growth. Trend and growth rate analysis shows that between January 2021 and June 2024, the highest growth rate in the energy sector is the primary industry electricity consumption and liquefied natural gas production, with a growth rate of more than 100%. The output of naphtha and coal bed gas increased by more than 50%. Other such as tertiary industry electricity consumption, urban and rural residential electricity consumption, liquefied petroleum gas production, fuel oil production, diesel oil production, kerosene production, raw coal production, crude oil processing volume, petroleum coke production, coke production, the growth rate is low, petroleum asphalt production growth trend is negative.

(3) New energy vehicles, lithium-ion batteries, electronic computer machines, optical cables, mobile communication base station equipment, the output of proprietary Chinese medicine, electrical instruments and meters, and nuclear power generation have the greatest impact on the power industry. Principal component analysis and correlation coefficient analysis show that from January 2021 to June 2024, the output of new energy vehicles has the largest positive impact on the electricity consumption of the secondary industry. The output of electronic computer has the greatest negative influence on the domestic electricity consumption of urban and rural residents. The output of optical cable, the output of mobile communication base station equipment and the output of proprietary Chinese medicine have the greatest positive influence on the domestic electricity consumption of urban and rural residents. Lithium-ion batteries,

electrical instrumentation, and nuclear power generation are highly correlated with the consumption of electricity in the secondary industry.

(4) New energy vehicles, solar cells, lithium-ion batteries, electronic computer machines, optical cables, chemical agents, nuclear power generation, wind power generation, solar power generation, industrial robots, optoelectronic devices have the greatest impact on the gas industry. Principal component and correlation coefficient analyses show that solar cell production has the largest positive impact on LNG production between January 2021 and June 2024. Computer and nuclear power generation have the greatest negative effect on gas production. The production of chemical agents has the most positive effect on the production of natural gas. The output of optical cable has the most positive effect on the output of gas. Wind power generation is highly correlated with natural gas production, solar power generation is highly correlated with liquefied petroleum gas production, new energy vehicle production, solar cell production, lithium-ion battery production, industrial robot production, optoelectronic device production, solar power generation is highly correlated with LNG production.

(5) In the new industries, solar cells, optical cables, service robots, industrial robots, solar power generation, new energy vehicles, and solar cells have the greatest impact on the oil industry. Principal component and correlation coefficient analysis showed that solar cells had the largest positive impact on naphtha between January 2021 and June 2024. The output of optical fiber cable has the greatest negative effect on the output of kerosene. The output of service robot has the most positive effect on the output of petroleum asphalt. The output of industrial robots has the greatest negative effect on the output of petroleum coke. The output of optical cable has the most positive effect on the output of fuel oil.

(6) Solar cells, electronic computer machines, service robots, Chinese patent medicines, industrial robots, new energy vehicles, lithium-ion batteries, optoelectronic devices, solar power generation, wind power generation have the greatest impact on the coal industry. Principal component and correlation coefficient analyses show that solar cell production has the largest positive impact on CBM production between January 2021 and June 2024. The output of electronic computer has the greatest negative effect on coke output. The output of service robot has the most negative effect on the raw coal output. The output of proprietary Chinese medicine has the greatest negative effect on the coke output. The output of industrial robots has the greatest negative effect on the output of raw coal.

5.2 Research Proposal

Based on the conclusions of this study, combined with the relevant contents of literature research, data research and policy research in the research process, there are five research suggestions as follows:

(1) The development of strategic emerging industries should focus on strengthening the investment and guarantee of the secondary industry and residential electricity consumption. Strategic emerging industries such as new energy vehicles and lithium-ion batteries have become the main factors driving the growth of the power industry. At present, the competent authorities of the power industry and relevant enterprises should

focus on strengthening the power guarantee of new energy vehicles, lithium-ion batteries, electronic computer complete machines, optical cables, mobile communication base station equipment, Chinese patent medicine production, electrical instruments and meters industry, and make plans in advance to ensure the power supply. At the same time, it is necessary to moderately accelerate the investment in nuclear power generation, accelerate the development of safe and reliable nuclear power generation units, and promote and guarantee the development of new industries.

(2) The development of strategic emerging industries should comprehensively strengthen the supply and price guarantee of liquefied natural gas and natural gas. Production of new energy vehicles, solar cell production, lithium-ion battery production, industrial robot production, optoelectronic device production, and solar power generation are highly correlated with liquefied petroleum gas and natural gas. At present, the government and industry departments should make great efforts to organize the supply guarantee and stable price of natural gas and liquefied petroleum gas, and provide good support for the development of the new industry.

(3) The development of strategic emerging industries should pay attention to the development of non-fuel products such as naphtha, petroleum asphalt and petroleum coke. The influence of traditional gasoline, kerosene and other oil products on strategic emerging industries is rapidly declining, and other non-fuel products that were originally by-products are beginning to be used in strategic emerging industries. The relevant competent departments and enterprises of the petroleum industry should pay attention to the development of strategic emerging industries in a timely manner, take the initiative to respond to changes, increase the petroleum products that meet the production needs of the new industry, and promote the coordinated development of the new industry and the petroleum industry.

(4) The development of strategic emerging industries should also pay attention to the security and development of China's traditional energy sources such as raw coal and coal-bed methane. From the research point of view, in the coal industry, the production of coking coal is not closely related to the war industry, but the production of new energy vehicles, solar cell production, lithium-ion battery production and raw coal production are also highly correlated. The output of new energy vehicles, solar cell output, optoelectronic device output, solar power generation, wind power generation is highly correlated with the output of coal bed methane, which shows that raw coal and coal bed methane, as China's traditional main energy sources, are still an important foundation to support the development of the new industry, and should continue to be retained, continue to promote their clean application, and promote the development of the new industry.

(5) To develop strategic emerging industries, we should promote the green and low-carbon transformation of energy and enhance the security guarantee ability of industrial chain and supply chain. In recent years, strategic emerging industries have developed rapidly. From the research point of view, the development of strategic emerging industries in the current open data has a high correlation with most products in the energy industry, and most of the data have a positive growth relationship, which means that the development of strategic emerging industries needs the support of the energy industry. Similarly, the development of solar energy, wind power and nuclear energy, which

belong to the new industry in the energy industry, also plays a greater role in promoting the transformation of the energy industry. At present, we should accelerate the green and low-carbon transformation of the energy industry and accelerate the construction of a new energy system that ADAPTS to the development of strategic emerging industries.

(6) The development of strategic emerging industries should pay attention to some lagging industries. For example, the development of the electronic computer industry and the Chinese patent medicine industry lags behind, should attract the attention of the competent government departments, formulate special measures to encourage development, and accelerate development. The low growth rate of tertiary industry electricity consumption and household electricity consumption of urban and rural residents reflects some hidden problems between the current service industry and residents' consumption power and the overall economic growth. Special policies should be formulated to promote the development of service industry and promote the growth of residents' consumption and income.

5.3 Research Prospect

In the context of the rapid development of the global economy and the increasingly severe environmental problems, strategic emerging industries are having a profound and significant impact on the energy industry with their strong innovation vitality and development potential. The development of strategic emerging industries will bring new opportunities and challenges to the energy industry. This paper conducts empirical analysis and finds some characteristics of trends, principal components and correlations between strategic emerging industries and energy industries. However, it is also found that at present, due to the rapid development of the new industry, some industry data can not be collected in time, and the relevant statistical data of the energy industry is too traditional and rough. This also precisely shows that there is an urgent need to improve the statistical index accounting system supporting high-quality development in accordance with the deployment of the Third Plenary Session of the 20th Central Committee, and strengthen the coverage of the new economy and new fields. Only in this way can the accuracy of research be further improved. In addition, in the research process, the author spent a lot of time processing data with large sample size. In the future, with the development of war industry and data acquisition technology, such research should develop towards artificial intelligence statistical analysis technology, using machine learning algorithms to automatically identify and extract key features in data, learn from new data and automatically adjust the model. In order to adapt to the change of data distribution, improve the accuracy of analysis, reduce the deviation and variance of the model, and improve the efficiency and accuracy of statistical analysis.

References

1. Xinhua. Decision of the CPC Central Committee on Further Comprehensively Deepening Reform. and Promoting Chinese-style Modernization [EB/OL]. http://www.news.cn/politics/20240721/cec09ea2bde840dfb99331c48ab5523a/c.html, 2024-07-21.
2. Dorestani, A., & Arjomand, L. H. (2006). An Empirical Study of the Relationship Between Consumer and Producer Price Index: A Unit Root Test and Test of Cointegration[J]. The Coastal Business Journal, 5(1), Article 4. https://digitalcommons.coastal.edu/cbj/vol5/iss1/4/, 2022-02-28.
3. National Bureau of Statistics. Consumer Price Index [EB/OL]. https://www.stats.gov.cn/zs/tjll/tjzs/202302/t20230224_1918473.html, 2023-02-24.
4. National Bureau of Statistics. How the Producer Price Index (PPI) is compiled [EB/OL]. https://www.stats.gov.cn/zs/tjws/tjzb/202301/t20230101_1903637.html, 2023-01-01. 105-122.
5. Sun Tianyang, Yang Danhui. The latest research progress and prospect of emerging industries: a. literature review [J]. Industrial economic review, 2022 (01) : 105-122, http://gjs.cssn.cn/kydt/kydt_kycg/202202/t20220223_5395170.shtml, 2022-02-23.
6. Dai Peichao, Zhang Rongjia, Zhou Yuan. Strategic emerging industry research hotspot and frontier. analysis [J]. Journal of management modernization, 2023 (4) : 175-182. http://iqte.cssn.cn/cycg/yjlw/202308/P020230823314496224301.pdf.
7. DENG Haiyun, Zhang Xi, Yang Junbo, Liu Qing, Yin Jiabin, et al. Research on Influencing factors and forecasting of CPI of urban and rural residents in China. Mathematical Practice and Understanding,2023(02):74-85.
8. International Monetary Fund.World Economic Outlook Database, April 2024[EB/OL]. https://www.imf.org/en/Publications/WEO/weo-database/2024/April,2024-05-21.
9. Ministry of Industry and Information Technology, Ministry of Science and Technology, National Energy Administration, National Standards Commission. New industry standardization pilot project implementation plan (2023-2035) [Z]. 2023-08-22.