CrossMark

ORIGINAL PAPER

Analysis of factors influencing the Henry Hub natural gas price based on factor analysis

Hong Li[1,2] · Hui-Ming Zhang[2] · Yuan-Tao Xie[3] · Di Wang[1]

Abstract Henry Hub as an important transaction hub for natural gas sets the gas price standard in the USA. In this paper, the factors influencing Henry Hub natural gas prices are analyzed, and the major factors determining the price levels in the period between January 1997 and December 2016 are studied. It is found that economic conditions, total energy demand, US dollar exchange rate and gas consumption are the major factors. The mechanism of each factor influencing the Henry Hub natural gas price is also explored in the paper.

Keywords Natural gas · Henry Hub · Factor analysis

1 Introduction

Natural gas is one of the most important types of energy used for industrial production and household use, and its price can have a great impact on the economic development of a country, including total production and inflation. The Henry Hub natural gas price is an important price level benchmark for the gas price not only in the USA but also in the worldwide market. Also, as a relatively cleaner energy than coal, wood and oil, natural gas can be exploited more broadly, and advancements in technology show a brighter future for the use of natural gas. This paper discusses the factors influencing the Henry Hub natural gas price.

van Goor and Scholtens (2014) mainly studied the explanation of the fluctuation of natural gas prices in England, and they found that the model based on supply and demand can successfully explain the instability in the gas price. Mu (2007) analyzes how weather influences the mean value and fluctuation of gas future prices and draws a conclusion that the market fundamentals are an important element affecting the gas price, but there are still some aspects of the price fluctuation that is not explained by the market fundamentals. Fazzio (2010) analyzed how the gas stock affects price and fluctuation. Nick and Thoenes (2014) studied the factors that influence the gas price in Germany in three cutoff periods and concluded that the lack of reserve and supply is the leading factor in short term, but in long term, the gas price is mainly influenced by the price of crude oil and coal and conditioned on the economic conditions and alternative energy sources.

Cong (2012) summarized the factors impacting on gas prices and also analyzes common factors that affect the global price and some special factors that influence the local price and found that the common factors include supply, demand, alternative energy and cost, and the major factors differ from area to area. Yu (2003) used the system dynamics method, established the ISM model, analyzing the superficial reasons, deep reasons and fundamental reasons for global gas prices, and found that the market supply, demand and competition are the most directly related reasons. Li and Wei (2003) compared the gas price within and outside China and put forward that the major

✉ Hong Li
  lihongpkupaper@126.com

[1] School of Economics, Peking University, Beijing 100871, People's Republic of China

[2] Center for National Resource Economy Studies, Beijing 100871, People's Republic of China

[3] School of Insurance and Economics, University of International Business and Economics, Beijing 100029, People's Republic of China

Edited by Xiu-Qin Zhu

Table 1 Factors of Henry Hub natural gas price

Variable	Unit	Mean	SD	Skewness	Kurtosis
Henry Hub natural gas price	Dollar/million Btu	4.66	2.27	1.27	4.91
US 48 states gas stock	Bcf	2376	790.67	− 0.09	2.08
US gas exploitation rate	MMcf	2,148,554	223,377	1.05	3.17
US gas total consumption	Bcf	64.56	13.6	0.84	2.75
US gas import	MMcf	236,917	72,993.81	− 0.77	2.49
CPI		198.92	24.94	− 0.02	1.64
US personal consumption	Billion dollars	8836	1953.56	− 0.12	1.79
US industrial production		96.4	6.02	− 0.26	2.64
GDP	Billion dollars	13,141	2705.76	− 0.09	1.79
Trade-weighted dollar index		110.0	9.27	0.36	2.07
US average pay level	Per month/per thousand non-agriculture	132,572	4199.49	− 0.30	3.00
Term premium	Percentage	1.76	1.2	− 0.25	1.89
US total population	Thousand person	297,582	14,405.38	− 0.11	1.80
US unemployment rate	Percentage	6.06	1.78	0.85	2.44
Stock market (S&P 500)		1268	289.94		
Risk-free interest rate	Percentage	2.28	2.11	0.33	1.47
Risk premium	Percentage	2.56	0.81	1.52	7.06
US carbon dioxide emissions	Million metric tons	473.46	35.09	0.25	2.63
US coal consumption	Thousand ton	86,591	9750.15	− 0.41	3.09
US 50 states alternative energy consumption	MMBBL/day	19.54	0.83	0.33	2.24
WTI oil price	Dollar/barrel	56.81	31.54	0.29	1.81

Table 2 Factor eigenvalue and variance contribution

Factor	Eigenvalue	Difference	Variance contribution	Accumulate variance contribution
Factor 1	6.64	0.96	0.36	0.36
Factor 2	3.67	0.33	0.20	0.56
Factor 3	3.34	1.05	0.18	0.74
Factor 4	2.29	0.89	0.12	0.86
Factor 5	1.40		0.08	0.94

impact factors of the domestic market are the upstream exploitation, transportation pipeline construction and distribution. Li et al. (2015) established a model for the relationship between the gas price and alternative energy, taking the theoretical price of gas as the correlation function of alternative energy (Ren and Sovacool 2014b), and made a forecast of the future price trend (Olsen et al. 2015).

Current literature from China and abroad mainly focuses on particular factors (Cheng and Guo 2007), such as the supply and demand fundamentals of natural gas price (Li and Wang 2007). But there are many other factors affecting natural gas prices, such as exchange rate (Ma 2011), financial market risk, total energy demand and alternative energy (Li et al. 2013). We should find out the main factors affecting the price of natural gas from these influencing variables. This paper refers to the major factors in the

present literature, including supply, consumption, GDP and oil consumption, and also the most commonly acknowledged economic index, including industrial production (Jiao et al. 2004). Twenty variables are selected, and a factor analysis and linear regression models are combined to study the factors (Ren et al. 2014; Ren and Sovacool 2014a).

2 Model and variable

Factor analysis is the analysis method that takes variance contribution as its choice index to select the most important underlying factors driving changes in the studied attributive variable. The general form of factor analysis is

Table 3 Factor loading

Variable	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Eigenvalue
US coal consumption	− 0.17	− 0.09	0.90	0.07	− 0.06	0.15
US 50 states oil consumption	0.25	− 0.26	0.76	− 0.29	− 0.15	0.19
US carbon dioxide emissions	− 0.05	− 0.15	0.83	− 0.10	0.50	0.03
US gas exploitation rate	0.58	0.28	− 0.53	0.25	0.27	0.17
Term premium	0.02	0.85	− 0.14	0.04	0.04	0.25
US gas imports	− 0.40	− 0.26	0.74	− 0.32	− 0.14	0.11
US total population	0.79	0.44	− 0.19	0.35	0.06	0.01
Risk premium	0.07	0.77	− 0.10	− 0.14	− 0.03	0.37
Risk-free interest rate	− 0.39	− 0.83	0.25	− 0.17	− 0.08	0.06
CPI	0.79	0.49	− 0.21	0.40	0.06	0.01
Trade-weighted dollar index	− 0.35	− 0.07	0.18	− 0.85	− 0.02	0.11
US average pay level	0.97	0.01	0.09	− 0.08	− 0.00	0.04
US GDP	0.84	0.35	− 0.17	0.36	0.05	0.01
US industrial production	0.97	− 0.11	0.09	0.02	0.03	0.03
US unemployment rate	0.10	0.77	− 0.24	0.52	0.01	0.07
US personal consumption	0.83	0.37	− 0.17	0.36	0.05	0.01
WTI oil price	0.70	0.24	− 0.10	0.59	− 0.04	0.09
US total gas consumption	0.16	0.09	0.00	0.04	0.93	0.09
US 48 states gas reserve	0.23	0.19	0.04	0.27	− 0.33	0.73
Stock market (S&P 500)	0.80	− 0.31	− 0.31	− 0.01	0.19	0.12

Table 4 Significance test

Independent variable	Coefficient	SD	t	p value
Economic factor	0.92***	0.10	9.06	0.00
Interest rate factor	− 0.13	0.10	− 1.31	0.19
Total energy demand factor	1.41***	0.10	13.77	0.00
Dollar factor	0.21**	0.10	2.11	0.04
Gas consumption factor	− 0.33***	0.10	− 3.15	0.002
Constant term	4.66***	0.10	45.94	0.00

***, **, *, represent 1%, 5% and 10% significance level, respectively

$$Y_{kt} = b_{k1}F_{1t} + b_{k2}F_{2t} + \cdots + b_{kn}F_{nt} + u_{kt}$$

where Y_{kt} is the tth observed value of the kth variable, b_{kn} is the loading of the nth factor of the kth variable, F_{nt} is the tth observed value of the nth factor and u_{kt} is the tth particularity of the kth variable.

In this paper, 20 variables are selected, including US total gas consumption, US gas imports and US personal consumption expenditure. We take monthly data, if there are only seasonal data, suppose the variable has average incremental change during the season, and if there are only daily data, we take the data of the last day in that month. The data come from the US Energy Information Agency (EIA) and the Federal Reserve Bank. There are in total 240 sets of data from January 1997 to December 2016. The descriptive statistics are listed in Table 1.

3 Empirical results and analysis

3.1 The applicability of factor analysis (Pan and Nie 2014)

(1) Bartlett's test proves that using the 20 variables to do factor analysis is applicable.
(2) The Kaiser–Meyer–Olkin (KMO) test value is 0.74, which suggests that it is in an ideal condition to do factor analysis.
(3) The sample size. In this paper, if we add more variables that might be related to the fluctuation of natural gas future price, such as the frequency of earthquakes, we may be able to find out more related variables. But the KMO will rapidly decline in this case, which means the applicability of factor analysis is smaller. The purpose of this work is to find out the most important factor influencing gas price by factor analysis and to study the influencing extent of each variable related to factors. So, only 20 variables are chosen. Besides, the size of samples should be large enough for the number of variables, and 240 sets of data are sufficient to carry the analysis.
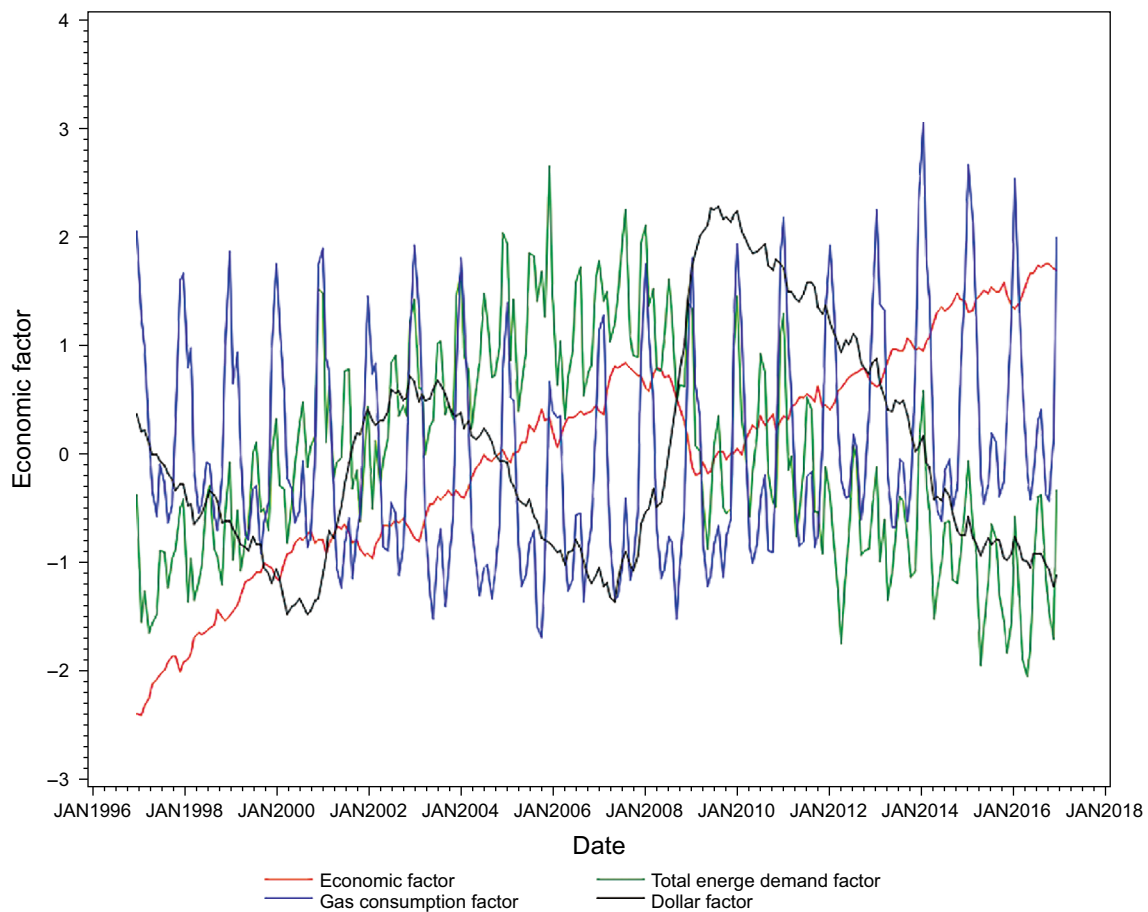
Fig. 1 Time series of variation trend (January 1997 to December 2016)

3.2 Factors selection

Factor eigenvalue and variance contribution of the model are listed in Table 2. In Table 2, we can see that the eigenvalues of the five factors are 6.64, 3.67, 3.34, 2.29 and 1.40, respectively. They are all greater than 1. The variance contributions of the five factors are 0.36, 0.20, 0.18, 0.12 and 0.08, respectively. The accumulate variance contribution of the total five factors is 0.93. These five factors are sufficient to represent all of the variables.

3.2.1 The choice of the number of factors

We sorted the 20 factors in descending order of eigenvalue. Among them, the first five factors' eigenvalues are all above 1. Table 2 shows that the accumulate variance contribution of the fifth factor is 93%. It is thus reasonable to choose the first five factors as the major price factors.

3.2.2 Matrix rotation of factors

The principle of each factor's varimax should follow that the variance of each variable's square will be a maximum, and no correlation is allowed between its factors.

3.3 Factor loading

Table 3 shows the factor loading after rotation. As listed in Table 3, the high absolute value of factor loading means that the variances of variable and factor overlap. The positive or negative sign of the value represents the difference of the change direction between variable and factor. The loading of factor 1 is ordered in the descending order of the absolute value—the pay level, industrial production, GDP, personal consumption and stock market. According to the naming principle of the maximum loading, we name factor 1 as the economic factor. The variables with comparatively low loading in the economic condition factors are term premium, risk premium and total carbon
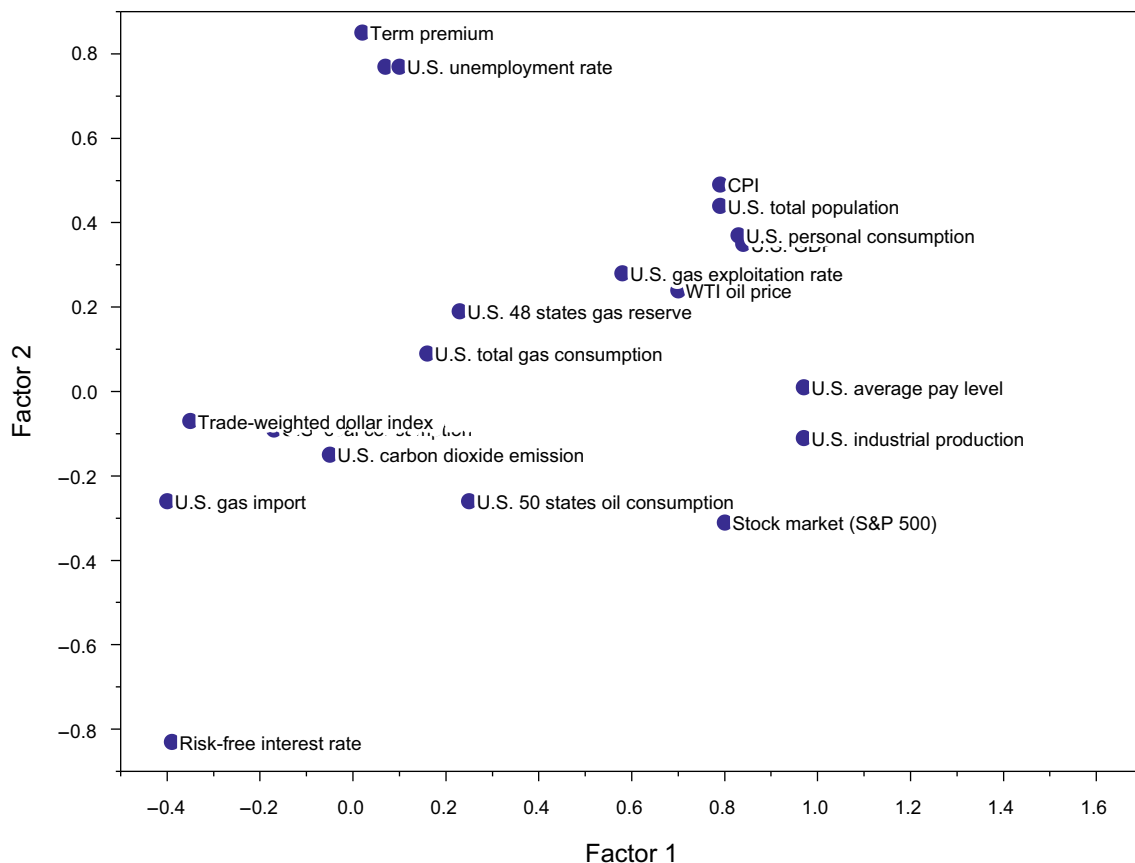
Fig. 2 Variables and factor loading matrix analysis (factor 1 and factor 2). *Notes*: *x*-axis and *y*-axis represent two corresponding factors; data point represents 20 variables

dioxide emissions. Thus, this naming is reasonable. Factor 2 has the biggest loading of term premium, risk-free interest rate and unemployment rate. It is thus named the interest rate factor. The variables with relatively low loading include total coal consumption, gas consumption, oil consumption, carbon dioxide emissions and gas imports. Thus, it is reasonable to name it total energy demand factor. Factor 4 has its biggest loading in trade-weighted dollar index, oil price, unemployment rate and CPI. The relationship between oil price and exchange rate is negative. CPI is related to exchange rate, and thus, it is named the dollar factor. The biggest loading of factor 5 appears in total gas consumption, and the loading is 0.93. The other loadings are all small. This means that the factor 5 should be taken as an independent factor, and it is named the gas consumption factor.

3.4 Significance test

Significance test results are listed in Table 4, where *t* stands for *t* statistics while *p* value is the probability for the statistical summary, which would be the same as or of greater magnitude than the actual observed results given the null hypothesis is true. Judging from the *p* value, except for the interest rate factor, the other four tests reject the null hypothesis. They have palpable effect upon the Henry Hub gas price, and the interest rate factor and gas consumption factor have opposite influences compared with others. Considering the economic factor, when the economic conditions are better (e.g., the higher the US industrial production), the economic factor carries more weight. Thus, the value 0.92 means that the gas price rises when US economy is strong. Judging from the variables and coefficient of the interest rate factor, the rise of risk-free interest will lead to a rise in the gas price. The rise of risk premium and term premium will lead to a decline in the gas price. The interest rate factor and the gas price fluctuation are independent, and the *p* value is 0.91. Thus, with the 10% significance level, the null hypothesis cannot be rejected. The total energy demand factor includes consumption of coal and oil, representing the amount of carbon dioxide emissions and gas imports. The coefficient is 1.41; thus, a rise in total energy demand leads to a rise in the gas price. The loading of the dollar index in the dollar
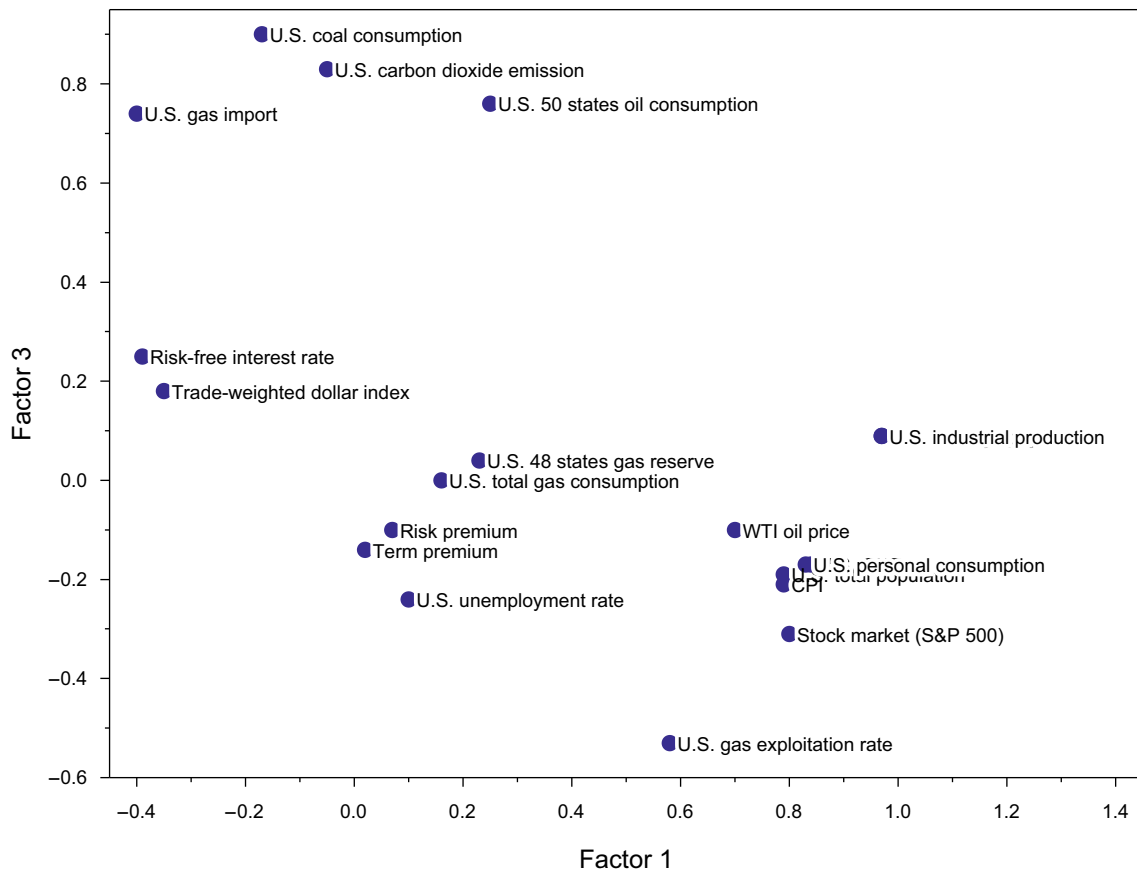
Fig. 3 Variables and factor loading matrix analysis (factor 1 and factor 3). *Notes*: *x*-axis and *y*-axis represent two corresponding factors; data point represents 20 variables

factor is negative, but the CPI is positive. The factor coefficient is 0.21, and the *p* value is 0.04; thus, when dollar appreciates, the gas price falls. And when CPI rises, the gas price also rises. The coefficient of total gas consumption is − 0.33. When total gas consumption rises, the gas price reduces.

3.5 Variation trend of factors and price fluctuation mechanism

The four factors are shown in Fig. 1.

The total energy demand factor and the gas consumption factor are obviously influenced by season. But the mechanisms are different. The total energy demand includes oil and coal, which are mainly used as industrial raw material and power-generating resources. The largest demand appears at November (back-to-school season) and December (Christmas consumption season). This is accordant with the US consumer habits. Thus, the choice and naming of factors are satisfactorily reasonable. The significance test shows that the total energy demand and

natural gas price are positively correlated. This means that the substitutional relationship between alternative energy and natural gas is not that obvious. The rise of other alternative energy does not mean a decline in the demand for gas. The impetus for alternative energy and gas consumption is probably the total demand for energy. When the total energy demand rises (including the gas demand), the gas price will rise. The gas consumption factor reaches its peak during January and March (winter) and reaches its valley in summer. Its variation trend is opposite to that of the total energy demand factor. But in winter, there is a maximum due to the heating energy consumption. This explains the difference existing among different kinds of energy. Natural gas is mainly used for house heating and cooking. It is also partly used as an industrial raw material. The peak that appears each November shows that the rise of demand for mass consumer product leads to the rise in gas demand.

The economic factor rises in accordance with time, but there is fluctuation sometimes. This is agreed with the developing trend of the US economy. The influence of the
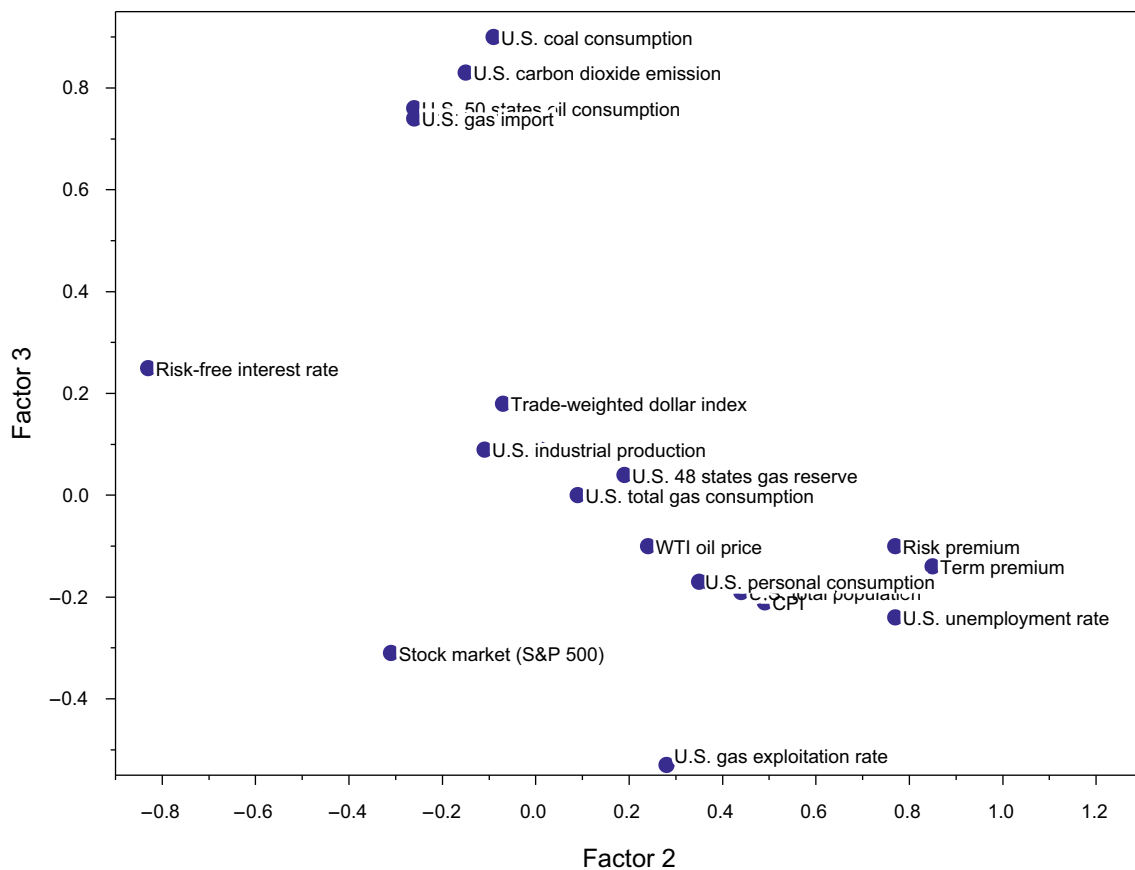
Fig. 4 Variables and factor loading matrix analysis (factor 2 and factor 3). *Notes*: *x*-axis and *y*-axis represent two corresponding factors; data point represents 20 variables

economic factor upon gas price is mainly like this—when the economy is strong, the rise of personal expenditure and demand of products will lead to a rise in demand for energy. In this way, the gas price rises.

The dollar factor shows an obvious declining trend from 2002 to 2008. Before 2002, it reduces gradually. From 2002 to 2008, it rises gradually. After a short-term fluctuation in 2008, it falls rapidly. The exchange rate is related to the economy and financial markets, and also currency policy and interest rates. The dollar exchange rate affects the cost of imports and exports of gas. If the dollar exchange rate rises, the nominal price of gas will decline. The actual export price of gas will rise compared with that of the overseas gas. A decline in demand lowers the gas price. CPI is also another influential factor. It has a positive correlation with gas price. WTI oil price makes up a large part in the dollar factor. But considering the negative correlation between oil price and exchange rate, the dollar factor is mainly affected by the exchange rate.

3.6 Analysis of factor loading matrix

The analytical result of variables and factor loading matrix is shown in Figs. 2, 3 and 4. A factor loading is the correlation between a variable and a factor that has been extracted from the data. From the factor loading matrix analysis, we can see the relationship between a variable and a factor obtained from orthogonal rotation. In Fig. 2, when compared with other variables, the coal consumption, oil consumption and carbon dioxide emissions are more independent and irrelevant with the other variables. The US GDP, industrial production, exchange rate and average pay level are more closely related. These are the indexes to measure US economy (Lin and Wesseh 2013). Variables that are related to interest like the risk-free rate and term premiums are also the same in their loading.

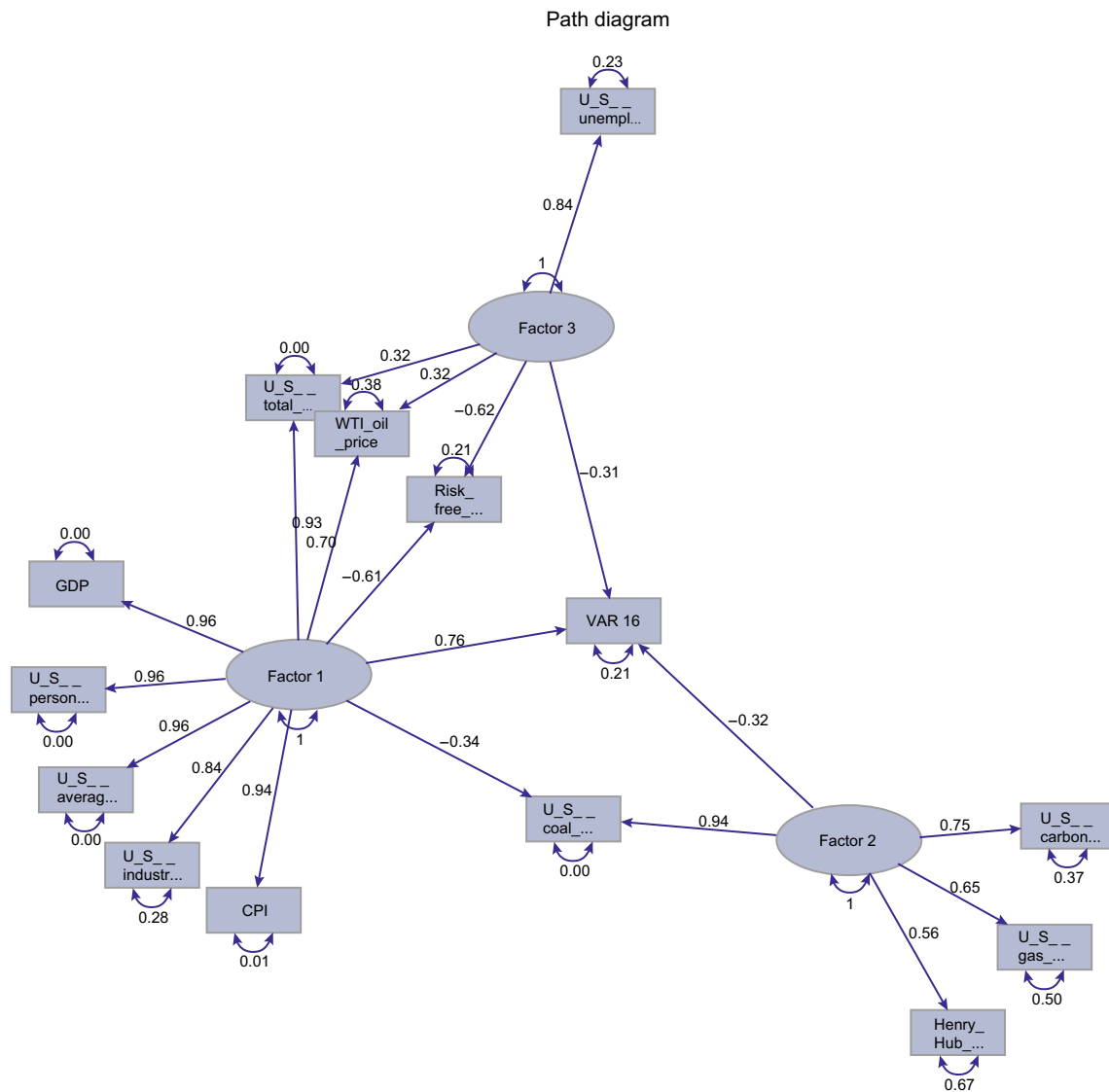The path diagrams are shown in Fig. 5 when we focus on the first three factors.

Path diagram

Fig. 5 Path diagram analysis (factor 1–factor 3)

4 Conclusion

This paper uses factor analysis to test 20 variables that may influence the gas price from January 1997 to December 2016. Among them, five major factors are selected out from the 20 variables—economic factor, interest rate factor, total energy demand factor, dollar factor and gas consumption factor. Linear regression analysis finds that the gas price is mainly affected by natural gas demand, alternative energy consumption, dollar exchange rate and the economic condition. The interest rate factor only influences the gas price little. Variables that represent the economic conditions are similar in their relationship with each factor, like US GDP, industrial production and average pay level. The loading of coal consumption, oil consumption and carbon dioxide emissions is similar (Li et al. 2015). The relationship between interest rate and exchange rate is quite independent when compared with other variables (Wang 2012).

References

Cheng P, Guo Y. Determination of natural gas price. GAS HEAT. 2007;27(8):36–41. doi:10.3969/j.issn.1000-4416.2007.08.012 **(in Chinese)**.

Cong R. The study of impact factors of global natural gas price. West Forum Econ Manag. 2012;23(04):61–69, 82. doi:10.3969/j.issn. 2095-1124.2012.04.012 **(in Chinese)**.

Fazzio TJ. A statistical analysis of the natural gas futures market: the interplay of sentiment, volatility and prices. M.B.A's degree, Massachusetts Institute of Technology. 2010. http://hdl.handle. net/1721.1/59118.

Jiao J, Fan Y, Wei Y. Review of oil price study. Energy China. 2004;26(04):34–40. doi:10.3969/j.issn.1003-2355.2004.04.009 **(in Chinese)**.

Li F, Wei D. Discussion on price of natural gas. GAS HEAT. 2003;23(10):579–82. doi:10.3969/j.issn.1000-4416.2003.10.001 **(in Chinese)**.

Li H, Tian X, Zou T. Impact analysis of coal-electricity pricing linkage scheme in China based on stochastic frontier cost function. Appl Energy. 2015;151(15):296–305. doi:10.1016/j. apenergy.2015.04.073.

Li H, Xie M, Zhang T. Promote the development of renewable energy: a review and empirical study of wind power in China. Renew Sustain Energy Rev. 2013;22:101–7. doi:10.1016/j.rser. 2013.01.025.

Li T, Wang X. Discussion on the determination of natural gas price. Nat Gas Technol. 2007; 1 01: 33–38+90. http://kns.cnki.net/kns/detail/ detail.aspx?FileName=TRJJ200701007&DbName=CJFQ2007. **(in Chinese)**.

Lin B, Wesseh PK. What causes price volatility and regime shifts in the natural gas market. Energy. 2013;55(1):553–63. doi:10.1016/ j.energy.2013.03.082.

Ma H. Research into domestic and overseas natural gas price formation mechanisms. Master's degree, China University of Geosciences (Beijing). 2011. http://cdmd.cnki.com.cn/Article/ CDMD-11415-1011078554.htm. **(in Chinese)**.

Mu X. Weather, storage, and natural gas price dynamics: Fundamentals and volatility. Energy Econ. 2007;29(1):46–63. doi:10.1016/ j.eneco.2006.04.003.

Nick S, Thoenes S. What drives natural gas prices?—a structural VAR approach. Energy Econ. 2014;45(5):517–27. doi:10.1016/j. eneco.2014.08.010.

Olsen KK, Mjelde JW, Bessler DA. Price formulation and the law of one price in internationally linked markets: an examination of the natural gas markets in the USA and Canada. Ann Reg Sci. 2015;54(1):117. doi:10.1007/s00168-014-0648-7.

Pan Y, Nie J. Research on the M & A performance difference based on factor analysis. Econ Probl. 2014;9(01):115–8. doi:10.16011/ j.cnki.jjwt.2014.01.023.

Ren J, Andreasen KP, Sovacool BK. Viability of hydrogen pathways that enhance energy security: a comparison of China and Denmark. Int J Hydrog Energy. 2014;39(28):15320–9. doi:10. 1016/j.ijhydene.2014.07.150.

Ren J, Sovacool BK. Enhancing China's energy security: determining influential factors and effective strategic measures. Energy Convers Manag. 2014a;88:589–97. doi:10.1016/j.enconman. 2014.09.001.

Ren J, Sovacool BK. Quantifying, measuring, and strategizing energy security: determining the most meaningful dimensions and metrics. Energy. 2014b;76:838–49. doi:10.1016/j.energy.2014.08.083.

van Goor H, Scholtens B. Modeling natural gas price volatility: the case of the UK gas market. Energy. 2014;72(7):126–34. doi:10. 1016/j.energy.2014.05.016.

Wang Z. The price mechanism of fossil fuels in China. Beijing: China Social Sciences Publishing House; 2012. http://www.bookask. com/book/180365.html. **(in Chinese)**.

Yu X. Analysis of factors influencing the natural gas price by using ISM. Chem Techno-Econ. 2003;21(10):30–2. doi:10.3969/j.issn. 1673-9647.2003.10.012 **(in Chinese)**.