

## An Empirical Study on the Cointegration Relationship between Energy Consumption, Energy Supply and Their Influencing Factors: Postprint

**Authors:** Li Mingjia, Junjie Yan, Tao Wenquan

**Date:** 2017-10-17T00:00:00+00:00

### Abstract

This study employs cointegration theory from econometrics to quantitatively investigate the degree of fluctuation impact among energy consumption, energy supply, and factors including energy prices, economic development, industrial structure, population, and urbanization. First, the Augmented Dickey-Fuller model is adopted to test the stationarity of time series data for indicator variables. Second, the Lagrange multiplier model is employed to quantitatively calculate the cointegration relationship between energy consumption and energy supply at the national level and their relevant influencing factors. The results indicate that both energy consumption indicators and population and urbanization indicators have significant impacts on energy supply; energy consumption is primarily influenced by industrial structure and economic indicators, while being less affected by energy price fluctuations.

### Full Text

#### Preamble

#### A Cointegration Analysis among Influential Indicators, Energy Consumption and Energy Supply in China

**Ming-Jia Li, Jun-Jie Yan, Wen-Quan Tao\***

*Key Laboratory of Thermo-Fluid Science and Engineering of Ministry of Education, School of Energy and Power Engineering, Xi'an Jiaotong University, Xi'an, Shaanxi, 710049, China*

#### Abstract

This paper employs cointegration theory from econometrics to quantitatively analyze the volatility relationships among energy consumption, energy supply,

and their influencing variables including energy price, economic development, industrial structure, population, and urbanization. First, the Augmented Dickey-Fuller (ADF) model is used to examine the stationarity of time series data for all indicator variables. Second, the Lagrange Multiplier (LM) model is applied to quantitatively calculate the cointegration relationships between energy consumption and its influencing variables, as well as between energy supply and its influencing factors. The results indicate that energy supply is significantly affected by energy consumption, population, and urbanization indicators. Energy consumption is primarily influenced by industrial structure and economic indicators, while being only slightly affected by energy price fluctuations.

**Keywords:** Energy demand and supply; Cointegration analysis; ADF model; LM model

**Funding:** This research was supported by the National Key Basic Research Development Program (973 Program) (2013CB228304) and the China Postdoctoral Innovative Talent Support Program.

**Corresponding Author:** Wen-Quan Tao, Professor, wqtao@mail.xjtu.edu.cn

---

In April 2017, China's National Energy Administration released the *Energy Production and Consumption Revolution Strategy (2016-2030)*, which emphasizes that from 2017 to 2020, China's total energy consumption must be controlled within 5 billion tons of standard coal equivalent [1]. The *Energy Work Guidance* issued by the National Energy Administration in 2016 further stated that the "dual control" task of controlling both total energy consumption and energy intensity is the primary priority for China's energy development [2]. Broadly speaking, numerous factors influence energy consumption and supply, and these factors mutually affect each other through volatility effects. Any macroeconomic influencing factor can produce volatility effects on both energy supply and consumption [3]. This paper deeply investigates the volatility cointegration relationships between national energy supply, energy consumption, and key influencing factors, providing a foundation for subsequent optimization and development recommendations.

The analysis examines five specific aspects: the impact of energy consumption-related indicators on energy supply, the influence of economic indicators on energy consumption, the effect of industrial structure-related indicators on energy consumption, the impact of energy price-related indicators on energy consumption, and the influence of population and urbanization-related indicators on energy supply.

## 1. Indicator Construction and Data Selection

Given that this study aims to investigate the mutual constraints between national energy consumption, energy supply, and their influencing factors under conditions of undesirable outputs, we selected a total of 12 indicators across

four dimensions: energy supply, energy consumption, economy, and industrial structure. Panel data for these 12 indicators were collected for China from 1990 to 2015. Due to missing data, Tibet, Hong Kong, Macau, and Taiwan were excluded from the evaluation scope. Data for each indicator were primarily sourced from authoritative publications including the *Energy Statistical Yearbook*, *China Population Statistical Yearbook*, *China Industrial Economy Statistical Yearbook*, and *China Statistical Yearbook*, as well as the National Bureau of Statistics progressive database [4]. All data were directly collected from statistical sources, ensuring scientific reliability and quality. The following sections provide detailed explanations of indicator selection and data collection methods.

### 1.1 Energy Supply Indicators

Based on all current normative documents referenced, data for “energy available for consumption in this region” were extracted sequentially from the *Energy Balance Sheets* for China from 1990 to 2015.

### 1.2 Energy Consumption Indicators

Energy consumption comprises four indicators: total energy consumption, energy price, energy consumption per unit GDP, and per capita energy consumption. Data for all four indicators were collected nationally from 1990 to 2015. Energy price includes domestic coal prices, international coal prices, and the raw materials, fuel, and power price index. Per capita energy consumption was calculated as follows: First, annual total energy consumption from 1990 to 2015 was obtained from the *China Statistical Yearbook*. Then, year-end total population for each year was queried from the National Bureau of Statistics website. Per capita energy consumption for China from 1990 to 2015 was derived by dividing annual total energy consumption by the corresponding year-end total population.

### 1.3 Economic Indicators

Economic indicators consist of five metrics: per capita GDP, regional GDP deflator, capital stock, total fixed asset investment, and population and urbanization. All five indicators were obtained from the National Bureau of Statistics website for China from 1990 to 2015.

### 1.4 Industrial Structure Indicators

The industrial structure subsystem comprises two indicators: the proportion of secondary industry in GDP and the proportion of tertiary industry in GDP. Data for both indicators were collected nationally from 1990 to 2015.

## 2. Model Specification

The empirical models in this paper are calculated as follows: First, the Augmented Dickey-Fuller (ADF) test model is employed. This model tests the stationarity of all indicator variable data to determine whether the data can be applied in subsequent calculations. Second, the Lagrange Multiplier (LM) model is used with a lag order of 1 to conduct cointegration analysis calculations for energy consumption, energy supply, and their influencing factors.

### 2.1 Stationarity Test

In econometrics, stationarity of time series data is a necessary condition, requiring that curves fitted from time series sample data maintain their existing “inertia” volatility for a future period [5]. If data are non-stationary, it indicates that the curve lacks continuity characteristics, meaning the new curve fitted from future time series data will exhibit different volatility patterns than the current sample fitted curve. Such data cannot be used for further research and must undergo higher-order differencing before re-evaluation. Since factors affecting energy consumption and supply fluctuate over time, stationarity testing forms the foundation for cointegration analysis and subsequent model calculations.

The ADF test model formulas extend from formula (1) to formula (3). Adding a constant term to Model 1 yields the specification. Adding a trend variable to Model 2 yields the specification. In these formulas,  $X_t$  represents the difference of the explained variable (starting from first-order difference);  $\mu$ ,  $\alpha$ ,  $\beta$ , and  $\gamma$  are constant terms;  $t$  represents the year;  $k$  is the lag order of  $X_t$ ; and  $\epsilon_t$  represents residuals.

The null hypothesis of this test is:  $H_0: \alpha = 0$ , indicating that variable  $X_t$  is a non-stationary time series requiring higher-order differencing. If the test statistic rejects the null hypothesis, the variable is a stationary time series.

### 2.2 LM Model

Following the time series stationarity test, the LM model is calculated. The calculation model is shown in formula (4), where  $\mu = \sigma^2$  represents the variance of  $X_t$  (calculated from formula (3));  $x_{t-i}$  represents the time series of each explanatory variable;  $q$  is the lag number; and  $\rho$  represents the correlation coefficient. This model proposes the hypothesis that there is no time series correlation at order  $p$ . In this test,  $p$  is assumed to equal 1, testing for first-order serial correlation [6]. If the chi-square calculated value is less than the observed goodness-of-fit, the null hypothesis is rejected, indicating that the time series values of each indicator variable exhibit correlation.

### 3.1 Data Collection and Processing

First, all indicator variable data were homogenized. In comprehensive indicator evaluation calculations, each indicator represents different physical meanings, resulting in dimensional differences and varying magnitude trends. By analyzing the nature of each indicator variable, six indicators were identified as inverse indicators: energy available for consumption in this region, total energy consumption, energy price, energy consumption per unit GDP, per capita energy consumption, and proportion of secondary industry in GDP. Six indicators were identified as positive indicators: regional per capita GDP, regional GDP deflator, capital stock, total fixed asset investment, population and urbanization, and proportion of tertiary industry in GDP. Second, since the units of all indicator variables differ, data dimensionless processing was required.

### 3.2 Stationarity Test

The ADF test model was applied to calculate stationarity for 12 indicator variables including total energy consumption, energy price, and energy consumption per unit GDP. Variable descriptions are provided in Table 1. Test results are shown in Table 2. All stationarity and non-stationarity conclusions in Table 2 were derived at the 5% significance level. The results show that test values are below the 5% significance level, indicating that all variables have no unit root and are stationary time series data suitable for further model calculations.

### 3.3 LM Model Cointegration Analysis

The model first assumes no time series correlation at order  $p$  for the variable time series data. In this test,  $p$  is assumed to equal 1, testing for first-order serial correlation. The chi-square calculated value is less than the observed goodness-of-fit, thus rejecting the null hypothesis and indicating that the time series values of each indicator variable exhibit correlation.

#### 1) Impact of Energy Consumption Indicators on Energy Supply

Following the stationarity test in Section 3.2, this study further establishes a panel data model with the logarithm of energy supply as the dependent variable and the logarithms of total energy consumption and energy consumption per unit GDP as explanatory variables. The model specification is:

$$\ln(S) = C + \beta_1 \ln(PTC) + \beta_2 \ln(EGDP) + \varepsilon_t$$

where  $C$  is the constant term;  $S$  represents energy supply;  $PTC$  is per capita energy consumption;  $EGDP$  is energy consumption per unit GDP; and  $\beta_1$  and  $\beta_2$  are coefficients for the two variables; and  $\varepsilon_t$  is the residual error term representing unestimable random disturbance.

After dimensionless and normalization processing of explanatory variables, the panel data analysis results for the impact of energy consumption-related indi-

icators on energy supply are shown in Table 3 . The results indicate that all these variable indicators exhibit a unidirectional strong correlation with energy supply. Historical data analysis reveals that China' s energy intensity peaked at 1.6 in 2005, after which it declined rapidly and fluctuated slowly. However, due to the high growth rate maintained by the total economic volume, energy consumption remains high. Figure 1 [Figure 1: see original paper] shows the energy intensity trend from 1990 to 2015, while Figure 2 [Figure 2: see original paper] illustrates the national energy consumption and economic growth trends from 1990 to 2014.

## 2) Impact of Economic Indicators on Energy Consumption

A panel data model was established with the logarithm of energy consumption as the dependent variable and the logarithms of per capita energy consumption, per capita GDP, regional GDP deflator, capital stock, and total fixed asset investment as independent variables. The model specification is:

$$\ln(EC) = C + \beta_1 \ln(PTC) + \beta_2 \ln(PGDP) + \beta_3 \ln(GDPD) + \beta_4 \ln(CAPITAL) + \beta_5 \ln(INV) + \varepsilon_t$$

where C is the constant term; EC represents total energy consumption; PTC denotes per capita energy consumption; PGDP is regional per capita GDP; GDPD is the regional GDP deflator; CAPITAL represents capital stock; INV denotes total fixed asset investment;  $\beta_1$  through  $\beta_5$  are coefficients for the five variables; and  $\varepsilon_t$  is the residual error term.

After dimensionless and normalization processing of explanatory variables, the panel data analysis results for the impact of economic indicators on energy demand are shown in Table 4 . The "Coefficient" column in the table represents the degree of volatility impact between the corresponding independent and dependent variables. The model requires taking absolute values of coefficients: values less than 0.5 indicate weak impact; values between 0.5 and 0.8 indicate moderate impact; and values greater than 0.8 indicate strong impact. For example, the coefficient for regional per capita GDP (PGDP) is 3.435 (greater than 0.8), and the coefficient for capital stock (CAPITAL) is 1.620 (greater than 0.8), proving that economic growth significantly affects per capita energy consumption. In contrast, the coefficient for total fixed asset investment (INV) is 0.473 (less than 0.5), indicating that fixed asset investment has minimal impact on energy consumption. During economic growth periods, numerous investment projects with long construction cycles create energy supply shortages, leading to long-term supply-demand imbalances. During economic downturns, since most energy consumption originates from industrial sectors and the overcapacity resulting from large-scale investment projects during growth periods, industrial energy consumption rates decrease, creating supply surpluses [8].

## 3) Impact of Industrial Structure Indicators on Energy Consumption

A panel data model was established with the logarithm of energy consumption as the dependent variable and the logarithms of secondary industry proportion

in GDP and tertiary industry proportion in GDP as independent variables. The model specification is:

$$\ln(EC) = C + \beta_1 \ln(WJGDP2) + \beta_2 \ln(WJGDP3) + \varepsilon_t$$

where C is the constant term; EC represents total energy consumption; WJGDP2 denotes the proportion of secondary industry in GDP; and WJGDP3 represents the proportion of tertiary industry in GDP.

After dimensionless and normalization processing of variables, the calculation results are shown in Table 5 . The coefficient for secondary industry proportion (WJGDP2) is 1.223, and the coefficient for tertiary industry proportion (WJGDP3) is 0.829, both greater than 0.8. This proves that both secondary and tertiary industry proportions currently have strong significant impacts on energy consumption, with the secondary industry being particularly dominant.

#### 4) Impact of Energy Price Indicators on Energy Consumption

In economic theory, energy price plays a crucial role in affecting commodity demand. If market energy price mechanisms can influence regional energy demand to a certain extent, it indicates relatively large price elasticity. Therefore, this section adopts domestic coal prices, international coal prices, and raw materials, fuel, and power price indices as indicator variables to analyze the impact of price factors on energy demand.

A panel data model was established with the logarithm of provincial total energy consumption as the dependent variable and the logarithms of domestic coal price, international coal price, and raw materials, fuel, and power price index as independent variables. The model specification is:

$$\ln(EC) = C + \beta_1 \ln(PC) + \beta_2 \ln(PI) + \beta_3 \ln(K) + \varepsilon_t$$

where EC represents provincial total energy consumption; C is the constant term; PC denotes domestic coal price; PI is international coal price; K represents raw material price; and t is the residual error term.

After dimensionless and normalization processing of variables, the cointegration results are shown in Table 6 . The coefficient for domestic coal price (PC) is 0.423, the coefficient for raw materials, fuel, and power price (K) is 0.676 (between 0.5 and 0.8), and the coefficient for international coal price (P) is 0.292 (less than 0.5). This indicates that domestic coal price has a moderate impact on provincial energy demand, international energy price has minimal impact, and changes in the raw materials, fuel, and power price index have relatively large impact. From a computational perspective, these conclusions do not fully conform to commodity demand theory under market economics. The reasons are as follows: First, it relates to the current characteristics of China' s energy market, where energy prices are regulated by the government and not completely influenced by international market economics [9][10]. Second, the marketization

of new energy is not fully mature. Since most enterprises in the power system are state-owned, they can receive government subsidies even when operating at a loss. Consequently, enterprises lack strong incentives to choose new energy as an alternative, resulting in weak cointegration between energy demand and international energy prices.

### 5) Impact of Population and Urbanization Indicators on Energy Supply

Population growth, urbanization rate, and employment proportion are all closely related to energy supply and regional economic growth. This model examines the cointegration relationships between energy supply, energy consumption, and their influencing factors through quantitative calculations, analyzing five aspects: the impact of energy consumption on energy supply, the influence of economy on energy consumption, the effect of industrial structure on energy consumption, the impact of energy price on energy consumption, and the influence of population and urbanization on energy supply. This section adopts the proportion of secondary industry employees, urbanization rate, proportion of tertiary industry employees, population density, and total regional population as indicator variables to analyze the impact of population and urbanization-related indicators on energy supply.

A panel data model was established with the logarithm of energy supply as the dependent variable and the logarithms of urbanization rate, proportion of secondary industry employees, population density, and total regional population as independent variables. The model specification is:

$$\ln(S) = C + \beta_1 \ln(UR) + \beta_2 \ln(WJ2) + \beta_3 \ln(WJ3) + \beta_4 \ln(ROU) + \beta_5 \ln(P_{total}) + \varepsilon_t$$

where S represents energy supply; UR denotes urbanization rate; WJ2 is the proportion of secondary industry employees; WJ3 represents the proportion of tertiary industry employees; ROU is population density; Ptotal is total regional population; and t is the residual error term.

After dimensionless and normalization processing of variables, the cointegration results are shown in Table 7. During periods of population growth, strong consumption demand directly affects supply. The coefficient for secondary industry employee proportion (WJ2) is 1.165, the coefficient for tertiary industry employee proportion (WJ3) is 1.274, the coefficient for population density (ROU) is 4.848, and the coefficient for total regional population is 4.410, all indicating strong significant relationships. This proves that these factors significantly impact energy supply. The coefficient for urbanization rate is 0.521, indicating a moderate impact on energy supply under the existing industrial structure. China's current urbanization process is in an expansion phase, thus rising urbanization rates further drive increases in energy supply.

## Conclusions and Recommendations

This study employed cointegration theory and econometric methods from econometrics to analyze the cointegration relationships between energy consumption, energy supply, and their influencing factors. The following conclusions were obtained: First, both energy consumption and economic indicators currently have strong cointegration relationships with energy consumption and significantly influence its trends. Second, industrial structure has a long-term equilibrium relationship with energy demand. Both secondary and tertiary industry proportions currently have strong significant impacts on total energy consumption, particularly the secondary industry. Third, the cointegration relationship between energy price and energy demand is relatively weak, primarily due to characteristics of China's domestic market and incomplete maturity of new energy marketization. Fourth, population and urbanization currently have significant impacts on energy supply, and under the existing industrial structure, secondary industry employees also increase significantly in densely populated and growing regions.

Based on this research, the following recommendations are proposed: First, given the current industrial structure, coal still dominates China's primary energy consumption. Therefore, technological investment should be increased to develop clean coal technology to effectively address environmental pollution from coal utilization. This technology will solve three problems: (1) reducing pollutant and greenhouse gas emissions; (2) gradually reducing dependence on imported oil and coal; and (3) improving combustion efficiency of coal. Second, energy consumption diversification should be promoted. The transformation of residential energy consumption should be effectively incorporated into renewable energy utilization plans, with increased support for commercial markets. For example, government financial institutions should provide low-interest loans for energy-efficient equipment investments and residential investments (such as insulation materials) while implementing tax reduction measures for these devices [11]. Third, new energy transportation should be encouraged. For the existing new energy transportation industry, the government should implement corresponding policies and market expansion funding programs aimed at encouraging collaborative cooperation through active innovation and systematic R&D projects, thereby creating closer linkages between the energy sector, transportation sector, and commercial markets in terms of technology and innovation policy.

## References

- [1] Energy Production and Consumption Revolution Strategy (2016-2030) [EB/OL] [2017.5] <http://www.sdpc.gov.cn/zcfb/zcfbtz/201704/W020170425509386101355.pdf>
- [2] World Economic Outlook [EB/OL] [2017.5] <http://www.imf.org/external/chinese/pubs/ft/weo/2017/update>
- [3] Li, M.J., He, Y.L. and Tao, W.Q., Modeling a hybrid methodology for evaluating and forecasting regional energy efficiency in China [J]. *Applied Energy*, 2017, 185:
- [4] National Bureau of Statistics, *China Energy Yearbook* (1990-2015) [Z].

Beijing: China Statistics Press.

- [5] Bollerslev, T., Generalized autoregressive conditional heteroskedasticity [J]. *Journal of Economics*, 1986, 31:
- [6] Efimova, O., Serletis, A., Energy markets volatility modelling using GARCH[J]. *Energy Economics*, 2014, 43:264-273.
- [7] Li, M.J., Tao, W.Q., Review of methodologies and polices for evaluation of energy efficiency in high energy-consuming industry [J]. *Applied Energy*, 2017, 187: 203-215.
- [8] Geng W, Ming Z, Lili P, et al. China' s new energy development: Status, constraints reforms[J]. *Renewable and Sustainable Energy Reviews*, 2016, 53:
- [9] Peersman G, Van Robays I. Cross-country differences in the effects of oil shocks [J]. *Energy Economics*, 2012, 34(5): 1532-1547.
- [10] Ibor BI, Gerschenkron' s structuralist hypothesis effect on corporate restructuring and economic development: implications for regulatory incentives [J], *Journal of Finance and Bank Management*, 2015, 3(1): 90-99.
- [11] Jeong J, Hong T, Ji C, et al. Development of a prediction model for the cost saving potentials in implementing the building energy efficiency rating certification [J]. *Applied Energy*, 2017, 189: 257-270.

*Note: Figure translations are in progress. See original paper for figures.*

*Source: ChinaXiv –Machine translation. Verify with original.*