Research on Relationship Between Urbanization and Coal Consumption Based on Cluster Analysis Techniques

Mingzhu Liang[1], and Ziwen Zhang[2, *]

[1]Department of Financial Mathematics, Dongbei University of Finance and Economics, Liaoning, China

[2]Department of statistics, The University of Warwick, Coventry, United Kingdom

* Corresponding Author Email: u2106667@live.warwick.ac.uk

Abstract. Since the initiation of the reform and opening-up policy, urbanization in China has seen rapid growth. At the same, the consumption of coal has also been on the rise year by year. This study selects the coal consumption of 30 provinces as the research subject. By employing the cluster analysis method, this paper compares the coal consumption in different regions and summarizes the correlation between the level of urbanization and coal consumption. The research reveals a nonlinear correlation between coal consumption and the level of urbanization: at lower urbanization rates, there is a positive correlation between urbanization and coal consumption. However, once the urbanization rate reaches a certain threshold, the relationship becomes inversely correlated. This paper offers insights for addressing industrial relations and energy level regulation during the urbanization process.

Keywords: Cluster Analysis Techniques, Coal Consumption, Urbanization.

1. Introduction

Since the reform and opening-up, China's urbanization has advanced at a pace surpassing the global average during the same period. However, considering China's vast territorial area, large population and heightened economic growth rate, our current level of urbanization does not adequately meet the demands of the current situation. Urbanization is an important strategy in China's modernization process, which gathers funds, population and technology, promotes industrial structure adjustment, and industrial structure optimization will further affect coal consumption.

The urbanization process has aggregated capital, population and technology, facilitating the restructuring of the industrial framework. Moreover, the optimization of the industrial structure will further influence coal consumption. From 2000 to 2020, accompanying China's surge in urbanization rate from 36.2% to 63.9%, coal consumption also increased simultaneously, growing from 1,356.90 to 4,048.60 million tons.

Research by Liu indicates that Shanxi, as a major coal-producing province and also the one with the largest coal reserves in China, experienced an upward trend in coal consumption from 1997 to 2017. However, the growth rate has generally decelerated over time [1].

Recent research indicates that while the influence of economic growth on China's carbon emissions is progressively diminishing, there remains a positive correlation. Based on the empirical study using PSTR by Zhang and Wang, when the urbanization rate is below a certain threshold, economic growth significantly exacerbates carbon emissions. However, as the urbanization rate exceeds this threshold, the impact of economic growth on carbon emissions gradually weakens. As of 2019, only Beijing, Shanghai, and Tianjin have urbanization rates exceeding this threshold [2]. This suggests that coal consumption, being a primary energy consumption and a major source of carbon emissions, is to some extent influenced by economic growth. The research by Wang and Wang reveals that in the central and western regions of China, urbanization initially promotes and subsequently suppresses coal consumption. Coal consumption also exerts a long-term inhibitory effect on the upgrading of the industrial structure [3]. Furthermore, recent studies indicate that from 2008 to 2018, coal consumption has been gradually decoupling from economic growth under the influence of industrial structure transformation. By 2013, due to national coal control policies, the degree of decoupling transitioned

from weak to strong and economic development is gradually no longer at the cost of high coal consumption [4].

From a more micro perspective, the research conducted by Wang and Su demonstrates that urbanization suppresses coal consumption in Hebei; however, its impact on coal consumption in the Beijing-Tianjin region is not significant. Regarding Shanxi's coal consumption, both economic growth and urbanization serve as principal influencing factors [5]. The studies by Ding and Zhang indicate that in the long run, the influence of urbanization levels in the Beijing-Tianjin-Hebei region on coal consumption is more significant than in the Yangtze River Delta region [6]. Yu's study analyzed the mechanism of the impact of industrial structure changes on urbanization. Empirical research reveals that the advanced industrial structure in the eastern region significantly promotes urbanization development. The urbanization development in the central region is influenced by both the rationalization and the advanced industrial structure, whereas in the western region, the rationalization of the industrial structure evidently promotes urbanization development [7].

Based on the previous literature, this paper employs cluster analysis to examine four indicators: gross domestic product (GDP), total investment in fixed assets, urban population and urbanization rates. The goal is to investigate their influence on urban coal consumption. By summarizing the correlation between China's coal consumption and urbanization, this study lends support and evidence to existing literature. Concurrently, it can also clarify the specific values that transform the relationship between coal consumption and urbanization, providing a reference for the measures that the country needs to take to promote urbanization in order to control the process of urbanization better.

2. Data, Variables and Method

2.1. Data

The dataset for this article which includes coal consumption data utilized as an indicator is derived from the China Economic Net's Industry Database and data is selected from the China Statistical Yearbook. The remaining four metrics, including regional gross domestic product (GDP), total investment in fixed assets, urban population and urbanization rate. The dataset for this study comprises annual data for 2019 from 30 Chinese provinces. During data analysis, the Z-score standardization method is applied to each data point.

Figure 1 shows a combination of bar charts which represent regional urban population distribution, coal consumption volumes, GDP and total investment in social fixed assets, and a line graph which represents urbanization rates. Gansu, Yunnan, and Guizhou are cities with lower urbanization rates and correspondingly low coal consumption volumes. In contrast, cities like Jiangsu and Guangdong, with higher urbanization rates, demonstrate higher coal consumption than cities with lower urbanization rates. Beijing, Shanghai, and Tianjin, which have the highest urbanization rates, but their coal consumption is significantly low. Anomalies to these observations are Shandong and Inner Mongolia; their coal consumption volumes are the highest, yet their urbanization rates aren't as so high, leading to a mismatch between urbanization and coal consumption.

2.2. Variables

This article mainly uses five indicators for research: coal consumption, regional gross domestic product, total investment in social fixed assets, urban population and urbanization rate. Specifically, coal consumption volume is defined as the cumulative coal consumed within a province in 2019; regional gross domestic product (GDP) represents the final output of all resident units' production activities within a specific region in 2019; total investment in social fixed assets refers to the aggregate monetary value representing efforts and associated costs to construct and acquire fixed assets throughout society in 2019; urban population is characterized by the count of individuals residing in cities or towns in 2019; urbanization rate represents the proportion of the urban population in 2019.

The above metrics and their definitions are tabulated as shown in Table 1.

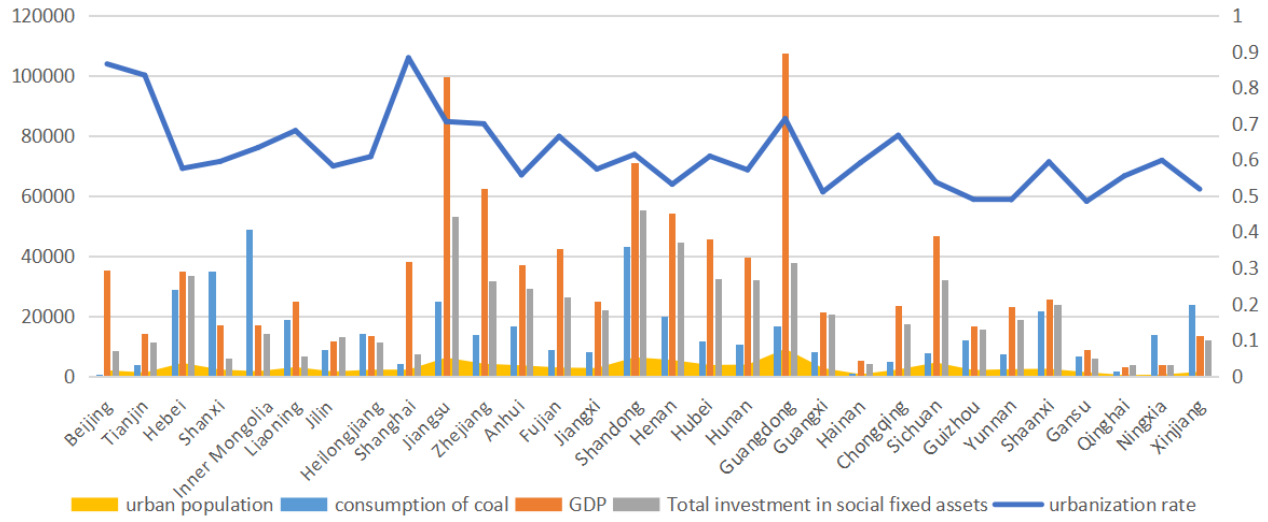


Fig. 1 Data descriptive chart

Table 1. Variables and definition

Variable	Description	Definition
y	Coal consumption	Total coal consumption of a certain province in 2019.
X1	Regional GDP	The final outcome of production activities for all permanent units in the region in 2019. The GDP of a region is equal to the sum of the added values of various industries.
X2	Investment in fixed assets	The total amount of work and related expenses incurred in the construction and purchase of fixed assets by the whole society in 2019, expressed in monetary form.
X3	Urban population	Population residing in cities and towns in 2019.
X4	Urbanization rate	Proportion of urban population in 2019.

2.3. Method

This study employs the cluster analysis method from the SPSS software to categorize 30 provinces. Cluster analysis is a quantitative method used to classify multi-factor entities, integrating contemporary taxonomy with multivariate analysis. The fundamental principle is to quantitatively determine the relationship between samples based on their own attributes, using mathematical methods according to certain similarity or difference indicators, and cluster the samples according to the degree of this relationship.

The basic idea of hierarchical clustering method is to initially treat each sample as a distinct cluster, then define distances between individual samples and among clusters, and then pair and merge the two samples or clusters that are closest in distance, recalculating the distances between this newly formed cluster and the others. This iterative merging based on minimal distance continues and successively reducing the number of clusters until all samples ultimately amalgamate into a single cluster. The advantage of hierarchical clustering is that it clusters based on the principle of shortest distance among samples. This systematic classification process is related to the specified classification index and the specific clustering method employed. The entire clustering process can be visually represented through a clustering (tree) graph. In this study, hierarchical clustering is applied to categorize coal consumption across various provinces. Importantly, when employing hierarchical cluster analysis, there is no need to weigh individual metrics; instead, an aggregate score based on evaluation criteria suffices.

3. Results and Discussion

3.1. Results of cluster

Systematically cluster the standardized indicators using the inter group connection method and select Euclidean distance squared as the measurement interval. The clustering results are showed in the cluster tree graph presented in Figure 2. It indicates setting the relative distance threshold at 6 allows the coal consumption of the 30 provinces to be broadly categorized into six clusters. Cluster 1 consists of Guangxi, Yunnan, Guizhou, Jiangxi, Shaanxi, Xinjiang, Jilin, Heilongjiang, Ningxia, Hainan, Qinghai, Gansu, Fujian, Chongqing, and Liaoning. Cluster 2 consists of Hebei, Henan, Hubei, Hunan, Anhui, Sichuan, and Zhejiang. Cluster 3 consists of Shanxi and Inner Mongolia. Cluster 4 consists of Beijing, Shanghai, and Tianjin. Cluster 5 contains Jiangsu and Guangdong. Cluster 6 solely contains Shandong.

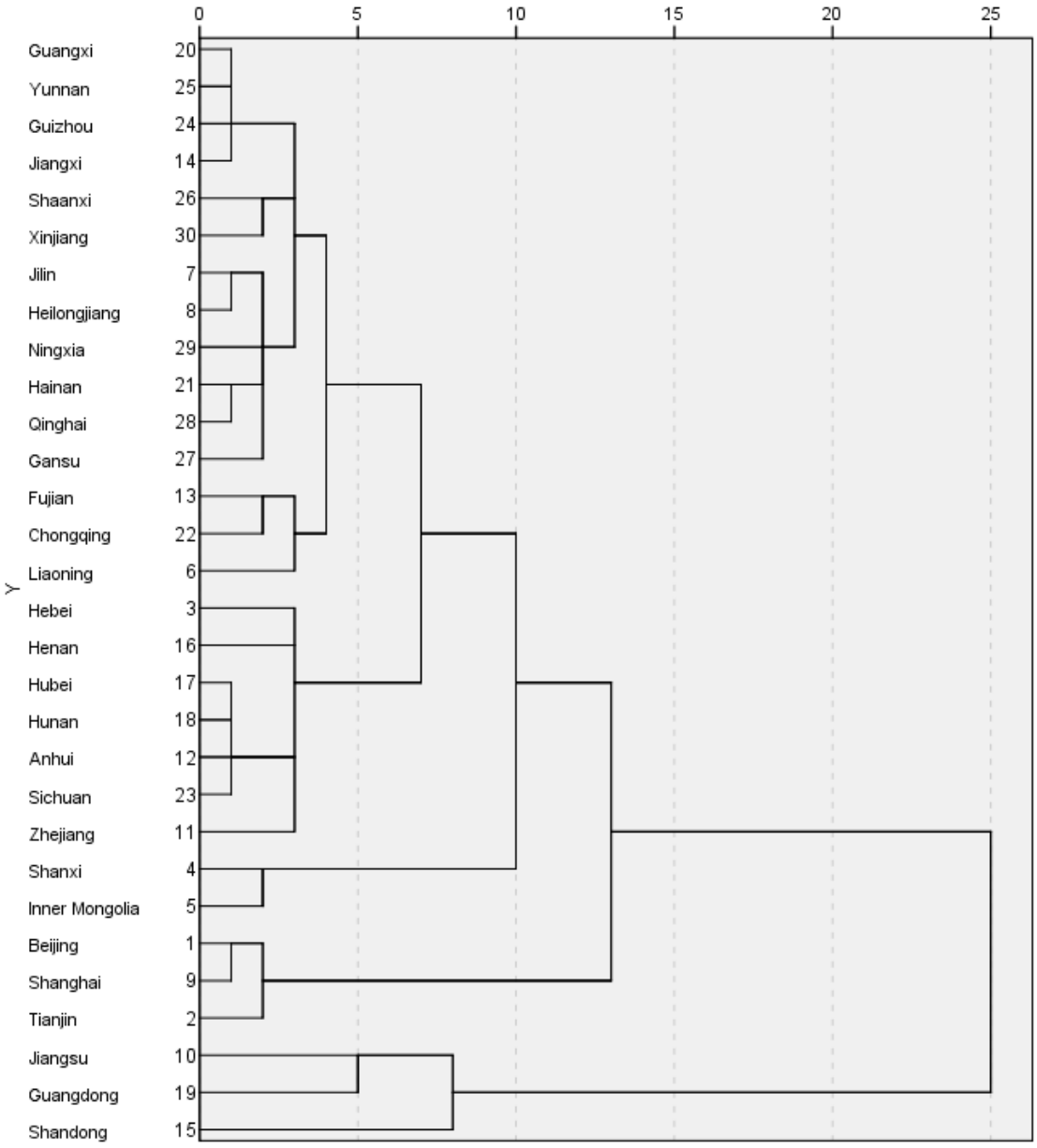


Fig. 2 Genealogy based on average joins (between groups)

The arithmetic mean of coal consumption and urbanization rate for these six clusters are computed and compared in Table 2. The arithmetic mean of coal consumption for Clusters 1 to 6 are 106.36, 156.15, 419.71, 27.29, 208.68 and 431.33 million tons respectively. Correspondingly, the arithmetic mean of urbanization rate for these clusters are 0.57, 0.58, 0.61, 0.86, 0.71, and 0.62 respectively.

Table 2. Provinces, average coal consumption, and average urbanization rate for each cluster

Cluster	Provinces	Coal consumption (million tons)	Urbanization rate (%)
1	Liaoning, Jilin, Heilongjiang Fujian, Jiangxi, Guangxi, Hainan, Chongqing, Guizhou, Yunnan, Shaanxi, Gansu, Qinghai, Ningxia, Xinjiang	106.3583	57
2	Hebei, Zhejiang, Anhui, Henan, Hubei, Hunan, Sichuan	156.1519	58
3	Shanxi, Inner Mongolia	419.7124	61
4	Beijing, Tianjin, Shanghai	27.2906	86
5	Jiangsu, Guangdong	208.6796	71
6	Shandong	431.3300	62s

Further, this study continues to visually analyze the relationship between coal consumption and urbanization rate. The findings are illustrated in Figure 3. As the urbanization rate gradually escalates, coal consumption displays a trend of initial growth followed by a decline. Specifically, when the level of urbanization is low, there's a positive correlation between urbanization and coal consumption, as evident from Clusters 1, 2, 3, and 6. However, once the urbanization level reaches a certain threshold, its relationship with coal consumption becomes negatively correlated, which is evident from Clusters 4 and 5.

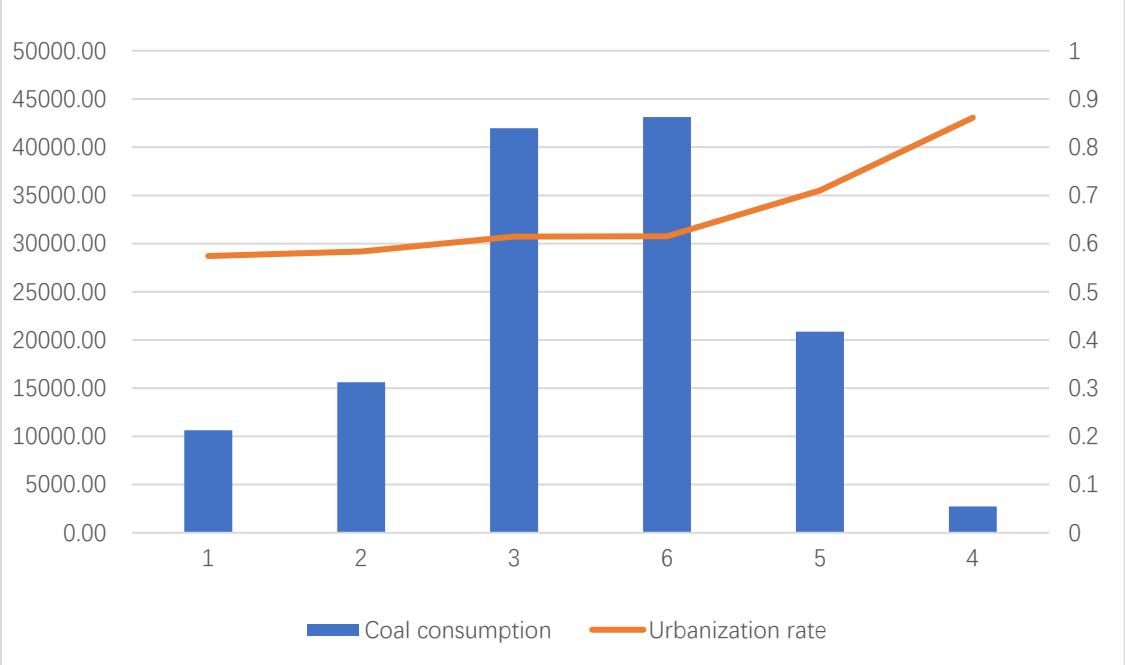


Fig. 3 Combined plot of coal consumption y and urbanization rate X4

3.2. Discussion

From the results, there's a positive correlation between coal consumption and the level of urbanization. Specifically, the population redistribution induced by urban growth leads to shifts in the industrial structure. On one hand, urban development needs to develop infrastructure construction and increase the demand for raw material production. This coincides with the rise of the secondary industry in the local area, which has the highest coal consumption among the three major industrial categories. Consequently, coal consumption will inevitably increase. On the other hand, as urbanization proceeding, there's a transition from rural to urban populations. This transition gradually expands urban consumption needs and indirectly increases coal consumption[3].

However, once urbanization surpasses a certain threshold, a negative correlation between coal consumption and urbanization levels occurs. High urbanization levels often signify increased income levels of residents, leading to an evident trend of consumption upgrade. This will promote the adjustment of China's industrial structure, increasing the proportion of the tertiary industry and reducing the proportion of the primary and secondary industries. Since coal consumption is mainly contingent on the growth of the secondary industry, any industrial restructuring and optimization could either reduce coal consumption or decease its growth rate [8,9]. Coal constitutes a significant portion of China's secondary industry, so a decrease in its proportion combined with an increase in the tertiary sector, will reduce overall coal consumption. Furthermore, with coal's low energy efficiency and significant pollution footprint, just like many countries worldwide, optimizing the energy composition and transitioning the mode of energy development has become an essential energy strategy for China at this stage[10]. Hence, with policy backing, developed cities in China are transitioning their energy infrastructures from coal-centric sources to cleaner and renewable energies.

4. Conclusions

This paper uses cluster analysis to study the relationship between coal consumption and urbanization rate in China in 2019. There are two main conclusions obtained, one is that the 30 provinces are divided into six categories using cluster analysis. For example, the first category is Guangxi, Yunnan, Guizhou, Jiangxi, Shaanxi, Xinjiang, Jilin, Heilongjiang, and so on, which are mostly in the northeast, southwest, south China, and northwest. These areas are mostly more remote and are in a state of development. The second category is Hebei, Henan, Hubei, Hunan, Anhui, Sichuan and Zhejiang, mostly in central and eastern China. The third category is Shanxi and Inner Mongolia. This category is located in North China, which not only has many coal enterprises, but also is the two largest coal-producing provinces. The fourth category is Beijing, Shanghai and Tianjin, which are developed coastal cities with high urbanization rates. The fifth category is Jiangsu and Guangdong, which are also developed coastal provinces. The sixth category is Shandong, a major industrial province and China's largest consumer of coal and electricity. Secondly, the study finds that when the level of urbanization is low, the higher the level of urbanization, the higher the coal consumption; when the level of urbanization is high, the higher the level of urbanization, the lower the coal consumption instead.

The paper provides two research implications. First, it is needed to focus on the reasonable allocation of social investment. In the case of China, excessive and unreasonable investment in fixed assets has caused an imbalance in the economy and China once suffered severe losses in the coal and oil industries. Thus, the industry needs to be optimized and achieve long-term development through innovation to prevent that kind of outcome. Secondly, it should promote industrial structure adjustment. There are two ways to achieve the rationalization of industrial structure. One is to transfer to high-tech industries. The other is the development of new equipment and technology to improve production efficiency and energy utilization. At the same time, it is essentially to actively develop clean energy to achieve "carbon emission reduction". There is also the need to focus on matching the employment of people with industrial development, reasonably allocating resources, reducing coal consumption and forming a healthy development cycle of economic growth and green ecology.

However, there are still some shortcomings in this paper. For example, due to data availability constraints, only data of year 2019 were selected for this paper, and the latest data were not available for research. In addition, only coal consumption, gross domestic product (GDP), total investment in fixed assets, urban population and urbanization rate were selected as indicators for this paper. Other indicators that may affect the conclusions, which need to be further explored in future research.

Authors Contribution

All the authors contributed equally and their names were listed in alphabetical order.

References

[1] Liu Shumao. Analysis of Coal Consumption Change in Shanxi Province. Energy and Energy Conservation, 2019, (11): 22-24.

[2] Zhang Sheng, Wang Jinying. Nonlinear Impact of Economic Growth on Carbon Emissions in China: An Empirical Study based on PSTR Model. Techno-economics and Management Research, 2022(11): 27-32.

[3] Wang Bofeng, Wang Huiwen. Research on the Dynamic Relationship between Coal Consumption, Urbanization and Industrial Structure in Midwest China. Xinjiang Finance and Economics, 2022 (01): 28-36.

[4] Cao Zhengxu, Zhang Xixi, Zhang Renjie, et al. Analysis on the Evolution of Coal Consumption Gravity and Its Decoupling Effect with Economic Growth in "2+26" Cities. Ecological Economy, 2023,39 (08): 86-94.

[5] Wang Fenyun, Su Yeqin. Energy consumption structure change and its influencing factors in Beijing-Tianjin-Hebei Region Area. Urban Issues, 2018(8): 59-67.

[6] Dong Jiefang, Wang Qiang, Wang Xingmin, et al. Empirical study on the influencing factors of coal consumption in Shanxi province. Journal of University of Chinese Academy of Sciences, 2016, 33(05): 641-650.

[7] Ding Rijia, Zhang Yibing. Comparative Analysis of the Impact of Coal Consumption in the Beijing Tianjin Hebei and Yangtze River Delta Regions: An Empirical Test Based on Panel Data. Enterprise Economy, 2017(2): 109-114.

[8] Yu Ji. Empirical Analysis of the Impact of Industrial Structure Change on Urbanization. Shanghai Economic Research, 2017(4): 11-16.

[9] Lei Qiang. Empirical Analysis of the Effect of the Industrial Structure on Coal Consumption Based Time - varying State Space Model. Industrial Technology and Economics, 2015,34 (05): 73-79.

[10] Wei Baolai. Identification and Prediction of Sudden Changes in Coal Consumption Demand in China. Shaanxi Normal University, 2017.