LUND UNIVERSITY
School of Economics and Management

Master program in Economic Growth,

Innovation and Spatial Dynamics

Chinese Energy Efficiency Analysis Based on DEA Model and the Influencing Factors---from a Regional Approach

Lu He

he.lu.364@student.lu.se

Abstract: This paper first measures the energy efficiency of 28 provinces from 2001-2009 with DEA model. Different from previous studies that include labor, capital and energy consumption as input indicators, GDP as the only desirable output, this paper also models the undesirable output by introducing environmental pollution index. The result shows that eastern area has highest energy efficiency while the western area has the lowest energy efficiency. The energy efficiency in all areas declined during the whole period and central area shares the highest energy input redundancy ratio. Then the econometric approach with fixed effect model is introduced to analyze the factors influencing energy efficiency. The result indicates that only producer price index (as the proxy to the energy price) has significant effect on the energy efficiency.

Key words ： Energy efficiency, DEA model, influencing factors,
Fixed effect model

EKHR72
Master thesis, second year (15 credits ECTS)
June 2012
Supervisor: Lennart Schön
Examiner: Håkan Lobell

Table of content

List of figures

List of Tables

1. Introduction

1.1 Background

Energy is the material base that human beings depend on. Without the utilization of highly-quality energy and advanced energy technology, economic growth cannot be achieved. However, due to the scarcity of energy resources and the global warming caused by burning fossil fuels, economic growth is also constrained by energy consumption. After the oil crisis in 1970s, the world oil price has been increasing dramatically, and many countries began to recognize how effectively the energy needs to be used to reach higher efficiency. In the late 1980s, the improvement of energy efficiency became another crucial step for many countries in reducing greenhouse gas emissions.

The improvement of energy efficiency seems more urgent in developing economies. The last two decades saw China's remarkable economic growth, of which the annual growth rate is over 9%. Thus, China has significant influence on the global energy profile and it is assumed that the influence will keep increasing in the future (Chaoqing, Sifeng, Zhigeng, & Naiming, 2010), however, as one of the biggest developing countries in the world, the energy use in China faces various problems such as limited supply, environmental pollution , relatively lower efficiency and substantial regional disparities.

Firstly, China is the second largest energy consumption country in the world, whose the energy consumption has increased from 0.57 to 3.35 billion ton coal equivalents (about 5.6% annual growth rate) since1978 (Ma, Oxley, & Gibson, 2010). However, the energy supply in China is severely limited, For example, there is one –third shortage of electricity in China in 2010 (Polsenberg, 2004). Secondly, China releases the largest amount of CO_2 in the world and the worsening environment pollution accounts for 5.8 % annual loss in total GDP (Yuan, Liu, & Wu, 2010). Thirdly, the energy efficiency is much lower than those of developed economies, IEA (2010) calculated the energy intensity in different countries, and it showed that the energy intensity is three times higher than the average level, nine times higher than Japan. The cause of low energy efficiency is due to the coal-based energy consumption structure, of which over 70% of the consumption relying on coal and relative laggard energy technology. Lastly, since Chinese government implements an economic strategy that favors eastern area, there is a substantial regional disparity in terms of energy use and energy efficiency. It is becoming increasingly clear that the unbalanced development is challenging the sustained growth in China.

1.2 Research focus

The contradictions between economic growth and energy consumption indicate that the improvement of energy efficiency is the key issue in addressing the problems. However, the significant regional disparity requires that the energy efficiency analysis focuses on the regional level. Recently, there is growing research concerning this issue. However, there is no uniform approach on evaluating energy efficiency. Some studies use energy intensity to indicate energy efficiency and some studies measure energy efficiency by multiple inputs-outputs indicators. Among them, data envelopment analysis (DEA) is the most widely used model in measuring energy efficiency with Chinese data, where most studies follow the way of Hu &Wang (2006) that measured the total-factor energy efficiency[1] based on DEA model. However, it is argued that the there is limitation with this approach since the undesirable output (like environmental pollution) is not take into account under multiple input-output model. Although Wang, Yu, & Zhang (2012), Shi, Bi & Wang (2010) added the pollution effect by modeling

[1]The total factor energy efficiency includes the labor and capital as two none-energy input indicators and together with energy input to produce a certain amount of GDP.

the undesirable outputs, they only considered the effect of green house gas emission and ignored other pollution effects; for example, the water pollution. What's more, most studies left the result that only telling the disparities of energy efficiency among regions, there is a lack of research on analyzing the factors influencing energy efficiency based on the efficiency disparities. Thus, in this thesis, the regional energy efficiency analysis will be explored with two major problems: ***how does the regional energy efficiency vary among regions over time and what are the factors influencing the energy efficiency.*** First I am going to measure the regional energy efficiency under DEA model with more comprehensive considerations on environmental effects: the measurement of environmental index will not only consider the waste gas, but also waste water and residues; second I will take a further step on the factors influencing energy efficiency through regression analysis. Now the improvement of energy efficiency is the major task of energy policy in China and the 12th five year plan has set the target of reducing 50% CO_2 by 2020 comparing with the level in 2005 .The regional energy efficiency analysis can offer some evidences on the regional energy disparities and shed some lights on how to improve energy efficiency through better use and allocation of energy resources.

1.3 Outline of thesis

The paper proceeds as follows: section 2 reviews the previous literature that focus on the discussion on energy efficiency definition; the different measurements in evaluating energy efficiency, which divides the measurements into single input-output ratio indicators and DEA model; the relationship between energy consumption, energy efficiency and economic growth; the factors that influence the energy efficiency. Section 3 is the empirical study of regional energy efficiency among 28 main provinces from 2001 to 2009. The analysis is based on DEA model, which includes capital, labor and energy consumption as input indicators, GDP and environmental pollution index as desirable and undesirable output indicators. The regional disparities among eastern, central and western area are presented. Section 4 is the further step in analyzing the factors that influence regional energy efficiency with longitudinal data, where the fixed effect model is introduced and measures the influence from degree of openness (the value-added of import and export in each region), technology development (numbers of patents), economic scale (the share of value added in big and medium enterprise), structural change (the share of value added in industry), energy price (the producer price index in each region). Section 5 is discussion and conclusion.

2. Literature review

2.1 The definition of energy efficiency

Energy efficiency is a kind of generic term and there is no uniform quantitative measure of energy efficiency. Economist, engineers and environmentalists may have different definitions on energy efficiency and they choose various indicators to measure energy efficiency. In general, energy efficiency can be simply expressed as a ratio between useful output of a process and energy input into a process:

$$\frac{\textbf{Useful output of a process}}{\textbf{Energy input into a process}}$$

The increase of energy efficiency means using less energy input to produce more useful output (Patterson, 1996). Besseboeuf (1997) explained the efficiency in terms of economic and technical perspective, the increase of economic energy efficiency means getting more valued products and services with the same amount of energy input, while the technical energy efficiency means the development of technology and management which allows less energy consumption in the production process.

2.2 The single input-output efficiency assessment VS multiple input-output efficiency measurement

The energy efficiency varies in different fields based on how to define the "energy input" and "useful output", thus sets of indicators are counted to serve a certain purpose. The previous researches show that the choice of different indicators will have significant impact on the results (Goldemberg & Siqueira Prado, 2011; Patterson, 1996). Thus, how to choose appropriate indicators becomes a crucial issue in efficiency analysis, which also has gained the increasing concerns by energy strategy researchers. The international Energy Agency (IEA) has been exploring the energy efficiency indicators since 1995 and now it is one of six broad focus areas of IEA's energy analysis (IEA, 2011). The energy efficiency indicators are also the subject of recent report by Chinese energy research center (Zhang, Li, Mu, & Ning, 2011). However, despite the various indicator selections, the calculations of energy can be also very different. The assessment can be divided into two categories in terms of the number of input and output indicators.

2.2.1 Single input-output efficiency measurement

The single input-output efficiency measurement is just a simple ratio that directly compares the useful output and energy input. Patterson (1996), Ang (2006), John (2004) and Boyd (2005) thought the energy efficiency can be measured in terms of thermodynamic indicators, physical-thermodynamic indicators, economic-thermodynamic indicators and pure economic indicators. Thermodynamic indicators measures the energy efficiency entirely on the science of thermodynamics of which the indicators are simple ratio, and the energy efficiency are measured in an "ideal" process. Physical-thermodynamic indicator is a kind of hybrid indicators. It measures energy input in thermodynamics units, but output is measured in physical units, which can reflect the end use service that the consumer require. The economic-thermodynamics indicators, however, are another kind of hybrid indicators, which still measure energy input with thermodynamics units but the output is measured by its market value. The economic indicators measure the both inputs and outputs in their related market value: energy input ($)/energy output ($). It is argued that given the energy price to the input instead of thermodynamic units can solve the energy quality problem (Berndt, 1975).

It is argued that the economic-thermodynamics indicator is widely used in the energy strategies studies, from which the energy-GDP ratio is known as the most sophisticated indictor that measures the energy intensity at most aggregated level. The formula

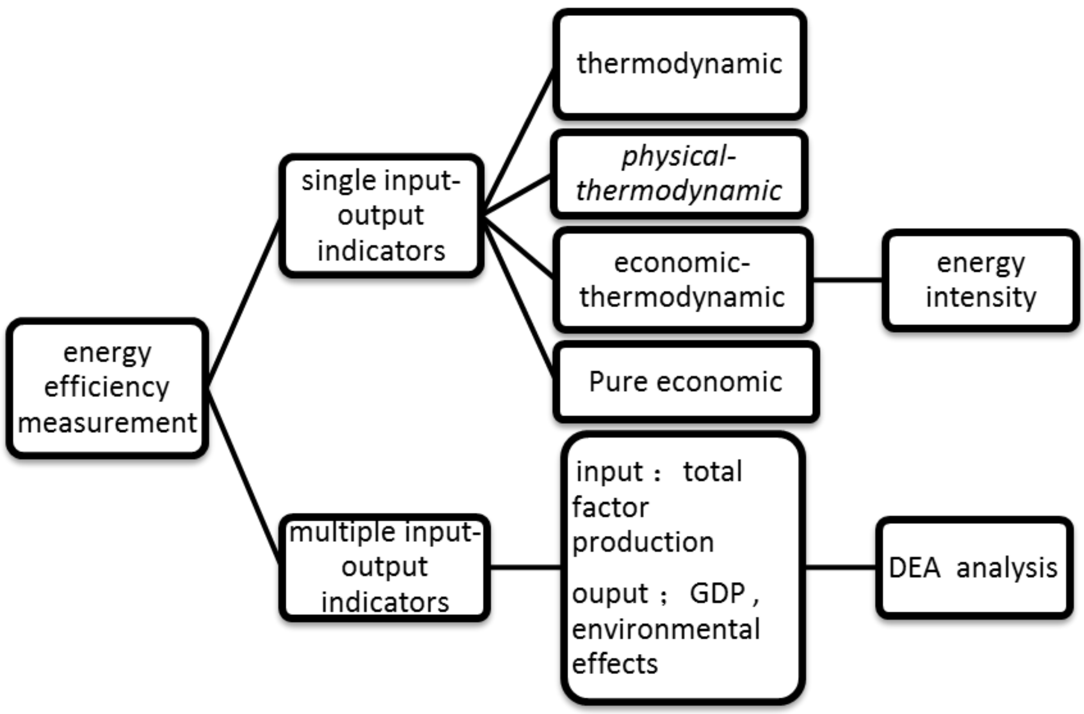
I=E/Y measures the amount of energy used (E) in producing on unit of GDP(Y). Thus, its inverse is used as a measure of energy efficiency. The lower intensity promises the higher energy efficiency and vice versa (Silveria & Luken, 2008). Due to its simplicity and understandability for policy makers, the energy-GDP ratios are often used to make across national comparison and it tells in the long run whether the economy is more energy insensitive or not. IEA and World Bank use energy intensity to evaluate the energy efficiency across countries and it shows there is a downward trend of energy intensities in the global perspective (Goldemberg & Siqueira Prado, 2011).

It is argued that GDP is an aggregate indicator which summarizes diverse activities. However, the energy intensities in different activities may differ widely and the change is mixed. Thus index decomposition analysis (IDA) approaches are developed to analyze the impact of energy intensity change (Ang, 2006).

2.2.2 Multiple input-output efficiency measurement

The single input-output efficiency measurement only presents a simple relationship between energy input and output, however, energy alone cannot produce any output and it must integrate with other input factors to produce output (Zhou & Ang, 2008). A nonparametric methodology—data envelopment analysis offers an ideal solution to the multiple inputs-outputs efficiency problem. Based on this concept, Hu &Wang (2006) developed total-factor production energy efficiency indicators that introduced labor and capital as other two none-energy input indicators. There are quite a few studies that follow the Hu and Wang's study. Xu & Liu (2008) analyzed the energy efficiency in China with eight economics zones and it indicated that the energy decreased gradually from southeast to northwest between 1998 and 2005. Sueyoshi & Goto (2011) studied the energy efficiency in Japan and indicated a U shape curve which is similar to environment Kuznets Curve (EKC) that explains the relation between energy efficiency and income per capita. Chein & Hu (2007) computed the renewable energy and macroeconomic efficiency of OECD and none-OECD countries and the results showed that OECD countries had general higher efficiency, but none-OECD countries shares higher growth rate on energy efficiency. However, many scholars argued that there is limitation on Hu & Wang (2006) since the GDP is taken as the only output indicator. As Shi (2010) argued, the production process not only produces desirable economic output, like GDP, but also generates undesirable output, like CO_2 emission. Thus the environmental effect should be taken into consideration when measuring energy efficiency. Without considering the undesirable output, the result cannot provide an appropriate benchmark and comparison, and the result of energy efficiency will be exaggerated (Banker & Morey, 1986; Zhou & Ang, 2008). In summary, an ideal multiple input-output method based on DEA model should include non-energy inputs, like labor and capital as input indicators and both desirable and undesirable output as output indicators.

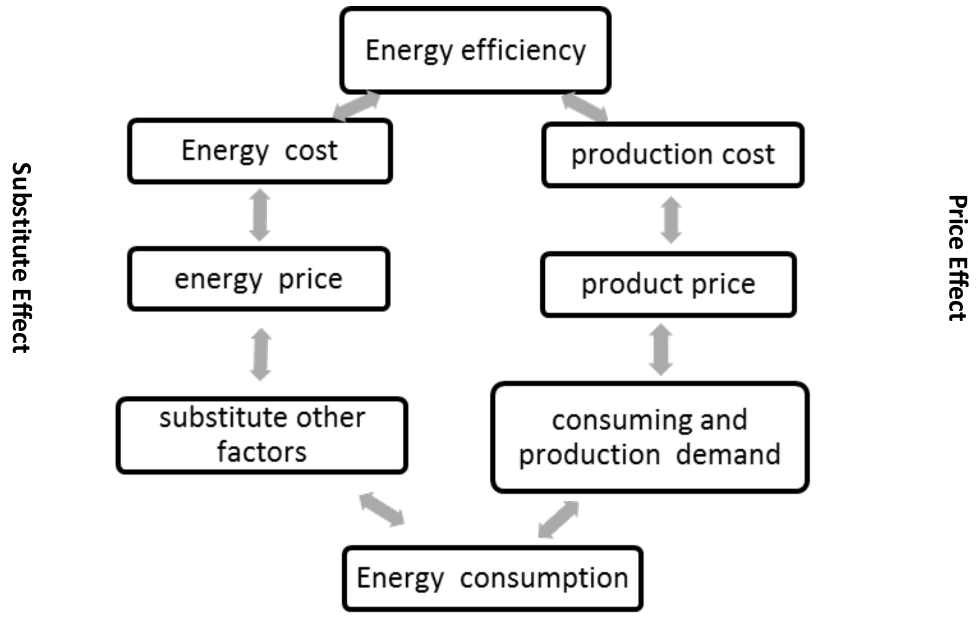
Figure 1: Different assessments of energy efficiency based on the number of indicators



2.3 The energy efficiency and economic growth

Energy is an essential input for the economic growth as well as labor and capital. In the modern history economic booms always involve keen demand on energy. However, the economic growth is also constrained by energy consumption. Arbex & Perobelli (2010) argued that the limited natural resources infer severe strain on growth which can eliminate the positive effects from technology improvements. The shortage of energy will increase the energy prices and the dependency on import, which threatens energy security. Serious environmental problem will also emerge and discourage the sustained economic growth. The increase of energy efficiency is an effective way to suppress energy consumption, however, more and more researches suggested there is rebound effect between energy efficiency and consumption, which indicates that the increased energy efficiency will also stimulate the growth on energy consumption (Ayres, Turton, & Casten, 2007). The energy intensity in China has decreased by 33% since 1978, however, the energy consumption have been growing much faster than the decrease of intensity (Crompton & Wu, 2005). Previous research divides rebound effects into substitution effects and price effects. The increase of energy efficiency will lower the cost of energy consumption and the energy price will drop. Thus, the possibility that energy substitutes other production factor, such as labor and capital will increase, and the energy consumption will be expanded. Secondly, the increased energy efficiency implies higher level of output, the cost of production and the price of product will decrease, which in turn will facilitate the energy consumption and production (Madlener & Alcott, 2009). Therefore, an ideal economic growth model is to increase productivity and decrease energy consumption simultaneously and Howarth (1997) argued that the energy efficiency will not increase the energy use unless the "the energy costs dominate the total cost of energy service and the expenditures on energy service constitute a large share of economic activity."

Figure 2: Rebound effect of energy efficiency on energy consumption



2.4 Factors that influence energy efficiency

It is necessary to understand the cause of energy efficiency change that shapes the energy policy guidance to achieve higher energy efficiency. The previous researches indicate that the causes are divers and the dominate factor varies among different economies. Jan (2008) inferred that the structural change that shift from high energy-intensive sectors to low energy-intensive sectors and the technology change are the main reasons lowering the energy intensity, of which the structure change had stronger influence in before 1980s and technology became the dominated factor after 1890s. Eyre (1998) analyzed the energy efficiency in UK and suggested the liberalization of energy market is an effective method in increasing energy efficiency. Sehieich (2004) found that the price of energy is an important reason that drove up energy efficiency. In the Chinese energy efficiency research, Shi (2002) and Fan (2007) indicated that the Open and Reform in 1980s, which introduced the market economy and competition mechanism, have significant influence in increasing energy efficiency. They studied the aggregated energy efficiency between 1980 and 2008, and found that the technology change is the driving force for energy efficiency but the relative lagging technology is one of the constrained factors in gaining higher energy efficiency in China. They also found that the structure change only had positive effect before 1998 and it had negative effect on energy efficiency after 1998. Wu (2010) thought that the scale of production has direct influence on the energy efficiency since under the economies of scale the long run cost will decline. Based on previous studies, however, this thesis summarizes five main factors that will influence energy efficiency.

1) The openness of market

The openness in the market perfects the market mechanism and creates the healthy competition environment (Jensen & Tarr, 2003). The enterprises with low energy efficiency will commit themselves to increasing energy efficiency. In addition, the liberalization of the market, especially the international trade will strengthen the spillover effects of technology and more enterprises get access to advanced technology and management skills to increase the energy efficiency

2) Price of energy

The rise of energy price will improve the energy efficiency (Birol & Keppler, 2000). This is because that the bumping-up energy price implies that the cost of production and consumption will increase, which will raise the awareness of the reduction of energy waste and the effort in increasing energy efficiency. What's more, as nonrenewable energy resources, the increase price of fossil fuels (like coal and oil) will facilitate the development renewable energy resources with higher efficiency to substitute the traditional energy resources (Chwieduk, 1997).

3) Structural change

The reason that structure change will influence the energy efficiency is because different sectors and industries have different energy intensities. Thus, the energy efficiency will improve if the share of high energy-intensive industries or sectors shifts to low energy intensive industries or sectors (Howarth, Schipper, Duerr, & Strøm, 1991). Structural change can take place either among sectors or within sectors. Assuming the whole structure is made up by three main sectors, agriculture, industry and service, of which the service has the lowest energy intensity and industry has highest intensity, if the share shift of industry shift to service, the energy efficiency will increase. Within the sectors, increasing energy efficiency in a specific highly energy-consuming industry will also increase the energy efficiency in industry.

4) Technology development

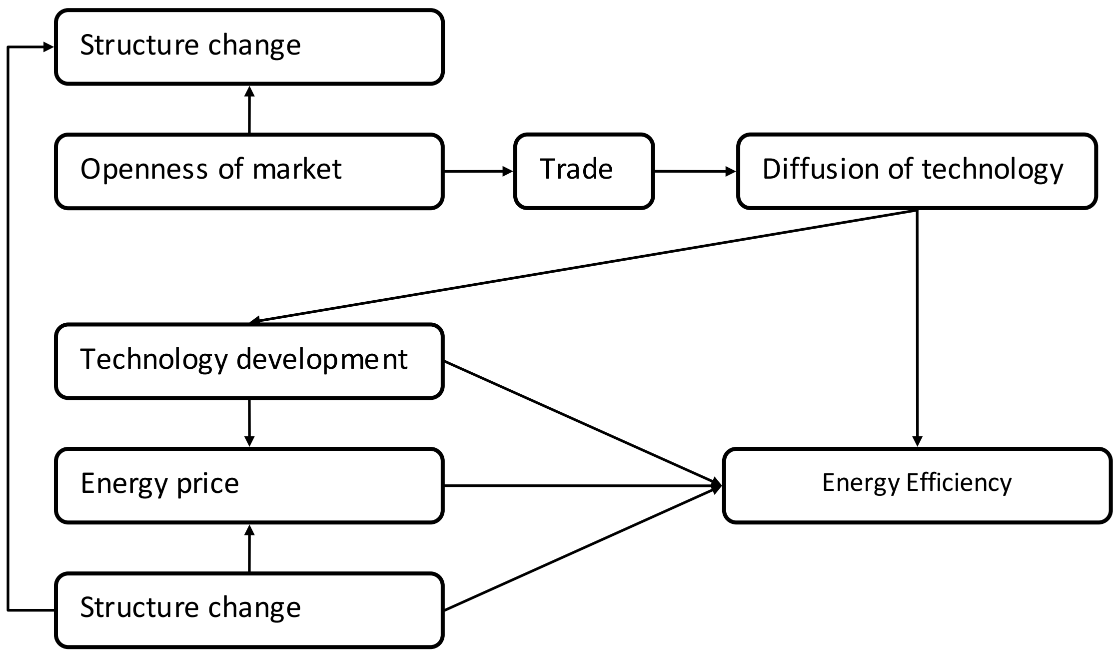
The improvement of technology is closely linked with the energy efficiency level. Firstly, new technology introduces effective manufacturing equipments that meliorate the production process and reducing the energy consumption. Secondly, advanced technology also promotes the development of renewable energy, like solar power and winder power, of which has higher efficiency and less environmental side effect.

5) Scale

Whether the production can be implemented under optimized scale will influence the energy efficiency directly (Trinh 2012). The economies of scale refer to the cost advantage that an enterprise obtains due to expansion. Enterprises with relative small scale will have higher cost in producing per unit of product and the energy efficiency is lower. However, the blinding investment in scale expansion will not increase the efficiency since return to scale can be either increased or decreased. The increase return to scale means by doubling the amount of input we can produce more than double of the output while under the decrease return to scale will generate less than double of the output.

As figure 3 shows, the five factors are integrated together that will not only influence the energy efficiency, but also have certain impact on others factors. For example, the openness of market will stimulates more trades and the diffusion of technology will promote the technology development. With the development of technology that induces more productive working machine and the large scale production, the cost of production will reduce and the energy price will also drop.

Figure 3: Factors influence energy efficiency



3. The measurement of energy efficiency based on DEA model

Data envelopment analysis is a non-parameter approach that evaluates the relative efficiency for a set of comparable entities with multiple inputs and outputs. The entities are called decision making unites (DMUs), and the envelopment frontier allows all the DMUs lying on or below the frontier. DMU that lies on the frontier is regarded as the efficient DMU, of which the value is 1. Thus, the efficient DMU functions as the benchmark that compares the distance between efficient DMUs and inefficient DMUs. The advantage of DEA is conspicuous: first, it doesn't require an indentified the relationship between input and output, which is useful to address problem if the relationship is not specified by concepts in some cases. Secondly, the DEA is an ideal approach to evaluate the efficiency with multiple inputs and outputs; the only information required is the physical quantities of input and output. Thirdly, it also suggests a way on how to transfer the inefficient DMUs towards efficient DMUs by slack adjustment.

3.1.1 The CCR model

CCR is the first DEA model introduced by Charnes, Copper & Rhodes (1978), even though various DEA models have been developed afterwards, the CCR model is still the basic and most widely applied model in the research, a brief explanation is listed below.

Assume there are n DMUs , each DMU has w inputs X_w and q outputs Y_q. Thus the input and output of DMU_j can be expressed as

$$X_j = (X_{1j}\ X_{2j},\ldots, X_{wj})^T > 0 \quad Y_j = (Y_{1j}\ Y_{2j},\ldots Y_{qj})^T > 0$$

The efficiency score of DMU_j is given by solving the following fractional programming model

$\text{Max}\frac{u^T y_0}{V^T y_0}$

s.t $\frac{u^T y_0}{v^T x_0} \leq 1, j = 1,2, \dots, n$

$u \geq 0$, $v \geq 0$ (1)

where $u=(u_1, u_2, \dots, u_w)^T$ and $v=(v_1, v_2, \dots, v_q)^T$ means the weights of input X_w and input Y_q.
This formula can transformed into the a liner program

$\text{Max}\mu^t y_0$

s.t $\omega^T x_j - \mu^t y_j \geq 0, j = 1,2, \dots, n$

$\omega^T x_0 = 1$

$u \geq 0$, $v \geq 0$ (2)

By duality, (2) is equivalent to linear programming model (3)

$\text{Min}\theta$

s.t $\sum_{j=1}^{n} \lambda_j x_j + s_i^- = \theta x_0$. i=1,2,...,w

$\sum_{j=1}^{n} \lambda_j y_j - s_r^+ = y_0$. r=1, 2…, q

$\lambda_j \geq 0$, j=1, 2…, n $s_i^- \geq 0$, $s_r^+ \geq 0$ (3)

Where θ is the value of overall efficiency; λ_j is the convex coefficient; s_i^- and s_r^+ are the slacks of input and output. They are also called input redundancy and output deficiency. The model (3) assumes two rules:

1), when θ=1 and s_i^- , s_r^+ =0, the DMU is efficient; 2), when $\theta \neq 1$ and s_i^-, $s_r^+ \neq 0$,

the DMU is inefficient. However, the level of slacks tells amount of input redundancy and output deficiency and by adjusting the slacks, the inefficient DMU can transfer into efficient DMU. According to the efficiency concept, the inefficient DMU can be achieved by either input-orientated (reduce the amount of input s_i^- *to* produce same amount of output) or output- orientated (increase the amount of output s_r^+ while keep input constant). Since in the production process, the control of input is more predictable than output, in this thesis, the model is based on the input orientated CCR model.

3.1.2 Modeling undesirable output

Even though the traditional DEA model can deal with multiple inputs and outputs that have different units. It doesn't distinguish the difference between desirable output and undesirable output and usually all the outputs are regarded as desirable. Keeping the input constant, the higher level output means higher efficiency. When it comes to the undesirable output, however, the amount should be reduced to increase the efficiency. In the energy perspective, besides the desirable output, it also produces undesirable pollution. For example, the burning of fossil fuels will generate wastes gas, CO_2 and SO_2. Thus, the desirable output and undesirable output should be treated differently in the production process. There are several ways in modeling the undesirable output. First, Faere, Grosskopf, Lovell, & Pasurka (1989) introduced a non-linear DEA program to model the paper production system by increasing the desirable output and decreasing undesirable output. Based on this model, Seiford & Zhu (2002) transformed this to a liner program. They first multiply the undesirable output by -1 and then find a proper translation vector $\boldsymbol{w}$ to let all undesirable output become positive. However, they didn't define the condition for a proper vector, and the results will be biased due to choosing different vector. Shi, Bi, & Wang (2010) thought that the most pollution is caused by using energy input. By reducing the energy consumption, the undesirable output also decreased; so they treat the undesirable energy output as input. However, it cannot reflect the true production process (Seiford & Zhu, 2002). Thus, it is argued perhaps the easiest and effective way is to apply a decreasing transformation---take the reciprocal value of undesirable output and treat them as desirable output (Lozano, Villa, & Brännlund, 2009). The model is displayed below

s.t $\sum_{j=1}^{n} \lambda_j x_j + s_i^- = \theta x_0.$ i=1,2,..,w

$\sum_{j=1}^{n} \lambda_j y_j - s_r^+ = y_0.$ r=1, 2,...,q

$\sum_{j=1}^{n} \lambda_j \frac{1}{y_j} - s_t^+ = y_0.$ t=1, 2,...,k

$\lambda_j \geq 0,$ j=1,2,…,n

$s_i^- \geq 0$, $s_r^+ \geq 0$

By using the computer program DEAP 2.1 with designed input-orientated CCR model, we can get the direct result of efficiency score.

3.2 Data

3.2.1 Description of regions

This thesis examines 28 regions in the provincial level in the mainland of China excludes Tibet, Hainan and Guangxi due to absent of complete data. In Chinese political and development perspective, 28 regions can be divided into eastern area, central area and western area. The analysis covers the period from 2001-2009. And all the data are collected from the Chinese statistical year book 2002-2010 and Chinese environmental statistical year book 2002-2010.[2]

The eastern area contains 10 regions (Beijing, Tianjin, Hebei, Liaoning, Shanghai, Jiangsu, Zhejiang, Fujian, Shandong and Guangdong). All of them are coastal provinces except Beijing. Due to the institutional unbalanced strategies which favor the eastern area since Open and Reform in 1980s, this area is the most developed area and shares about half of total GDP in China. Most industries are light industries and international trade is flourishing there. The central area constitutes 9 regions (Shanxi, Inner Mongolia, Jilin, Heilongjiang, Anhui, Jiangxi, Henan, Hubei, and Hunan). In the inland area, there is a large population and it is known as the home that based for agriculture. In the northeast part (Heilongjiang, Jilin); quite a few heavy industries are located and Inner Mongolia, Shanxi are known as two biggest fossil fuel centers in China. The Western area contains 9 regions (Sichuan, Guizhou, Yunnan, Shanxi, Gansu, Qinghai, Ningxia, and Xinjiang). The economy is less developed compared to the eastern area, but still higher than western area. Western area, which is known as fewer pollution and large storage of natural resources, is still underdevelopment

---

[2] Sources: National Bureau of Statistics of China

(http://www.stats.gov.cn/tjsj/ndsj/)

(http://www.stats.gov.cn/tjsj/qtsj/hjtjzl/index.htm) compare to other areas. But the environmental situation is better than other area due to less setting up of industries. More detailed comparisons are listed in table 1.

3.2.2 The choice of indicators and the summary of statistics

Input indicators

In this thesis, the choice of indicators will follow the idea of Hu &Wang (2006). Thus, three input indicators will be the total energy consumption, the number of total employed persons and the capital stock. All the energy consumption are transformed into millions tons of coal equivalent (Mtce) to make proper comparison. Since the data of capital stock in regional level is not receivable from the statistical year book. I used the results by Xie & Pan (2011), where the capital stock are calculated based on 1952 prices.

Output indicators

GDP is the most common output indicator in the aggregated level, however, it is risky to take the current price because of the inflation will exaggerate the level of output, thus, the GDP are calculated in 2001 constant price. Since the energy production will produce both desirable and undesirable output, the pollution data are also taken into consideration. In China, the pollutions generally divided into three categories, the waste water discharged (1000tons); the waste gas emission (100 million cu.m) and the solid wastes generated (10000tons). However, the data on regional level are only achievable for the industry, which generates largest amounts of pollutions. In this thesis, the data in industry pollution are chosen to represent the undesirable output. One problem in this environmental data is that all the pollution are obtained with different statistical units and the one limitation of DEA is the DMUS should be more than five times larger than the total number of indicators. Thus the pollutions data are transformed into index by weighted average with the data in Beijing 2000 as the base. However, there is no standard on how to give weight to each item and which item of pollution is more important than the others. Thus I give the same weight (1/3) to each item.

Table 1 shows the summary of statistics of inputs and outputs. It indicates that during the period 2001-2009, the average capital stock has increased more than three times (from 4148.99 to 13074.40 billion RMB); the energy consumption has increased more than two times (from 5547.20 to 12461.71 Mtce). However, the change of employment is relatively small, from 2214.46 to 2553.50 millions of workers. The standard deviation also shows that the regional disparities are more significant in capital stock and energy consumption compared with employment. Similarly, all the outputs have increased and the average GDP has grown nearly three times, as well as the pollution. The standard deviation also shows that the regional disparities are more significant in capital stock and energy consumption. Comparing the total average input and output from 2001 to 2009 in the three areas, the eastern region has much higher level in all input and output indicators and the western area is at the lowest level. For example, the average capital stock in eastern area is 1255.45 billion RMB, which is four times bigger than western area; the GDP is five times bigger than western area. The summary of input and output indicators showed two main features: 1) at national level, both input and output have increased remarkably in nine years; 2) at regional level, there is a distinct disparity among regions, which indicates a rather unbalanced development in different areas. However, the high level of input and output is not convincible that eastern area also has higher energy efficiency. And it is also questionable how the energy efficiency changed over time across different regions. In the next section, a DEA approach will be applied to address this issue.

Table 1: Summary statistics of inputs and outputs

Year	Non-energy input			Energy Input	Desirable output	Undesirable output
	Variables	Labor	Capital Stock	Energy	GDP	Pollution
	Units	Million workers	RMB billion	Million tons of coal Equivalent(Mtce)	RMB billion	Index
2001	Mean	2144.46	4168.99	5547.20	3770.34	2.59
	Std. Dev.	1428.54	2724.36	3221.07	2849.64	1.65
	Min	240.32	589.98	915.60	300.13	0.27
	Max	5516.59	9888.62	13778.54	12039.25	6.78
2002	Mean	2229.70	4672.22	6059.02	4180.86	2.73
	Std. Dev.	1437.38	3072.40	3618.62	3192.10	1.67
	Min	247.30	685.48	1018.83	336.38	0.25
	Max	5522.00	11195.75	16149.70	13528.88	6.82
2003	Mean	2206.32	5315.55	6862.27	4698.89	2.89
	Std. Dev.	1440.58	3535.69	4130.46	3648.00	1.74
	Min	254.26	793.51	1122.70	376.28	0.28
	Max	5535.68	13053.09	18195.76	15537.22	6.65
	Mean	2255.67	6063.09	7967.79	5343.53	3.28

Year		Non-energy input		Energy Input	Desirable output	Undesirable output
	Variables	Labor	Capital Stock	Energy	GDP	Pollution
	Units	Million workers	RMB billion	Million tons of coal Equivalent(Mtce)	RMB billion	Index
2004	Std. Dev.	1466.36	4085.13	4810.87	4193.09	2.16
	Min	263.08	909.97	1364.38	422.45	0.34
	Max	5587.45	15220.13	21398.25	17835.02	9.31
	Mean	2314.51	7061.28	9095.16	6033.40	3.71
2005	Std. Dev.	1518.49	4785.05	5598.05	4766.60	2.46
	Min	267.62	1041.61	1670.21	473.87	0.46
	Max	5662.41	17909.68	25104.79	20296.34	9.65
	Mean	2371.95	8191.93	10037.58	6862.73	3.71
2006	Std. Dev.	1557.85	5541.86	6271.41	5471.61	2.46
	Min	271.95	1169.80	1903.22	531.68	0.46
	Max	5717.56	20845.10	28249.86	23258.16	9.65
	Mean	2429.38	9528.72	10998.11	7848.43	4.54
2007	Std. Dev.	1598.92	6387.84	6851.04	6272.40	3.00
	Min	276.29	1311.62	2094.89	598.14	0.72

Year	Non-energy input			Energy Input	Desirable output	Undesirable output
	Variables	Labor	Capital Stock	Energy	GDP	Pollution
	Units	Million workers	RMB billion	Million tons of coal Equivalent(Mtce)	RMB billion	Index
	Max	5772.72	24112.52	30596.00	26667.87	12.71
	Mean	2484.70	11082.52	11636.46	8766.85	4.71
2008	Std. Dev.	1631.08	7320.53	7188.08	6944.66	3.08
	Min	276.79	1463.21	2256.52	674.10	0.86
	Max	5835.45	27906.20	32225.23	29348.64	11.83
	Mean	2553.50	13075.40	12461.71	10418.37	4.95
2009	Std. Dev.	1673.53	8600.24	7484.75	8154.41	3.19
	Min	285.54	1680.18	2348.17	808.44	0.89
	Max	5948.78	33121.85	32420.24	34619.03	13.76
Eastern Area	Mean	2679.15	12455.45	12607.54	10492.78	5.13
Central Area	Mean	2545.03	6210.48	8686.80	5223.38	3.76
Western Area	Mean	1722.19	3857.18	5189.13	2640.87	2.10

Source: Chinese statistical year book 2002-2010, Chinese environmental statistical year book 2002-2010, National Bureau of Statistics of China

3.3 Result and Analysis

3.3.1 The overall energy efficiency in different regions

Table2 shows the results of overall energy efficiency in different regions from 2001 to 2009 under CCR model. It shows that there are two eastern regions (Shanghai, Guangdong) and one western region (Qinghai) keeping themselves on the efficiency frontier in all years, of which the value is 1. Beijing is inefficient only in 2003 and Tianjin has reached efficiency frontier since 2004. However, Fujian has strayed from efficiency frontier since 2005. There is no central region that has touched the energy frontier during the whole period. By comparing the average energy efficiency in different areas, the eastern area has the highest energy efficiency in all years and the average energy efficiency in all years is 0.876. The second highest is central area, of which the average energy efficiency in the whole period is 0.725. The energy efficiency in western area is the lowest, the average energy efficiency is only 0.679.

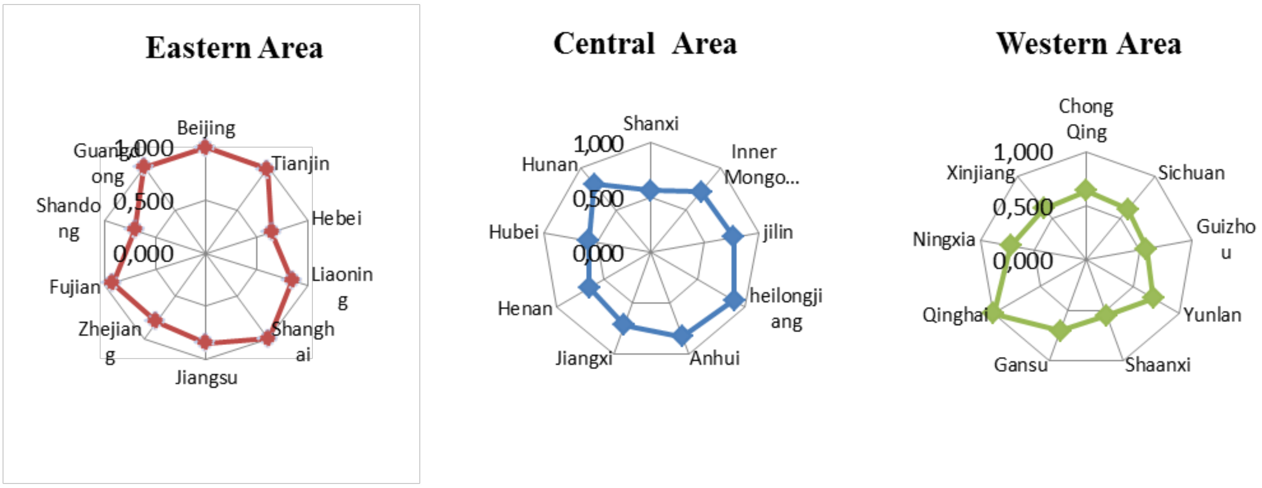
Figure 4 also show the average energy efficiency in different areas from 2001-2009, it indicates that all the regions in the eastern area have generally higher level of energy efficiencies and the distribution of energy efficiencies are more balanced than central area and western area. In eastern area, Hebei and Shandong are the regions that have relative low energy efficiency and they are the only two regions of which energy efficiency is less than 0.8. In central area, Heilongjiang is the region that is most close to the energy efficiency frontier. Shanxi and Hubei, however, the heavy industry is centralized there, are the worst performers in energy efficiency. The energy efficiency in these two regions is less than 0.6. The western area, however, has general relative low energy efficiency in all regions except Qinghai, which has reached the energy frontier. Shaanxi and Guizhou are the other two regions of which energy efficiency is less than 0.6 and Shaanxi is considered to have the lowest energy efficiency among all 28 regions in China 2001-2009.

Table 2: The overall energy efficiency in different regions 2001-2009

	2001	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**Average**
Eastern Area										
Beijing	1.000	1.000	0.982	1.000	1.000	1.000	1.000	1.000	1.000	**0.998**
Tianjin	0.877	0.966	0.974	1.000	1.000	1.000	1.000	1.000	1.000	**0.980**
Hebei	0.681	0.668	0.661	0.651	0.659	0.648	0.632	0.618	0.585	**0.645**
Liaoning	0.838	0.842	0.844	0.829	0.855	0.862	0.864	0.854	0.860	**0.850**
Shanghai	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	**1.000**
Jiangsu	0.868	0.883	0.872	0.840	0.827	0.833	0.835	0.846	0.852	**0.851**
Zhejiang	0.792	0.773	0.766	0.777	0.818	0.816	0.822	0.832	0.805	**0.800**
Fujian	1.000	1.000	1.000	1.000	0.881	0.878	0.874	0.863	0.846	**0.927**
Shandong	0.725	0.710	0.702	0.693	0.684	0.699	0.710	0.724	0.713	**0.707**
Guangdong	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	**1.000**
Average	**0.878**	**0.884**	**0.880**	**0.879**	**0.872**	**0.874**	**0.874**	**0.874**	**0.866**	**0.876**
Central Area										
Shanxi	0.560	0.567	0.578	0.585	0.585	0.565	0.559	0.548	0.514	**0.562**
Inner Mongolia	0.769	0.750	0.709	0.665	0.705	0.705	0.706	0.716	0.760	**0.721**
Jilin	0.835	0.808	0.787	0.769	0.771	0.751	0.751	0.743	0.723	**0.771**
Heilongjiang	0.944	0.920	0.905	0.898	0.915	0.896	0.860	0.846	0.769	**0.884**
Anhui	0.872	0.863	0.846	0.837	0.835	0.814	0.791	0.785	0.798	**0.827**
Jiangxi	0.740	0.714	0.694	0.680	0.724	0.722	0.719	0.712	0.742	**0.716**

	2001	2002	2003	2004	2005	2006	2007	2008	2009	Average
Eastern Area										
Henan	0.735	0.714	0.699	0.695	0.688	0.652	0.606	0.571	0.520	**0.653**
Hubei	0.597	0.580	0.572	0.569	0.580	0.578	0.577	0.587	0.606	**0.583**
Hunan	0.857	0.834	0.815	0.811	0.813	0.797	0.786	0.776	0.796	**0.809**
Average	**0.768**	**0.750**	**0.734**	**0.723**	**0.735**	**0.720**	**0.706**	**0.698**	**0.692**	**0.725**
Western Area										
Chong Qing	0.644	0.646	0.674	0.666	0.657	0.648	0.640	0.627	0.522	**0.636**
Sichuan	0.635	0.616	0.602	0.595	0.602	0.600	0.594	0.584	0.599	**0.603**
Guizhou	0.599	0.572	0.552	0.544	0.561	0.548	0.553	0.570	0.595	**0.566**
Yunnan	0.749	0.754	0.752	0.754	0.745	0.709	0.676	0.682	0.658	**0.720**
Shaanxi	0.576	0.563	0.555	0.551	0.554	0.546	0.534	0.534	0.571	**0.554**
Gansu	0.690	0.676	0.677	0.685	0.716	0.726	0.727	0.725	0.721	**0.705**
Qinghai	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	**1.000**
Ningxia	0.762	0.651	0.656	0.682	0.668	0.755	0.739	0.783	0.764	**0.718**
Xinjiang	0.628	0.618	0.611	0.597	0.613	0.613	0.610	0.632	0.589	**0.612**
Average	**0.698**	**0.677**	**0.675**	**0.675**	**0.680**	**0.683**	**0.675**	**0.682**	**0.669**	**0.679**
Total	**0.785**	**0.775**	**0.767**	**0.763**	**0.766**	**0.763**	**0.756**	**0.756**	**0.747**	**0.764**

Figure 4: Comparison of average energy efficiency among three areas in China



(a) (b) (c)

3.3.2 Comparison with Energy Intensity

Both table 2 and figure4 suggest that the energy efficiency has decreased slightly in all areas, which makes the national level of energy efficiency decrease from 0.785 to 0.764. Comparing with other two areas, central area has the more significant downwards trend in energy efficiency, the energy efficiency in central area has decreased from 0.768 to 0.725 between 2001 and 2009.

Energy intensity is known as the most common use indicator in measuring energy efficiency under single input-output formulation. It measures the energy consumption in producing the one unit of GDP (E/Y), thus the lower energy intensity implies the higher efficiency and vice versa. Figure 6 shows the changes of energy intensity in different areas. It shows that eastern area has the lowest energy intensity among three areas, which means eastern area has the highest energy efficiency. This result is consistent with the DEA calculation. However, the energy intensities in all areas actually have declined during the period, the energy intensity deceased from 1.3 to 0.9 in eastern areas, from 1.9 to 1.7 in central area. The energy intensity in western area increased from 2.3 to 2.7 in 2005 but decreased to 2.1 by the end of 2009. The declined energy intensity implies an uptrend of energy efficiency, which is inconsistent with the DEA result. One reason for this result could be that the energy intensity doesn't take other input factors (capital and labor) into account; GDP is the only output indicator and the environmental effect is excluded in the calculation. Thus the energy efficiency under the single input-output calculation may be overestimated.

Figure 5: The changes of overall energy efficiency in different areas 2001-2009

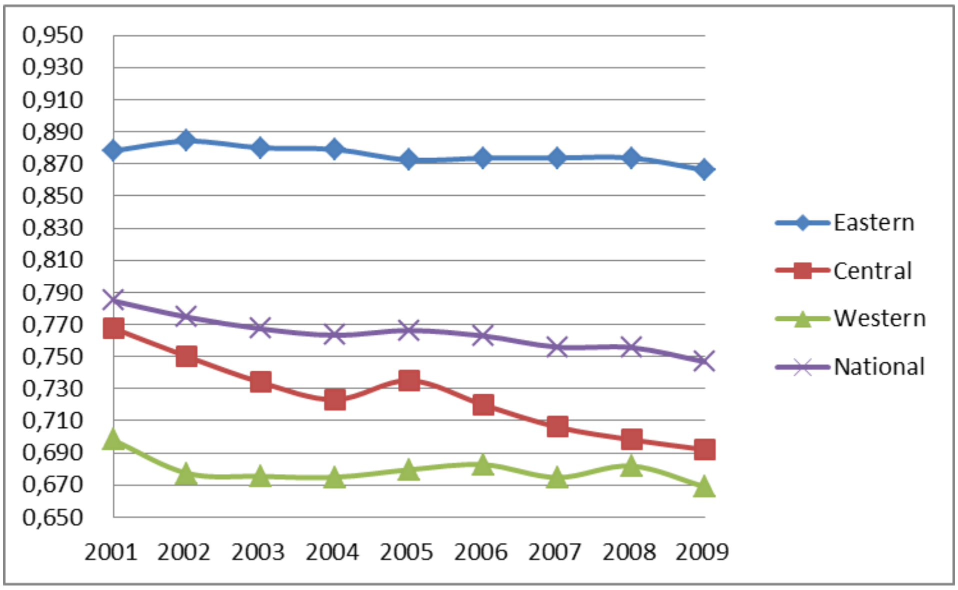
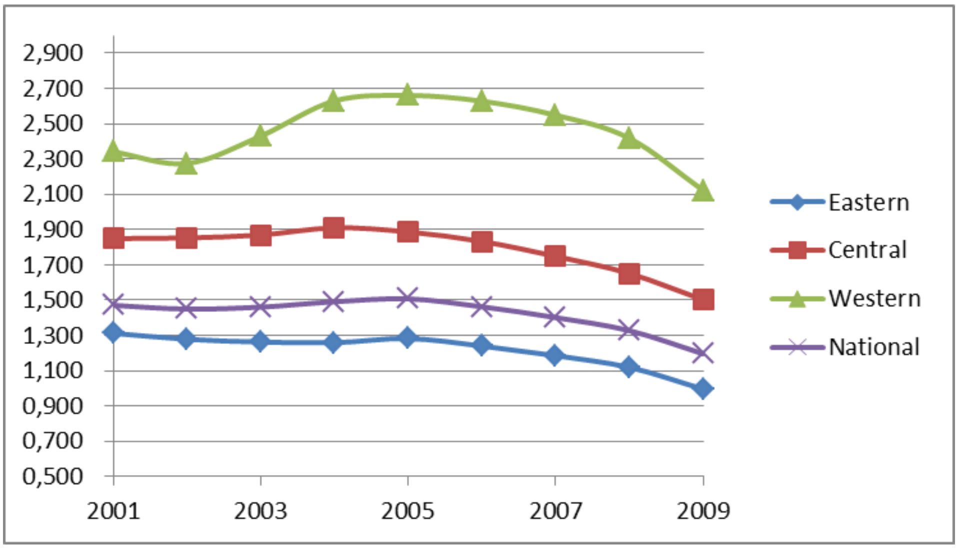


Figure 6: The energy intensity in different areas 2001-2009



3.3.3 The energy input redundancy and the regional saving potential

As it is mentioned before, one advantage of DEA method is that it not only tells the relative efficiency among DMUs, but also suggests the distance between inefficient DMU and efficient DMU through the slacks. By adjusting the slacks, the inefficient DMUs can reach the efficiency frontier and become efficient. In the input-orientated model, the slacks suggest the input redundancy, which is the difference between actual input and target input. Table 2 shows the slacks of energy input in different regions, it indicates that the energy input redundancy have increased during the period in all areas. The use of energy is not well structured and lots of energy is wasted during the production. Comparing the average energy input redundancy in nine years, eastern area has the highest energy input redundancy, 3310 Mtce, despite its relative high energy efficiency. Hebei and Shandong are the two regions that have the highest energy input redundancy among the eastern area, 10646 and 11011 Mtce respectively. These two regions are also the worst energy efficiency performers in eastern regions (see table 1). Central area has the second highest energy input redundancy, 3251Mtce. In this area, Shanxi has 6372 Mtce energy input redundancy, which is the highest in central area. The energy efficiency of Shanxi is also relatively low among all regions. In the western area, the average energy input redundancy in all regions is relatively low comparing with other areas. However, Sichuan shows much higher energy input redundancy, 4652 Mtce. In Sichuan, the energy efficiency is 0.603, which is not high. The result of energy input redundancy demonstrates that the energy input redundancy has significant influence on the overall energy efficiency. Regions have high energy input redundancy are usually accompanied with relative low efficiency. But the result is not completely consistent. For example, the energy efficiency in Shanxi is lowest among all regions, but the energy input redundancy is not the highest. This is because the energy efficiency is also affected by other factors, like capital, labor and the pollution.

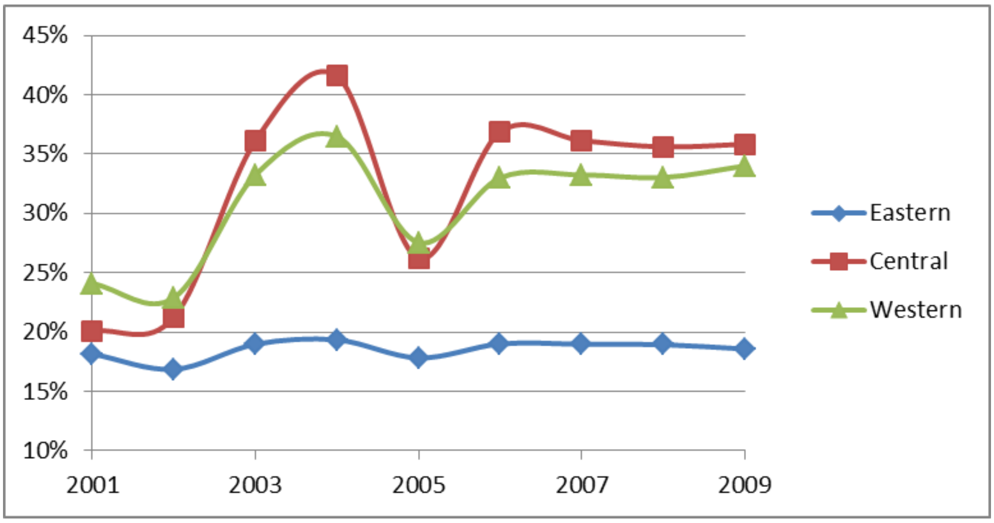
On the one hand, regions with higher energy input redundancy suggest severer waste of energy; on the other hand, the higher energy input redundancy also implies the higher energy saving potential. Table2 only reflects the absolute number of energy redundancy. However, the higher absolute number of energy redundancy doesn't reflect the energy saving potential properly since it is constrained by its level of total energy input. Namely, the share of energy redundancy in total actual energy input. Thus, the redundancy ratio (Redundancy energy/Actual energy input) is used to address this problem. Figure 7 shows the energy redundancy ratio in three areas. It shows that central area has the highest energy saving potential. There is about 35% energy input that can be saved to achieve the efficiency frontier. The second one is western area, 33% energy input can be saved to achieve the efficiency frontier. Eastern areas has the highest energy input redundancy in absolute number, but due to the energy input level is also much bigger than the other two, the share of redundancy in the total energy input is smaller. So the energy saving potential is relatively low, but still indicates that there are about 20% energy input can be saved to achieve efficiency frontier.

Table 3: The input slacks of energy (Mtce) in different regions 2001-2009

	2001	2002	2003	2004	2005	2006	2007	2008	2009	Average
Eastern Area										
Beijing	0	0	0	0	0	0	0	0	0	**0**
Tianjin	379	0	0	0	0	0	0	0	0	**42**
Hebei	5314	6101	8960	10560	9894	12793	13758	13941	14490	**10646**
Liaoning	5704	5253	5111	6042	6052	6750	7420	7883	8231	**6494**
Shanghai	0	0	0	0	0	0	0	0	0	**0**
Jiangsu	1110	968	1937	2128	3000	3393	3609	3529	3336	**2557**
Zhejiang	1349	1638	1574	2240	1913	2184	2402	2458	2508	**2029**
Fujian	0	0	225	0	165	247	419	753	1204	**335**
Shandong	5300	6842	8525	10213	11858	13552	14336	14784	13596	**11001**
Guangdong	0	0	0	0	0	0	0	0	0	**0**
Average	**1916**	**2080**	**2633**	**3118**	**3288**	**3892**	**4194**	**4335**	**4337**	**3310**
Central Area										
Shanxi	3911	4653	5426	6219	5355	6812	7287	7382	10300	**6372**
Innermongolia	1276	1726	4247	5479	6742	8182	9409	10502	8637	**6244**
Jilin	325	385	1646	1998	1402	2325	2537	2818	1955	**1710**
Heilongjiang	0	0	1262	3003	0	1812	1591	1529	3420	**1402**
Anhui	319	277	1037	1647	489	1403	1473	1577	1395	**1069**
Jiangxi	323	405	305	558	344	381	450	537	607	**435**
Henan	2008	2283	4657	6281	4728	7213	7841	8058	8704	**5752**
Hubei	2293	2498	3038	4177	3625	4342	4713	4866	5036	**3843**
Hunan	388	492	1938	2220	1112	3527	3763	3772	4681	**2433**
Average	**1205**	**1413**	**2617**	**3509**	**2644**	**4000**	**4341**	**4560**	**4971**	**3251**
Western Area	**967**	**988**	**1562**	**1983**	**1635**	**2191**	**2410**	**2546**	**2792**	**1897**

	2001	2002	2003	2004	2005	2006	2007	2008	2009	Average
Eastern Area										
Chongqing	486	454	513	565	578	669	896	1004	2487	**850**
Sichuan	2223	2554	3833	5175	3968	5311	5700	6148	6957	**4652**
Guizhou	2180	2078	3214	3832	3202	4118	4469	4525	3797	**3491**
Yunnan	565	508	1666	2037	1030	2540	2686	2789	2943	**1863**
Shanxi	908	1065	1305	1552	1642	1808	2289	2481	2554	**1734**
Gansu	823	782	1303	1711	1122	1804	1889	1942	1899	**1475**
Qinghai	0	0	0	0	0	0	0	0	0	**0**
Ningxia	0	0	450	886	776	833	891	888	938	**629**
Xinjiang	1520	1448	1775	2092	2401	2639	2872	3133	3548	**2381**
Average	**967**	**988**	**1562**	**1983**	**1635**	**2191**	**2410**	**2546**	**2792**	**1897**

Figure 7: The energy redundancy ratio in different regions 2001-2009



4. Factors that influence energy efficiency

4.1 Hypothesis and data

Section 2 reviews previous study on the factors affecting energy efficiency, and it can be sorted into five perspectives: market openness, energy prices, structural change, technology development and scale. Thus, based on the concepts and the available data, five variables are selected to represent those perspectives. I include the value added of import and export (impexp) as a measure of openness of economy. Since there is no complete data of energy price in regional level, I use producer price index (ppi) as proxy to energy price to capture the effects of changing raw material price on energy efficiency. Among different sectors, industry is the largest energy consumer and has the most significant effect on the energy efficiency, so I use the share of total value added from industry (ind) to indicate structural change on energy efficiency. For the technology, patent is the most common and direct reflection to the technology achievement, so I use the number of patents (patents) in each region to measure the level of technology. The scale effect (scale) is captured through the share of total value added from large and medium size enterprises.[3] All the explanatory variables data stem from the statistics yearbook of each regions from 2002-2010[4]. Import and export data is originally measured in U.S dollar, so I converts it into Chinese Yuan (CNY) by times the Federal Reserve Exchange Rate[5] between CNY and U.S dollar. Table 4

[3]According to the Chinese statistical principle, the annual turnover that is larger than 500 million RMB (Chinese Yuan) is considered as medium enterprise.

[4] Data are collected through the statistical data base in each region, more details see http://acad.cnki.net/Kns55/brief/result.aspx?dbPrefix=CYFD

[5] The yearly exchange rate is from Ecowin database, the source is: Exchange Rate, China, Fixings, Federal Reserve USD / CNY, Fixing, ew:chn19870, Ecowin.

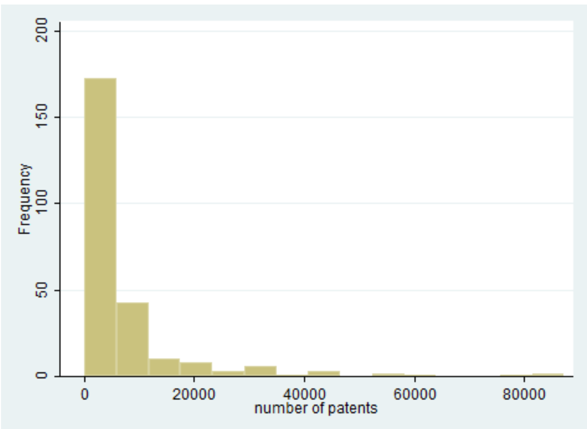
shows the statistic summary of the data across provinces during the year 2001 and 2009.

The panel data contains 28 provinces in mainland China (excluding Tibet, Guangxi and Hainan), and the data is strongly balanced. However, it is noticeable that the patent data has outliers (Figure 8), which may influence the estimation of error in our regression later.

Table 4: The summary statistics of Explanatory Variables

Variable	obs	Mean	Std. Dev.	Min	Max
impexp	252	3.92E+07	7.77E+07	162817.9	4.67E+08
ppi	252	102.7659	5.74238	85.5	122.4
ind	252	0.4797354	0.0616301	0.234966	0.614777
patent	252	7589.167	12842.65	70	87286
scale	252	0.7617967	0.2673849	0.279478	1.834432

Figure 8: the histogram of Patents



I include the result of energy efficiency from DEA measurement as the independent variable and the impexp, ppi, ind, patent, scale as explanatory variables. The main hypotheses are:

- Import and export share shall be positively related to the energy efficiency. A more opened economy policy could strengthen the technology exchange between countries, thereby, improving the technology innovation.

- Producer price index is expected to be negatively correlated with energy efficiency. An increasing energy price would motivate firms to use energy input in a more efficient way.

- The share of value added in industry is expected to be negatively correlated with the energy efficiency as discussed in section 2.

- Patents shall be positively related to energy efficiency due to the improvement of energy efficiency from the technology shocks.

- Production proceeds under optimized scale have positive effect on the efficiency; however it is inaccurate to assume that larger scale will always have higher efficiency since the overinvestment will also generate decrease return to scale. Thus, the relation between share of total value added from large， medium enterprises and energy efficiency is uncertain.

4.2 The fixed effect model

First, I take the logarithm of all of the variables. However, energy efficiency could also be impacted by other factors that are not directly observed among regions. Omitting those variables could make the OLS estimation biased and inconsistent. One way to overcome the problem is to add the unobserved effects into the linear equation along with other explanatory variables to capture the unobserved effects that are time-invariant, which leads to the fixed effects model. Another reason is the use of province data that makes the observations cannot be random draws from a large population (Wooldridge, 2003). The fixed effect model is described below

$Y_{it} = \beta_1 X_{it} + \alpha_i + \lambda_t + \mu_{it}$,Where α_i is treated as the regional-specific intercept and λ_t is the time fixed effect.

The advantages of using the panel data is to make us remove the time invariant unobserved characteristics of the regions that could both impact the efficient score and the explanatory variables. Such unobserved characteristics could come from the energy consumption habits across provinces. In cross-sectional data, the OLS estimator usually suffers from the omitted variable problem, therefore, it is biased. With consecutive series of efficiency score and the observable variables, such bias could be reduced by controlling for the region specific fixed effect. Moreover, the time fixed effect could eliminate the omitted variables bias from unobserved variables that are constant over the regions, such as the central government's policy towards the encouragement of energy innovation. However, in order to enable interpretation of the result as a causal, there are no other regional characteristics which correlate with the efficiency performance (score) and our explanatory variables change over time, which is a strong assumption. The model with Econometric Specification shows as follows:

$$lneff_{it} = \beta_1 lnimpexp_{it} + \beta_2 lnppi_{it} + \beta_3 lnind_{it} + \beta_4 lnpatent_{it} + \beta_5 lnscale_{it} + \alpha_i + \lambda_t + \mu_{it}$$

Where i indicates the specific province, and t indicates the specific year. For instance, $lnppi_{it}$ indicates the producer price index in the province i and year t.

4.3 Results and analysis

Table 5: Estimations of Pooled OLS Model and Fixed Effects Model

	(1)	(2)	(3)
VARIABLES	Pooled OLS	model 2	model 3
Lnimpexp	0.169***	-0.00464	0.0353
	(0.0139)	(0.0167)	(0.0305)
Lnppi	-0.732***	0.172**	0.191*
	(0.265)	(0.0676)	(0.0996)
Lnind	-0.152**	-0.0559	-0.0367
	(0.0606)	(0.0724)	(0.0704)
Lnpatent	-0.163***	0.000967	0.0120
	(0.0181)	(0.0100)	(0.0151)
Lnscale	-0.0269	-0.0573*	-0.0676
	(0.0358)	(0.0302)	(0.0465)
Observations	252	252	252
R-squared	0.450	0.156	0.189
Regional Effects?	No	Yes	Yes
Time Effects?	No	No	Yes
Number of Regions		28	28

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Model 1 shows the results from the pooled OLS regression without any control for the regional and time effects.[6] As we can see, the lnimpexp variable is positively significantly effects the energy efficiency at 1% level. The estimated parameter

[6]The constant term is not shown here since our main interests are the estimated explanatory variables.

associated with producer price index (ppi) is negative, which is not consistent with the theory and hypothesis. Moreover, the two standard errors also fall in a negative area. The patent has a negative and significant effect on the energy efficiency, which also contradicts with our common sense. The industry structure also has negatively effects on energy efficiency with significant level at 5% level. The sign of estimated parameter associated with scale effect is negative, which indicates that the expansion of industry scale discourages the energy efficiency. Overall, the pooled OLS explains 45% variation of our dependent variable.

Model 2 shows the estimation results if we control for the regional effects, which generally arises from the omitted variables but constant overtime. Such unobservable variables could comes from energy consumption habits of people, where the regions in north of China would certainly consume more energy in the winter compared with south. The F statistic is 2.24, which suggests the rejection of the null hypothesis that all the estimated coefficients are zero simultaneously at 10% level. Compared with pooled OLS model, the value of estimated parameters are generally smaller in absolute value, which concludes that the pooled OLS estimator is upward bias when suffering omitted variables. The estimated parameter of lnimpexp becomes no longer significant and it is negative, but the 95% confidence interval is so large that covers from negative to positive values. The estimated parameter associated with lnppi is still significant at 5% level, but the coefficient becomes positive as we expected. And it decreases to 0.172 in absolute value compared with -0.732. The variables lnind (industrial structure) and patents are not statistically significant at all. The scale effect is still significant at 10% level. The negative sign indicates that a more scaled economy could decrease the energy efficiency. However, the 95% confidence level covers both positive and negative numbers (from 0.003 to -0.1177), which leads us not to be sure about the impact of scale effects on the energy efficiency.

Model 3 shows the estimation results from controlling of regional effect as long as time effects that the unobserved variables are raised from time varying but constant over the regions, such omitted variables could come from the central government's policy, for instance, the increase tax of energy products and introduction of new regulations about pollutant emissions. Looking at the estimated parameters, they do not differ much compared with model 2 except for the patent variable. Among the five explanatory variables, only lnppi has a statistically significant effect on the energy efficiency at 10% level. The rest of the variables are not significant at all. The estimated coefficient associated with lnppi increases from 0.172 to 0.191. The F statistic of the model is 5.76, which indicates we reject the all the coefficients are jointly zero at 1% level. Testing the interceptions from time variables, the F statistic is 3.5 with p-value equals to 0.007, which indicates that the time effects are indeed significant. Except for the producer price index has a positive effect on the energy efficient, we are not certain about how the rest four variables affect the energy efficient since the standard error is so large that covers both positive and negative numbers.

To sum up, the pooled OLS model seems to suffer the omitted variables severely, and the estimated parameters tend to be upward bias. While the fixed effects models certainly improve the accuracy of the estimated parameters. However, the significance level of estimated variables differs from model 2 to model 3. The possible reason behind this could be the presence of outliers and small data set that makes the estimated parameters unstable. Therefore, we should carefully analyze the effects of the five explanatory variables on the dependent variable, even lnppi is significant at 10% level, we should be careful about the interpretation due to the small data size.

4.4 Robustness analysis

4.4 .1 Fixed Effects Model VS Random Effects Model

In the random effects model, we assume the unobserved variables α_i are not correlated with the explanatory variables. If this assumption is valid, then β_1 would be consistently estimated by OLS with a single cross section, but single cross section estimation would scarify the useful information in other time periods (Wooldridge, 2003).

$Y_{it} = \beta_1 X_{it} + \mu_{it}$, Where $\mu_{it} = \alpha_i + \varepsilon_{it}$. μ_{it} is the composite error term.

In the robustness analysis, firstly, we test the fixed effects model with regional control against the random effects model by using the Hausman test.

Table 6: Fixed Effects Model against Random Effects Model

VARIABLES	(1) Fixed Effects Model	(2) Random Effects Model
Lnimpexp	-0.00464	0.0195*
	(0.0115)	(0.0114)
Lnppi	0.172***	0.119*
	(0.0604)	(0.0649)
Lnind	-0.0559	-0.0654
	(0.0468)	(0.0496)
Lnpatent	0.000967	-0.0115
	(0.00929)	(0.00985)
Lnscale	-0.0573***	-0.0721***
	(0.0185)	(0.0190)

VARIABLES	(1) Fixed Effects Model	(2) Random Effects Model
Observations	252	252
R-squared	0.156	0.138
Number of Regions	28	28

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Performing the hausman test, the Chi-square statistic is 38.44 with p-value equals to 0, which indicates we should reject the null hypothesis that difference in coefficients are not systematic (the unobserved variables are not correlated with the explanatory variables), in other words, the fixed effects model is preferred.

4.4.2 Robustness test of the Fixed Effects model for the Eastern, Central and Western of China

In this part, we classify the provinces of China into three categories, which are east, center and west according to the geographical location. Table7 shows the fixed effects model regression for the three geographical regions of China. Among the three models, the lnppi still has a significant effect on regions of east and center. The estimated parameters associated with lnppi are covered by the two standard errors from table5 model 3 except for the center case. The variable patent is significant at 10% level in the east model. However, in such a small data size, we cannot confidently conclude the variable patent has a strong effect on energy efficiency even if it is significant at 10% level. All the F tests show we should reject the null hypotheses that the estimated time effects are jointly zero at 5% level. The big variations of the estimated parameters may come from the small observations. Another pitfall of the fixed effects model is that we cannot capture the unobserved variables that are time changing, thereby it is possible to violate the strict exogeneity assumption and lead to biased estimations. In general, our findings from the robustness analysis tend to agree with the estimation result in section 4.3, which indicates that an increase in producer price (as a proxy to the energy price) is more likely to improve the use of energy in a efficient way.

Table 7: Fixed Effects Model for Eastern, Central and Western Region

	(1)	(2)	(3)
VARIABLES	model East	model Center	model West
lnimpexp	-0.00377	0.102	0.0118
	(0.0740)	(0.0680)	(0.0404)
lnppi	0.271***	0.449*	0.148
	(0.0780)	(0.231)	(0.174)
lnind	0.0440	0.0194	-0.374
	(0.0871)	(0.156)	(0.260)
lnpatent	0.0629*	-0.0190	0.0332
	(0.0302)	(0.0199)	(0.0382)
lnscale	-0.103	-0.0645	-0.0471
	(0.0737)	(0.120)	(0.114)
Observations	90	81	81
R-squared	0.300	0.485	0.226
Number of Regions	10	9	9
Regional Effects?	Yes	Yes	Yes
Time Effects?	Yes	Yes	Yes

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.

5. Discussion and Conclusion

In this paper, I first examine the energy efficiency in China at the regional level from 2001-2009 by using the DEA model. Under the model with one desirable and one undesirable output, the results first show that the energy efficiency varies among regions in China. The eastern area has the highest energy efficiency (the average efficiency is 0.876) while the western area has lowest energy efficiency (the average efficiency is 0.679).

Secondly, there is a tendency that the energy efficiency was decreasing between 2001 and 2009, especially in the central area, where the average energy efficiency dropped from 0.77 to 0.69 during the nine years. However, the Chinese GDP was growing at around 9% per year over the same period, which indicates that the impressing growth was achieved at the expense of resources and environment. A more sustainable and energy-saving growth pattern need to be adjusted to the current energy policy. However, comparing the result with energy intensity in the same period, the energy intensity was decreasing, which means that the energy efficiency measured by single input-output ratio shares an upswing trend. The distinct results suggest that when excludes other input factors (labor and capital) and environmental effects, the energy efficiency will be overestimated.

Thirdly, there are only three regions (Shanghai Guangdong and Qinghai) that keep themselves on the efficiency frontier in all years, which indicates that most regions in China has not use the energy in efficient and effective ways. However, the slacks under DEA model suggest the distance between inefficient DMUs and the efficiency frontier. By adjusting the slacks, the inefficient DMUs can reach the efficiency frontier. The input-orientated CCR model refers slacks to input redundancy and through the comparison of energy input redundancy ratio among regions, regions in central area have the 35% energy redundancy ratio, which is the highest among all areas. However, the higher energy redundancy ration implies the higher energy saving potential. The government in central area should promote the energy policy with the aim at energy saving in order to shift the regions to energy efficiency frontier.

To further explore the factors that engender the energy disparities among regions, I apply the econometric analysis with fixed effects model to investigate how different factors impact energy efficiency. As depicted in figure 3, five different variables are introduced as proxies to the factors that could affect energy efficiency. The econometric results show that only the producer price index, which is a proxy to the energy price, has a positive significant effect on the energy efficiency. We are not certain how the other four variables influence the energy efficiency since the standard error is too large. Moreover, the existence of the measurement error in either explanatory variables or dependent variable could affect our fixed effects model estimation. An ideal possible solution could be appending the data by including more time periods (years), using 252 observations are certainly too small so that the estimated parameters are not stable. However, the data limitation should be considered before 2001. At the same time, the fixed effects model cannot control the exogeneity that is time varying, which could make our fixed effects estimation still biased.

For future studies, a more precise measurement of DEA model with both desirable and undesirable output could be implemented compared with using a simplified approach to handle the undesirable output in this paper. And also more complicated DEA model can be introduced in the analysis, like the super DEA model that measures the rank of efficient DMUs, and the window analysis that evaluates the dynamic effect of efficiency change. In the econometric analysis, instead of using the fixed effects model, another approach to obtain the unbiased and consistent OLS estimation could be using instrumental variables (two stage least square estimation), where the instrument should be correlated with the explanatory variables and uncorrelated with the error term. However, due to the time limitation this approach is not done in this paper which could be done otherwise. Without finding a good instrument, it is so easy for the model to suffer the weak instrument problem, which in turn doe not benefit our estimation better than the fixed effects model. With more time periods data at hand, we could even investigate the dynamic effects of the energy efficiency by using the Generalized Method of Moments estimation.

Reference

I. E. A. (2010). World energy outlook. International Energy Agency. (Paris : International Energy Agency, 2010), 731.

Ang, B. W. (2006). Monitoring changes in economy-wide energy efficiency: From energy–GDP ratio to composite efficiency index. *Energy Policy, 34*(5), 574-582.

Arbex, M., & Perobelli, F. S. (2010). Solow meets Leontief: Economic growth and energy consumption. *Energy Economics, 32*(1), 43-53.

Ayres, R., Turton, H., & Casten, T. (2007). Energy efficiency, sustainability and economic growth. *Energy (Oxford), 32*(5), 634-648.

Banker, R. D., & Morey, R. C. (1986). Efficiency Analysis for Exogenously Fixed Inputs and Outputs. *Operations Research, 34*(4), 513-521.

Berndt, E. R. (1975). Aggregate energy, efficiency and productivity measurement *Annual Review of Energy* 3225-3249.

Besseboeuf D, C. B., Lapillone B,. (1997). Cross-country comparison on energy efficiency indicators: the on-going European effort towards methodology *Energy Policy, 25(9)*, 673-682

Birol, F., & Keppler, J. H. (2000). Prices, technology development and the rebound effect. *Energy Policy, 28*(6–7), 457-469.

Boyd, G. A. (2005). A Method for Measuring the Efficiency Gap between Average and Best Practice Energy Use: The ENERGY STAR Industrial Energy Performance Indicator. *Journal of Industrial Ecology, 9*(3), 51-65.

Chaoqing, Y., Sifeng, L., Zhigeng, F., & Naiming, X. (2010). The relation between Chinese economic development and energy consumption in the different periods. *Energy Policy, 38*(9), 5189.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research, 2*(6), 429-444.

Chien, T., & Hu, J.-L. (2007). Renewable energy and macroeconomic efficiency of OECD and non-OECD economies. *Energy Policy, 35*(7), 3606-3615.

Chwieduk, D. (1997). The role of Polish Energy Conservation Agency in promoting energy efficiency and renewable energy in Poland. *Renewable Energy, 10*(2–3), 323-326.

Council, W. E. (2010). Energy Efficiency Policies and Indicators-A report by the World Energy Council

Crompton, P., & Wu, Y. (2005). Energy consumption in China: past trends and future directions. *Energy Economics, 27*(1), 195-208.

de Vos, R. (2004). Energy efficiency: The forgotten option? *Refocus, 5*(4), 58-59.

Erye, N. (1998). A golden Age for A False Dawn? Energy Efficiency in UK Competitve Energy Market *Energy Policy, 26(12)*, 963-972.

Faere, R., Grosskopf, S., Lovell, C. A. K., & Pasurka, C. (1989). Multilateral Productivity Comparisons When Some Outputs are Undesirable: A Nonparametric Approach. *The Review of Economics and Statistics, 71*(1), 90-98.

Fan, Y. (2007). Can Market Oriented Economics Reforms to Energy Efficiency Improvment ? Evidence from China *Energy Policy 35(4)*, 2287-2295.

Goldemberg, J., & Siqueira Prado, L. T. (2011). The decline of the world's energy intensity. *Energy Policy, 39*(3), 1802-1805.

Guoquan, X. (2008). The energy efficiency in eight economic zones in China based on DEA model *Chinese energy policy* 23-45.

Howarth, R. B. (1997). ENERGY EFFICIENCY AND ECONOMIC GROWTH. *Contemporary Economic Policy, 15*(4), 1-9.

Howarth, R. B., Schipper, L., Duerr, P. A., & Strøm, S. (1991). Manufacturing energy use in eight OECD countries: Decomposing the impacts of changes in output, industry structure and energy intensity. *Energy Economics, 13*(2), 135-142.

Hu, J.-L., & Wang, S.-C. (2006). Total-factor energy efficiency of regions in China. *Energy Policy, 34*(17), 3206-3217.

IEA. (2011). G-20 Clean Energy, and Energy Efficiency Deployment and Policy Progress-. 52. Retrieved from http://www.iea.org/publications/free_new_Desc.asp?PUBS_ID=2457

Jan, W. s. (2008). Explaining the declining energy intensity of the U.S economy. *Resources and Energy Economics 30(1)*, 21-49

Jensen, J., & Tarr, D. (2003). Trade, Exchange Rate, and Energy Pricing Reform in Iran: Potentially Large Efficiency Effects and Gains to the Poor. *Review of Development Economics, 7*(4), 543-562.

John, G. F. (2004). ENERGY EFFICIENCY. *Mechanical engineering (New York, N.Y. 1919), 1*(3), 2.

Lozano, S., Villa, G., & Brännlund, R. (2009). Centralised reallocation of emission permits using DEA. *European Journal of Operational Research, 193*(3), 752-760.

Ma, H., Oxley, L., & Gibson, J. (2010). China's energy economy: A survey of the literature. *Economic Systems, 34*(2), 105-132.

Madlener, R., & Alcott, B. (2009). Energy rebound and economic growth: A review of the main issues and research needs. *Energy, 34*(3), 370-376.

Patterson, M. G. (1996). What is energy efficiency?: Concepts, indicators and methodological issues. *Energy Policy, 24*(5), 377-390.

Polsenberg, J. (2004). Reducing Energy Consumption in China. *Frontiers in Ecology and the Environment, 2*(7), 345.

Schleich, J. (2004). Do Energy Audits Help Reduce Barriers to Energy Efficiency *International Journal of Energy Technology and Policy 2(3)*, 226-239.

Seiford, L. M., & Zhu, J. (2002). Modeling undesirable factors in efficiency evaluation. *European Journal of Operational Research, 142*(1), 16-20.

Shi, G.-M., Bi, J., & Wang, J.-N. (2010). Chinese regional industrial energy efficiency evaluation based on a DEA model of fixing non-energy inputs. *Energy Policy, 38*(10), 6172-6179.

Sueyoshi, T., & Goto, M. (2011). DEA approach for unified efficiency measurement: Assessment of Japanese fossil fuel power generation. *Energy Economics, 33*(2), 292-303.

Trinh, T. A., Hlavacs, H., & Talia, D. (2012). Energy efficiency in large-scale distributed systems. *Future Generation Computer Systems, 28*(5), 743-744.

Wang, K., Yu, S., & Zhang, W. China’s regional energy and environmental efficiency: A DEA window analysis based dynamic evaluation. *Mathematical and Computer Modelling*(0).

Wu, Q. (2010). Research on Provincial Energy Efficiency Evaluation in China (Dalian University of Technology).

Xie Q, P. Y., J. (2011). The evaluation of regional capital stock from 1952-2009 *Contemporary Economic Policy*

Yuan, C., Liu, S., & Wu, J. (2010). The relationship among energy prices and energy consumption in China. *Energy Policy, 38*(1), 197-207.

Zhang, M., Li, G., Mu, H. L., & Ning, Y. D. (2011). Energy and exergy efficiencies in the Chinese transportation sector, 1980–2009. *Energy, 36*(2), 770-776.

Zhou, P., & Ang, B. W. (2008). Linear programming models for measuring economy-wide energy efficiency performance. *Energy Policy, 36*(8), 2911-2916.

Data:

Total Employed Persons, Number of Employed Person s at Year-end by Three Strata of Industry and Region, Chinese statistical yearbook 2002-2010, National Bureau of Statistics of China (http://www.stats.gov.cn/tjsj/ndsj/)

Capital stock, Xie Q, P. Y., J. (2011). The evaluation of regional capital stock from 1952-2009 *Contemporary Economic Policy*

Total energy consumption (10000 tons of SCE), Indicators of Energy Consumption by Region, Chinese statistical yearbook 2002-2010, National Bureau of Statistics of China (http://www.stats.gov.cn/tjsj/ndsj/)

Total Volume of industiral Waste Water Discharged(10000 tons), Chinese environmental statistical year book 2002-2010, National Bureau of Statistics of China (http://www.stats.gov.cn/tjsj/qtsj/hjtjzl/index.htm

Total Volume of Industrial Waste Gas Emsission(100million cu.m), Chineseenvironmental statistical year book 2002-2010, National Bureau of Statistics of China (http://www.stats.gov.cn/tjsj/qtsj/hjtjzl/index.htm)

Industrial Solid Wastes Generated(10000 tons), Chineseenvironmental statistical year book 2002-2010, National Bureau of Statistics of China (http://www.stats.gov.cn/tjsj/qtsj/hjtjzl/index.htm)

Appendix: The map of China