UNIVERSITY OF CALGARY

An Economic Evaluation of Energy Substitution in China

by

Yuanying Huang

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF ARTS

DEPARTMENT OF ECONOMICS

CALGARY, ALBERTA

May, 2006

UNIVERSITY OF CALGARY

FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled "An Economic Evaluation of Energy Substitution in China" submitted by Yuanying Huang in partial fulfilment of the requirements of the degree of Master of Arts

Supervisor, Frank J. Atkins, Department of Economics

Daniel V. Gordon, Department of Economics

External Examiner ,
Jess Chua, Haskayne School of Business

May 12th, 2006

Date

Abstract

China has been playing a more and more important role in the world energy market for the past few years. With the largest population in the world, rapid economic growth and soaring energy demand, China will likely continue to be an influential factor in the world energy market. Despite the great importance and urgency of this issue, there has been little econometric empirical study of Chinese energy structure. There has been especially little study of energy substitution, probably because of the unavailability of data.

This thesis seeks to narrow the empirical gap by studying Chinese production structure, significant energy variables and energy substitution. Two approaches are used. The first approach takes advantage of duality theory, and the production function is studied through the dual cost function. Flexible functional form is introduced, economic regularity conditions are tested, and more efficient system estimation is used. Elasticity of Substitution, own and cross price elasticities are also calculated. The second approach uses the model of Vector Autoregression (VAR) to uncover some data properties. Impulse response function and dynamic correlation are calculated. Overall these two approaches give consistent results.

Interestingly enough, the theoretical regularity condition of curvature is violated in the early sample points but satisfied for recent years. This might be an indication that the Chinese economy is behaving more consistently with economic theory lately, but not in the early sample years.

The relative low value of elasticity of substitution indicates that two main fuels for China, coal and oil, are not easily substituted for each other. Overall empirical results support positive own price elasticity for oil. It has to be noted that these conclusions are based on a sample period from 1960 to 2003, and there is reason to believe that the Chinese economy is behaving differently now than during previous years.

Acknowledgements

First of all, I would like to express my sincere gratitude to my supervisor, Dr. Atkins, who always give me guidance and support. Without his wisdom and help, I wouldn't be able to complete my research and walk so far in my life.

I want to give special thanks to Dr. Gordon, who is always open and helpful for my problems.

I also want to thank Dr. Chua for suggestions to make the thesis better.

Last but not least, thanks for all my classmates and friends for supporting me in various ways.

Dedication

To My Parents,

Huang Jiansheng

and

Du Shuhua

Table of Contents

List of Tables

List of Figures

CHAPTER 1. INTRODUCTION

1.1 A Review of Past Energy Trend and Policies in China

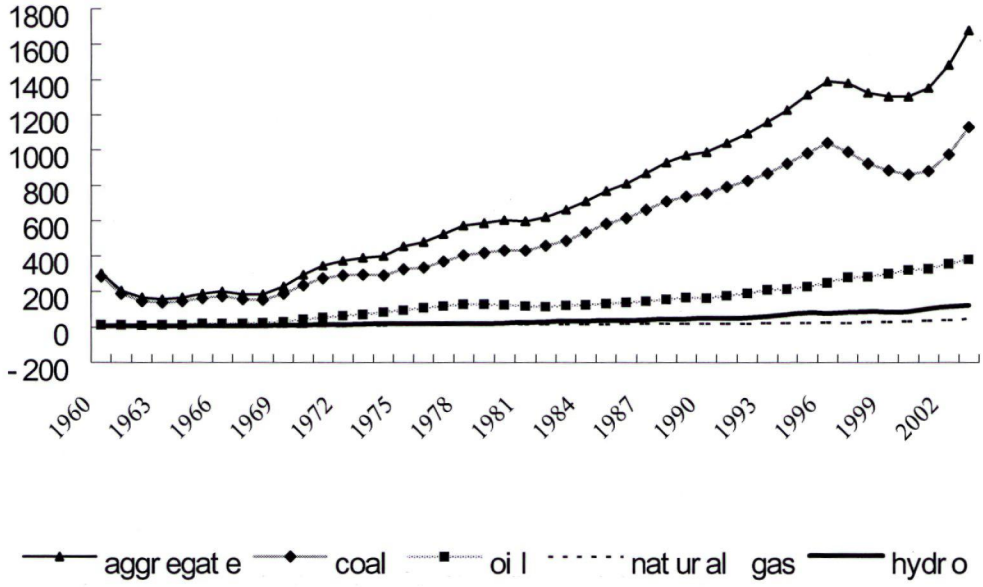
With the largest population in the world and rapid economic growth, China is the second largest energy consumer in the world, and is playing a more and more important role in the world energy market. However, contrary to most other developing countries at China's income level, China's income elasticity of energy consumption is well below one (Zhang, 2000), implying an energy intensity fall. This intensity fall may be attributed to energy policies, energy usage efficiency, technological progress, and structural changes in the economy. In one word, energy status in China is interesting and important.

Aggregate energy consumption in China is comprised of four primary energy types: coal, oil, natural gas and hydropower, with coal being the dominant fuel. From Figure 1.1, it is clear that total consumption and the consumption of coal have an increasing trend until 1997, a decline until 2002 and then the increasing trend continues. The reasons for this decline are discussed later. In 2003, China's coal consumption was 1.13 billion tons, which counts for 28% of the world total coal consumption (EIA, 2005).

Figure 1.2 shows the proportion of coal consumption to total energy consumption in China has a decreasing trend, with 93.9% in 1960 to 67.1% in 2003. In contrast, the proportion of oil consumption increased, though not monotonically, over time, from 4.1% in 1960 to 22.7% in 2003. China became a net importer of oil in 1993 and was the world's second largest consumer of petroleum products in 2004, with total demand of 6.5 million barrels per day (bbl/d). China's oil demand is projected by the EIA to reach 14.2 million bbl/d by 2025, with net imports of 10.9 million bbl/d.

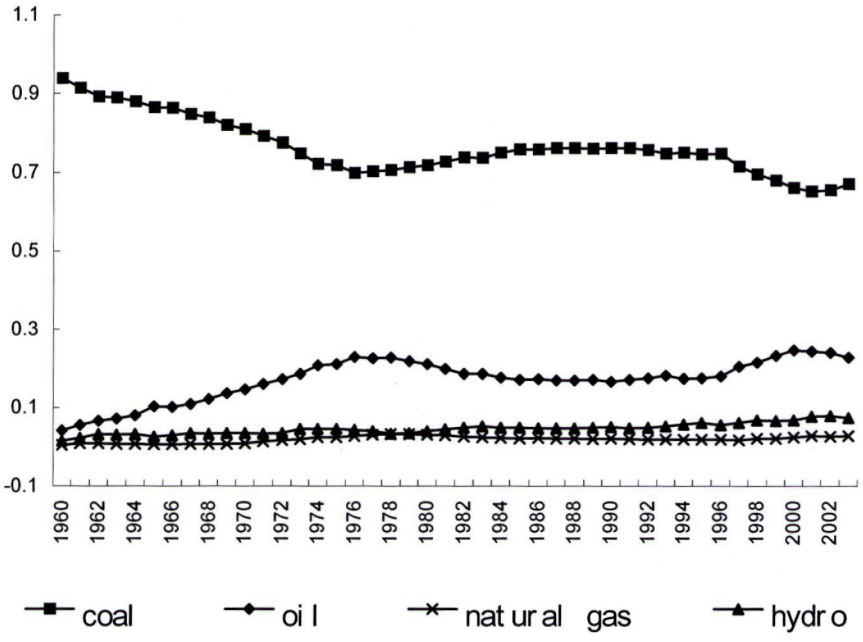
Figure 1.1 Aggregate and Primary types of Energy Consumption

(Million tons of coal equivalent)



Source: China Statistical Yearbook, various issues.

Figure 1.2 Share of Different Energy Types



Hydro-electricity consumption makes up 7.4% of China's total energy consumption in 2003. Although it's a relatively small proportion of total energy consumption, it has been increasing through time and is expected to grow in the future. The largest project under construction is the Three Gorges Dam, which, when fully completed in 2009, will include 26 separate 700-MW generators, for a total of 18.2 GW. Another large hydropower project involves a series of dams on the upper portion of the Yellow River. With eventual construction of 25 generating stations, there will be a combined installed capacity of 15.8 GW. (EIA, 2005)

Historically, natural gas has not been a major fuel in China, but given China's domestic reserves of natural gas, which stood at 53.3 trillion cubic feet (Tcf) at the beginning of 2005 (EIA 2005), and the environmental benefits of using natural gas, China has embarked on a major expansion of its gas infrastructure.

A closer scrutiny of Figure 1.2 implies three stages of China's energy consumption pattern, which seem to be separated by the years 1978 and 1997. This phenomenon is a direct consequence of China's economic growth, energy demand and supply, and governmental policy intervention.

The first stage can be defined as the period from 1960 to 1978. In the 1960s, coal has absolute dominance in energy consumption. However, the share of coal had been decreasing and the share of oil had been increasing until around 1978. This is a result of the increased domestic oil production output in the 1960s and early 1970s, which is supported by the discovery and exploration of huge oil fields in the Northeastern region. Oil was believed to have had abundant supply at the time and consequently, demand was stimulated. Therefore, coal consumption declined drastically, oil consumption increased.

The second stage is from 1978 to 1997. It is well known that in 1978, after years of state controlled economy, China commenced a series of economic reforms. It encouraged the formation of rural enterprises and private businesses, liberalized foreign trade and investment, relaxed state control over some prices, and invested in industrial production and the education of its workforce. The strategies worked spectacularly and the economic reform brought into a sharp increase in energy

consumption, particularly in the industrial sector. Oil supply was then constrained by production capacity limitations. Subsequent demands for energy would have to be met by increasing coal consumption. This explains why there seems to be a different energy consumption pattern before and after 1978.

The third stage is marked by the year 1997. Energy consumption in China reached a maximum in 1996, and then the aggregate energy consumption shows a clear drop, which is mainly caused by the decrease in coal consumption. Why did this happen?

From table 1.1, it can be seen that starting from 1997, the growth rate of real GDP is slower than in previous years. The slow-down of economic growth may have resulted in less demand for energy, an important factor input.

Table 1.1 Growth Rate of Chinese GDP

(Constant Local Currency)

1991	9.20%
1992	14.20%
1993	13.50%
1994	12.60%
1995	10.50%
1996	9.60%
1997	8.80%
1998	7.80%
1999	7.10%
2000	8.00%
2001	7.50%
2002	8.30%
2003	9.29%

Source: World Development Indicator, World Bank

Other reasons for the aggregate energy consumption decline may be attributed to relevant national policies such as "Energy Conservation" and "Environmental Protection". The Energy Conservation Law (ECL) was approved on 1 November 1997 and came into effect on 1 January 1998. Actually, "Energy Conservation" has long been on China's national planning agenda because of the sustainability of economic development and domestic energy supply limit. China did constitute its energy development strategy in 1981, the strategy could be briefed stated as "to

produce and save energy simultaneously, with priority on saving energy". In the "Proposal to Formulate the Ninth Five-year Plan and the Long-term Target of Social and Economic Development of China" in 1995, this energy development strategy was repeated. Later, there were appeals from researchers to upgrade the energy conservation strategy to the level of being a basic national policy. They encouraged the formation of a basic national policy of "Controlling Population, Conserving Resources, Protecting the Environment" in the coming new development period, to guarantee a sustainable development in China. In 1997, ECL was approved. Energy conservation could be considered, in a broad sense, as a strategy to improve energy system efficiency, to save raw materials of production, labor force and funds, to slow down population expansion, to increase productivity and benefits of industrial enterprises, and to make reasonable adjustment and optimization of industry mix and product mix (Wei, 2004). In fact, on the supply side, China has undertaken a campaign to close or curtail production at thousands of coal mines. On the demand side, the slowing economic growth and reforms in the state-owned sectors led to reductions in output and closures of many enterprises, which in turn decreased the energy demand. Households have continued their rapid switch from coal to gas fuels and electricity. Moreover, across all sectors, equipment turnover has continuously raised efficiency. Trade policies were also affected by energy policies. Importing more energy intensive products saved more domestic energy consumption (Garbaccio,1999).

More specifically, as discussed in Wei (2004), studies on the situation of energy conservation in the years of 1996-2000 reveal some preliminary findings that about 80% of energy saving and energy use reduction were attributed to effective measures: reconstruction of national economic structure such as industry, sector, and product adjustment, enhancement of energy management, imports of energy and energy-intensive products. 20% of energy saving were directly derived from construction of new large-scale power generation units, co-generation, technical renovation in energy intensive sectors of iron and steel, building materials, chemistry, and the retrofit of power grids in urban and rural areas.

Part of the sustainable economic development strategy is "Environmental Protection". The combustion of coal leads to serious environmental problems, which attracts more and more domestic and international attention. The reduction in coal consumption can also be attributed to environment related policies. The "Environment Protection Law" was passed in 1989, and the Chinese government has been trying to control pollution, which is entangled with energy consumption.

The above reasons may explain why there is a clear drop in energy consumption from 1997. After 2001, with recovered economic growth, energy demand went up high and consumption increases.

1.2 Future Trend: Energy Development and Substitution

Oil resources in China are very limited considering the rapid increase demand in the past decade. Oil imports occupied about 30-40% of total oil demand in recent years, and the proportion will likely to increase, e.g. more than 80% in 2030 according to IEA's forecast. Thus, China's energy source, especially for the source of oil, will depend largely on imports.

Besides renewable energy, in substituting coal for improving energy efficiency and reducing pollution, natural gas and nuclear power are the most feasible options in China (Gan, 1998). China's total natural gas resources are estimated to be 38 trillion cubic meter, of which 30 trillion are in the mainland, and8 trillion under the sea. Although the proven reserve is only about 3% of total resources, it shows a bright prospect for developing natural gas in the long-term. On the one hand, China has approved a project of "Transporting Natural Gas from West to East", because the West is where the resources are located, while the East is the developed and most energy-consuming part. On the other hand, China sought international co-operation for natural gas exploitation and later on imported natural gas. For example, China signed a government agreement with Russia to import natural gas and to build a pipeline from the natural gas fields in Keweikejin in Irkutsk to China. The project construction is expected to put into operation in 2010.

1.3 Literature Review

The previous overview and discussion of Chinese energy status present several issues. First, there is a natural conflict between the energy requirements for economic development and the limitations of domestic energy supply. Rapid economic growth, increasing demand for energy from industry and households, and sustainable development policies are tangled and fostered a number of studies on China's energy sector. Most of them have focused on policy issues (Yanrui Wu, 2002; Zhongxiang Zhang, 1995), energy intensity (Richard F. Garbaccio, Mun S. Ho and Dale W. Jrgenson 1999 , Zhongxiang Zhang, 2003), and energy forecast (Hing Lin Chan, Shu Kam Lee, 1997; Paul Crompton, Yanrui Wu, 2004).

Second, there are concerns about pollution accompanied by the use of energy, the corresponding deterioration of the environment, and international pressure to abate emissions. Relevant research includes abatement in emission and pollution (Solveig Glomsrod, Wei Taoyuan 2003; Lin Gan 1998).

Third, sustainable development requires energy conservation, efficiency, diversification and substitution. There is some research regarding the issues of energy conservation, energy efficiency (J. Sinton, M. Levine, D. Fridley, F. Yang, and J. Lin), and alternative energy possibilities (Eric D. Larson, Wu Zongxin, Pat DeLaquil, Chen Weiying, Gao Pengfei, 2003).

To the best of the author's knowledge, there has been no formal econometric research conducted on input factor demand and interfuel substitution in China's aggregate production sector. This is an important issue for understanding the Chinese economic structure and energy status. It is also important for the world energy market because of China's significant role in world energy market and demand. Furthermore, it is closely related to China's future energy demand and strategy.

1.4 Purpose and Structure of the Thesis

The purpose of this study is to econometrically measure production structure and energy substitution in China. We are interested in the extent of energy substitution among different types of energy and possible substitution with other factor inputs. Two perspectives have been used. The first approach starts from economic theory and proceeds to empirical results. The importance of duality between production and cost is discussed, flexible functional form is chosen, system estimation method is used and empirical results are developed. The second approach is less theoretical. It is the time-series approach, which uses the method of Vector Autoregression (VAR). This approach tries to uncover the data properties and adds more dynamic analysis. The purpose of using different perspectives and analyzing methods is to see whether consistent results can be achieved and whether we can discover any possible complementary or comparative results.

The structure of the thesis is as follows: In Chapter 2, the theoretical foundation used to measure the structure of production in China is presented. In particular, duality theory is explored to facilitate the indirect study of production through the cost function. Flexible functional form is introduced. Chapter 3 describes the econometric techniques that will be used in the analysis. Chapter 4 provides an overview of Chinese energy data and, this chapter also examines empirical results for energy substitution and substitution for other input factors, price elasticity, and concludes with discussion and analysis of the results. Chapter 5 introduces the rationale of Vector Autoregression (VAR). Impulse response function and dynamic correlation are constructed and calculated. Attempts to discover data properties have been made. Chapter 6 concludes and summarizes the major results of the research. Detailed empirical results are presented in the Appendix.

CHAPTER 2 THEORETICAL FRAMEWORK: DUALITY, ELASTICITY OF SUBSTITUTION AND FUNCTIONAL FORM

2.1 Duality Theory

2.1.1 Production Function

The production process can be considered as a projection from a set of inputs to a set of outputs. There are various combinations of inputs to give the desired outputs. We can summarize the feasible combinations of inputs and outputs by *production possibilities, input requirement set, isoquant, and production function.*

Suppose the firm has n possible goods for inputs and outputs, and they constitute the space R^n. The set of all feasible production plans constitute the firm's *production possibilities* (Y), and Y is a subset of R^n.

If x is a vector of inputs and y is the output. Denoting input as negative and output as positive, *input requirement set* is a special case of a restricted production possibilities set: V(y)={x in R^n_+: (y, -x) is in Y}. It gives all input bundles that produce at least Y.

The isoquant gives all input combinations that produce exactly y : Q(y)={x in R^n_+ : x is in V(y), x is not in V(y') for y'>y}

If there is one output, the *production function* can be defined as: f(x)={y in R, y is the maximum output associated with –x in Y}

As described in Varian (1978), there are some basic assumptions on input requirement set V(y):

(1) Regular. V(y) is not empty, y>0 implies 0 is not in V(y), and V(y) is a closed subset of R^n_+;

(2) Monotonicity. If x is in V(y) and x' ≥ x, then x' is in V(y);

(3) Strict convexity. So if x and x' can both produce y, then any weighted average of x and x' can also produce y.

The assumptions of monotonicity and convexity ensure the technology is well behaved, and they arise naturally under duality theory, which is to be discussed.

Empirically, how to estimate the production function? First of all, we need to specify a functional form, which satisfies the basic assumptions and which does not present estimation problems. We can start by the most widely used Cobb-Douglas production function, which satisfies basic assumptions. For Cobb-Douglas technology with inputs capital (x_1), labor (x_2), and material(x_3), the production function is $y=f(x_1, x_2, x_3)=A x_1^b x_2^c x_3^d$, where A, b, c, d are parameters. Take logarithm to both sides and the statistical model can be written as:

$$\ln y=a+b\ln x_1+c\ln x_2+d\ln x_3+\varepsilon$$

Where ε is a random variable with zero mean and constant variance.

It is tempting to estimate the above production function directly using ordinary least squares. However, this approach is only appropriate when the right-hand side variables are uncorrelated with the error term ε, or inputs can be considered exogenous, which is usually not the case. In fact, a firm typically selects input and output levels simultaneously and inputs may be correlated with the disturbance term. When input levels are endogenous, direct estimators of the production function may be inconsistent (Mundlak, 1996).

2.1.2 Duality between Production and Cost

Duality theory facilitates the analysis of production function through cost function. When certain conditions are met, production function and cost function can be uniquely represented by each other. More specifically, define the producer's (total) cost function C: $C(y, p)=\min_x \{p'x: f(x)\geq y\}$, where p' is a vector of input prices. Assuming that solutions x to the producer's cost minimization problem exist, irrespective of what properties of production function f(x) have, it turns out that the cost function C(y,p) satisfies the following regularity conditions (Diewert 1986):

(1) C is a nonnegative function. That is, $C(y,p) \geq 0$;

(2) Homogeneous of degree one in input prices for each fixed output level y. This condition implies that a proportional increase in all input prices

results in a proportionate increase in total cost. C(y,tp)=tC(y,p) for $t \geq 0$, $y \succ 0$, $p \succ 0$;

(3) C is nondecreasing in p for fixed y. An increase in input price will cause total cost to increase although the firm substitute away from higher priced inputs. Mathematically, if $p_2 \geq p_1 \succ 0$, $C(y, p_2) \geq C(y, p_1)$;

(4) Concave in p for fixed y. An increase in input price increases total cost, but at a decreasing rate as substitution occurs; i.e. for $p \succ 0$, $0 \prec t \prec 1$, C(y, $t p_1$+(1-t) p_2) $\geq$tC(y, p_1)+(1-t)C(y, p_2);

This restriction may be satisfies on a local basis or a global basis. Local concavity is a weaker restriction. If the intent is to predict behavior outside the observed region of the sample, it is more suitable to use global concavity (Gordon, 1996);

(5) Nondecreasing in the output level y for fixed input prices p; i.e., $y_2 \geq y_1$, $p \succ 0$ implies C(y_2,p)$\geq$ C(y_1,p).

It has been shown that the cost function C can be defined given the production function f. A reverse question is: Given a cost function satisfying regularity conditions (1)-(5), is it possible to define a underlying production function f?

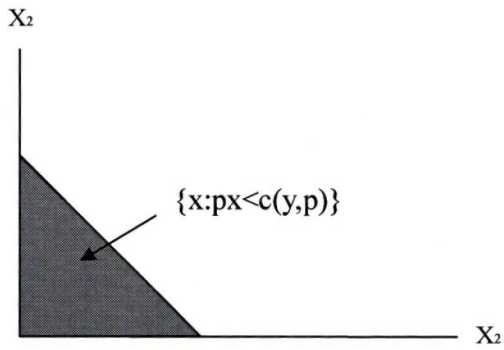
As shown in Varian (1978), the answer is yes, under some conditions. The intuition is shown in the following figures.

Define V*(y)={x in R^n_+ : px$\geq$C(y, p) for all p$\geq$0}

Graphically, as shown in figure 2.1, the shaded area denote px<c(y,p), and is therefore eliminated from constructing the production function.

Given a production input requirement set V(y), it is possible to construct V*(y) from the cost function by repeating the process of eliminating the shaded area in Figure 2.1.

Figure 2.1 Constructing a Technology from a Cost Function



What is the relationship between V(y) and V*(y)? Intuitively, it can be seen from the following figures.

Figure 2.2 Relationship between V(y) and V*(y)-1

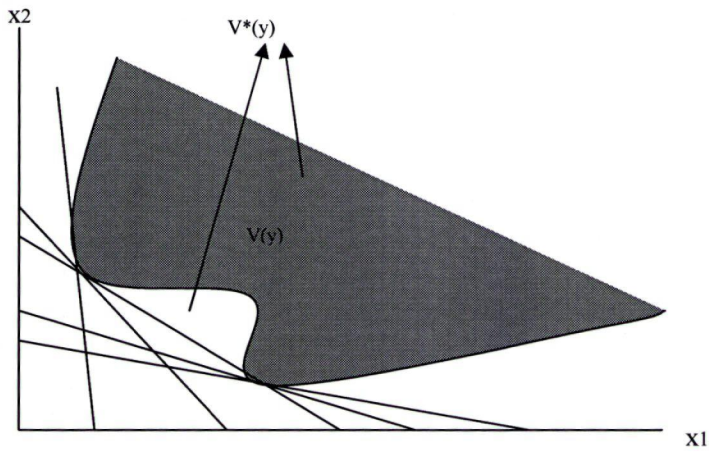
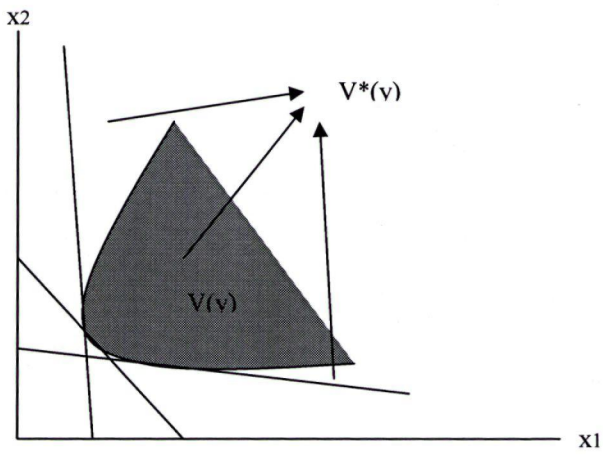
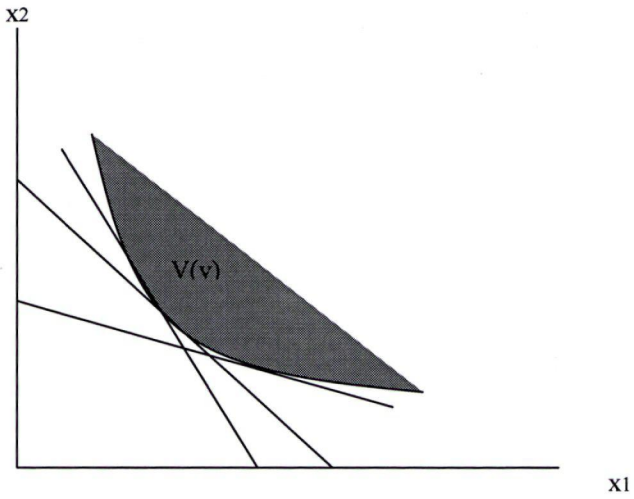


Figure 2.3 Relationship between V(y) and V*(y)-2



From figure 2.2 and 2.3, it can be seen that intuitively, V*(y) includes V(y). But they are not the same, depending on the shape of V(y). In fact, if V(y) is a regular, convex and monotonic technology, then V*(y)=V(y), as can be seen in Figure 2.4.

Figure 2.4 Proof that V*(y)=V(y) under Convexity Assumption



To this point, it is clear why there are basic assumptions of regularity, convexity and monotonicity on input requirement sets V(y) and how they arise naturally in the process of constructing production function from cost function. These regularity conditions on V(y) is equivalent to the regularity conditions on production function f(x):

(1) f(x) is nondecreasing;

(2) f(x) is a quasiconcave function; i.e. $V(y)=\{x: f(x)\geq y\}$ is a convex set.

The duality property between cost and production was first proved by Shephard (1953). The Shephard duality theorem states that technology may be equivalently represented by a production function satisfying the production function regularity conditions, or a cost function satisfying the cost function regularity conditions.

2.1.3 A Demonstration of Duality Using Cobb-Douglas Technology

In this section, Cobb-Douglas technology is used to demonstrate the concept of duality discussed in the previous section.

Cobb-Douglas technology with two inputs and a single output can be represented as:

$Y=\{(y, -x_1, -x_2) \text{ in } R^3 : y \leq x_1^{\,a} x_2^{\,b}\}$

It satisfies the two regularity conditions on production function. It can be seen how to derive cost function from the production function and vice versa, as shown in Varian (1978)

Under the assumption of cost minimization, the problem can be set as:

$C(p,y)=\min \; p_1 x_1 + p_2 x_2 \quad \text{s.t.} \; A x_1^{\,a} x_2^{\,b} = y$

This is equivalent to:

$$\min \; p_1 x_1 + p_2 A^{-\frac{1}{b}} y^{\frac{1}{b}} x_1^{-\frac{a}{b}}$$

The first order condition is:

$$p_1 - \frac{a}{b} p_2 A^{-\frac{1}{b}} y^{\frac{1}{b}} x_1^{-\frac{a+b}{b}} = 0$$

which gives the conditional demand for factor 1:

$$x_1(p_1, p_2, y) = A^{-\frac{1}{a+b}} \left[\frac{a p_2}{b p_1}\right]^{\frac{b}{a+b}} y^{\frac{1}{a+b}}$$

By substitution the conditional demand for factor 2 is:

$$x_2(p_1, p_2, y) = A^{-\frac{1}{a+b}} \left[\frac{a p_2}{b p_1}\right]^{-\frac{a}{a+b}} y^{\frac{1}{a+b}}$$

Therefore the cost function is:

$$C(p_1, p_2, y) = p_1 x_1(p_1, p_2, y) + p_2 x_2(p_1, p_2, y) =$$

$$A^{-\frac{1}{a+b}} \left[\left(\frac{a}{b}\right)^{\frac{b}{a+b}} + \left(\frac{a}{b}\right)^{-\frac{a}{a+b}}\right] p_1^{\frac{a}{a+b}} p_2^{\frac{b}{a+b}} y^{\frac{1}{a+b}}$$

We can measure units so that A=1. Under the assumption that a+b=1, the cost function reduces to:

$C(p_1, p_2, y) = K p_1^{\,a} p_2^{\,1-a} y \quad$ where $K=a^{-a}(1-a)^{a-1}$

Therefore, the cost function is derived from a production function for the

cost-minimizing firm.

Next, suppose a cost function C(p,y)= $yp_1^{\ a}p_2^{\ 1-a}$ is given. Is it possible to construct the associated technology?

According to the derivative property:

$$x_1(p,y) = ayp_1^{\ a-1}p_2^{\ 1-a} = ay\left(\frac{p_2}{p_1}\right)^{1-a}$$

$$x_2(p,y) = (1-a)yp_1^{\ a}p_2^{\ -a} = (1-a)y\left(\frac{p_2}{p_1}\right)^{-a}$$

Rearraging gives:

$$\frac{p_2}{p_1} = \left(\frac{x_1}{ay}\right)^{\frac{1}{1-a}}$$

$$\frac{p_2}{p_1} = \left(\frac{x_2}{(1-a)y}\right)^{\frac{1}{a}}$$

Setting these equal to each other and raising both sides to the –a(1-a) power,

$$\frac{x_1^{\ -a}}{a^{-a}y^{-a}} = \frac{x_2^{\ 1-a}}{(1-a)^{1-a}y^{1-a}}$$

or

$$\left[a^a(1-a)^{1-a}\right]y = x_1^{\ a}x_2^{\ 1-a}$$

which is the Cobb-Douglas technology.

Now we have shown the duality between production function and cost function. In empirical studies, this duality is important because it facilitates the study of production function through cost function.

2.2 Regularity Conditions and Imposition of Curvature

Regularity conditions are important to ensure the duality between production function and cost function. Positivity, homogeneity, monotonicity, concavity can all be tested. Empirically, curvature condition is the most possible condition to be violated. Sometimes Imposition of curvature is needed.

Concavity of cost function can be checked by negative semi-definiteness of

Hessian matrix. The Hessian matrix of a cost function is the square matrix of second partial derivatives of the cost function with respect to prices. If Hessian matrix is not negative semi-definiteness at any data point, curvature condition is violated. In a recent article, Ryan and Wales (2000) suggested a relatively simple procedure for imposing at a point (that is, locally) the curvature conditions implied by economic theory.

Diewert and Wales (1987, p. 48) show that the Hessian of the Translog cost function will be negative semidefinite, providing C(p, y, t)>0, if and only if the matrix G below is negative semidefinite. The ijth element of G, evaluated at t*, is defined as

$$g_{ij} = a_{ij} - a_i\delta_{ij} + a_i a_j \qquad \text{where} \quad i,j= C,O,L \qquad (2.1)$$

with $\delta_{ij} = 1$ if i=j and 0 otherwise.

As is argued by Ryan and Wales (2000), imposing global concavity on translog cost function destroys its flexibility. But it does not necessarily mean that the translog cost function should be abandoned. Instead, we can impose concavity at one chosen point, that is, impose concavity locally. Imposing curvature at the reference point is attained by setting G=-DD' with D triangular, and solving for A as

$$a_{ij} = -(DD')_{ij} + a_i\delta_{ij} - a_i a_j \qquad (2.2)$$

where $(DD')_{ij}$ is the ijth element of DD'. The elements of A are replaced in the estimating equations with the right hand side of (2.2) thus guaranteeing that G, and hence H, is negative semidefinite. It is easy to see that reparameterizing according to (8) will not destroy the flexibility of the functional form because the n(n-1) /2 elements of D just replace the n(n-1) /2 elements of A in the estimation. Negative semi-definiteness of G guarantees concavity of the cost function at the reference point.

2.3 Elasticity of Substitution

The concept "elasticity of substitution" measures the degree of substitutability

between pairs of factors, and was first introduced independently by John Hicks (1932) and Joan Robinson (1933). Hicks emphasizes that the pattern of income redistribution between two inputs due to a change in their relative price depends upon their substitution elasticity. Formally, the elasticity of substitution measures the percentage change in factor proportions due to a change in marginal rate of technical substitution, or a change in relative prices, since at cost-minimizing point given prices, price ratio equals to marginal rate of technical substitution. Hicks' key insight was that the effect of changes in factor price ratio on the distribution of income for a given output can be completely characterized by a scalar measure of curvature of the isoquant.

To demonstrate the concept in more detail, first consider two input case. For production function with two inputs capital and labor: Y =f(K, L), the elasticity of substitution between capital and labor is given by:

$$\sigma = d \ln (L/K) / d \ln (f_K / f_L) = [d(L/K) / d(f_K / f_L)] \cdot [(f_K / f_L)/(L/K)]$$

Intuitively, the elasticity of substitution is "a measure of the ease with which the varying factor can be substituted for others" (Hicks, 1932), and is a measure of the curvature of an isoquant (Lerner, 1933), as presented in Figure 2.5

Moving along the isoquant Y* from E to E', elasticity of substitution (σ) is represented by the change of labor-capital ratio (ΔR) divided by the change of marginal rate of substitution (ΔM), that is $\sigma = \Delta R / \Delta M$.(Refer to "The History of Economic Thought" Website.)

As an illustration, for Cobb-Douglas production function: $Y = f(K, L) = A x_1^{\ a} x_2^{\ b}$,

$$f_K = aY/K, \quad f_L = bY/L$$

Thus,

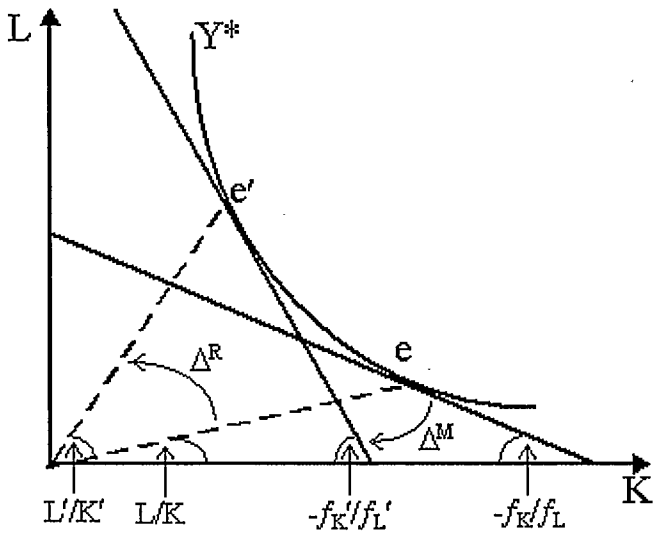
$$\frac{f_K}{f_L} = \frac{a}{b} \cdot \frac{L}{K}$$

Consequently,

$$\sigma = \frac{b}{a}.\left[\frac{a}{b}.\frac{L}{K}\right].\frac{K}{L} = 1$$

Therefore, the Cobb-Douglas production function gives constant elasticity of substitution and is equal to one.

Figure 2.5 Diagrammatic Representation of Elasticity of Substitution



In more general case, there are more than two inputs. When measuring elasticity of substitution, cross effects must be taken care of. Assume the production function with n inputs is $f(x) = f(x_1,....x_n)$, which is subject to constant return to scale and diminishing marginal rate of substitution. Allen (1938) proposed the Allen elasticity of substitution (AES), or partial elasticity of substitution as follows:

$$\sigma_{ij}^{A} = \frac{\sum_i f_i x_i}{x_i x_j}.\frac{F_{ij}}{F}$$

where $f_i = \frac{\partial f}{\partial x_i}$, $f_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}$, $F = \det\begin{bmatrix} 0 & f_1 & \cdots & f_n \\ f_1 & f_{11} & \cdots & f_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ f_n & f_{n1} & \cdots & f_{nn} \end{bmatrix}$,

F_{ij} is the co-factor of the element f_{ij} in the determinant F.

It is clear that in the two-variable case, AES coincides with the original two-input elasticity of substitution. AES is symmetric in the sense that $\sigma_{ij} = \sigma_{ji}$.

Alternatively, Allen elasticity of substitution can be expressed in terms of cost

function as (see Uzawa):

$$\sigma_{ij}^{A} = \frac{C(y,p)C_{ij}(y,p)}{C_i(y,p)C_j(y,p)}$$

By Shephard's lemma, $C_i(y,p) = \dot{x}_i$, where $\dot{x}_i$ is the optimal quantity of the ith input.

Blackorby C. and Russell R (1989) show that:

$$\sigma_{ij}^{A} = \frac{\varepsilon_{ij}}{s_j}$$

where $\varepsilon_{ij} = \frac{\partial \ln x_i}{\partial \ln p_j}$, which is the elasticity of demand for factor i with respect to the price of the factor j. $s_j = \frac{p_j x_j}{\sum_i p_i x_i}$, which is the jth factor's share of total cost.

However, as argued by Blackorby C. and Russell R (1989), while Allen elasticity of substitution (AES) reduces to original Hicksian concept in two-dimensional case, there is no intuition to believe that it's a generalization of the two-variable elasticity of substitution. In general it has none of the salient properties of the Hicksian notion. As a quantitative measure, it has no meaning; as qualitative measure, it adds no more information than cross-price elasticity given output. AES is completely uninformative. Nevertheless, AES has been a standard statistic reported in empirical studies.

Alternatively, they reason that Morishima elasticity of substitution (MES) originally formulated by M. Morishima (1967) is the natural generalization of the original Hicksian concept. More specifically, First, MES is a measure of curvature, or ease of substitution; Second, it is a sufficient statistic to assess the effects of changes in price or quantity ratios on relative factor shares; Third, it is a logarithmic derivative of a quantity ratio with respect to a marginal rate of substitution or a price ratio.

It is shown that MES can be expressed as:

$$M_{ij}(y,p) = \frac{p_i C_{ij}(y,p)}{C_j(y,p)} - \frac{p_i C_{ii}(y,p)}{C_i(y,p)}$$

$$= \varepsilon_{ji}(y,p) - \varepsilon_{ii}(y,p)$$

$$M_{ji}(y,p) = \varepsilon_{ij}(y,p) - \varepsilon_{jj}(y,p)$$

It is worth noting that MES is asymmetric. Consider the change of optimal quantity ratio $\frac{x_i}{x_j}$ induced by a change on relative price $\frac{p_i}{p_j}$. The same change in relative price can be caused solely by p_i, or solely p_j, but they usually have different effects on compensated demand, which result in asymmetry.

2.4 Functional Forms

Model selection is crucial in the fulfillment of theoretical conditions and the validity of empirical evaluation. In order to approximate the unknown dual cost function, properties of different functional forms need to be investigated.

As shown above, the most widely used Cobb-Douglas production function restricts elasticity of substitution equal to one. This nature of Cobb-Douglas production function is restrictive in the sense that there is no reason to believe at priori that the elasticity of substitution between input pairs is equal to one.

Another frequently used functional form is constant elasticity of substitution (CES). The n-variable CES function is:

$$f(x_1,...x_n) = (\alpha_1 x_1^{-\beta} + ... + \alpha_n x_n^{-\beta})^{-1/\beta}$$

It is shown that the MES, not the AES, is the elasticity of substitution, and is also restrictive in the sense that elasticity of substitution is equal between pairs of inputs.

Therefore, functional forms such as Cobb-Douglas and CES incorporate restrictions in their specifications, and are appropriate only if it is reasonable to believe these restrictions are valid. Contrary to the restrictive nature of the above functional forms, flexible functional forms pose no priori restrictions on elasticity of

substitution. Flexible functional form is defined as it "provides a second-order approximation to an arbitrary function at an arbitrary point in the sample space" (Diewert, 1974).

The translog cost function (Christensen et al. 1971, 1973) has been widely used in the literature. A general translog production function with n inputs can be written as:

$$\text{Ln(y)}= a_0 + \sum_{i=1}^{n} b_i \ln(x_i) + \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij} \ln(x_i)\ln(x_j)$$

Where x_i's are inputs. Y is level of output.

The corresponding cost function can be written as:

$$\text{Ln(C)}= \ln(\alpha_0) + \alpha_y \ln(y) + \sum_{i=1}^{n} \alpha_i \ln(w_i) + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_{ij} \ln(w_i)\ln(w_j)$$

$$+ \sum_{i=1}^{n} \alpha_{yi} \ln(y)\ln(w_i) + \alpha_{yy}[\ln(y)]^2$$

where w_i is the price of x_i.

For a well-behaved cost function, regularity conditions have to be satisfied. Positivity condition requires that $C \geq 0$. Homogeneity condition requires that: $\sum_{i=1}^{n} \alpha_i = 1$, $\sum_{i=1}^{n} \alpha_{yi} = 0$, $\sum_{i=1}^{n} \alpha_{ij} = \sum_{i=1}^{n} \alpha_{ji} = 0$. Concavity in prices can be checked by negative semi-definiteness of the Hessian matrix.

Combine theoretical reasoning with empirical tests, appropriate functional form can be chosen.

CHAPTER 3 SYSTEM OF EQUATIONS AND ESTIMATION PROCEDURE

3.1 System of Equations

The economic model specification is considered to be significant and yield valid estimates only if all the relevant information affecting dependent variable has been included in the variables specified in the model. In other words, all the left-out information is considered pure random and is included in the error term. If systematic relevant information is not considered in the estimation process, the estimated results are not significant. More efficient estimates can be gained by incorporating all systematic relevant information in the model specification and estimation process.

Duality theory allows the possibility of studying the primal production function through the dual cost function. From the dual cost function, all economic relevant information about production can be derived, such as marginal cost of production, elasticity of substitution, or economics of scale. It is possible to estimate the dual cost function alone. However, the single equation estimation may not be the optimal estimation strategy since there is more systematic information left out.

Under the assumption of cost minimization, factor demand functions can be derived through partial differentiation of the cost function with respect to input prices by Shephard lemma.

Assume the cost function is C=C(y,p) is a continuous and twice differentiable function. Then the factor demand xi is

$$x_i(p,y) = \frac{\partial C(p,y)}{\partial p_i}$$

where xi is the ith input demand, p is a vector of all input prices, pi is ith input price.

Input share equations can be derived as follows:

$$S_i = \frac{\partial \ln C(p,y)}{\partial \ln p_i} = S_i(p,y)$$

where $S_i = \frac{p_i x_i}{C}$ is the cost share of ith input.

Share equations are derived from the cost function and in inherently related to the firm's cost-minimization decision-making process. Furthermore, as discussed by Christensen and Greene (1976), including the cost share equation has the effect of adding many additional degrees of freedom without adding any unrestricted regression coefficients. This will result in more efficient parameter estimates then applying ordinary least squares to the cost function alone. Therefore, it is crucial to include the share equations in the estimation process of cost function.

3.2 Seemingly Unrelated Regression and Generalized Least Squares

It is likely that there exists contemporaneous cross-equation correlation in the system comprised of cost function and share equations. Firstly, there are cross-equation restrictions imposed by economic theory. Secondly, common parameters are present in different equations of the system, and it is likely that the disturbance terms contain factors not specified in the model, which result in the correlation of disturbance terms. Therefore, the equations in the system are "seemingly unrelated", and system estimation method is required.

To be more specific, consider the following system:

$$y_1 = X_1\beta_1 + \varepsilon_1$$

$$y_2 = X_2\beta_2 + \varepsilon_2$$

.....

$$y_M = X_M\beta_M + \varepsilon_M$$

Assume there are T observations used in estimating the M equations. Disturbances are uncorrelated across observations, that is:

$E[\varepsilon_{it}\varepsilon_{js} \mid X_1, X_2, ...X_M] = \sigma_{ij}$ for t=s and 0 for t≠ s

where i=1,...,M

Or

$E[\varepsilon_i\varepsilon_j' \mid X_1, X_2, ...X_M] = \sigma_{ij}I_T$

Or

$$E[\varepsilon\varepsilon' | X_1, X_2, ..., X_M] = \begin{bmatrix} \sigma_{11}I & \sigma_{12}I & ... & \sigma_{1M}I \\ \sigma_{21}I & \sigma_{22}I & ... & \sigma_{2M}I \\ & ... & ... & \\ \sigma_{M1}I & \sigma_{M2}I & ... & \sigma_{MM}I \end{bmatrix} = \Omega$$

where $\varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ ... \\ \varepsilon_M \end{bmatrix}$

The generalized least squares estimator is:

$$\hat{\beta} = [X'\Omega^{-1}X]^{-1}X'\Omega^{-1}y$$

where $X = \begin{bmatrix} X_1 & 0 & ... & 0 \\ 0 & X_2 & ... & 0 \\ & ... & ... & \\ 0 & 0 & ... & X_M \end{bmatrix}$, $\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ ... \\ \beta_M \end{bmatrix}$

At this point, the equations are related only through the disturbance terms. Zellner (1962) and Dwivedi and Srivastava (1978) have analyzed the efficiency gain through the GLS estimation. The greater the correlation of the disturbances or the less correlation there is between X matrices, the greater is the efficiency gain by using GLS. On the other hand, if the equations are unrelated, or the equations have identical explanatory variables, or the regressors in one block of equations are a subset of those in another, then OLS and GLS are identical, and there is no efficiency gain in GLS.

3.3 Zellner's Approach to Generalized Least Squares

The GLS estimator can only be realized if the error process, or the error covariance matrix is known. Unfortunately, the error covariance matrix is usually unknown, which makes the GLS estimation not feasible. In such cases, Zellner's (1962) iterative multi-stage estimation method can be used. The idea of this method is as follows: Firstly, the reduced form of the system model is estimated. Secondly, the fitted values for the endogenous variables are used to construct least square estimates for all equations in the system. Least squares residuals for each equation are also obtained. Thirdly, the proxy residuals are used in the determination

variance-covariance matrix. This matrix is employed to get coefficient estimates.

A consistent covariance estimator is:

$$\hat{\sigma}_{ij} = \frac{e_i' e_j}{T}$$

where $e_i = y_i - X_i \hat{\beta}_i$

The sample, or estimated covariance matrix is:

$$\hat{\Omega} = \frac{1}{T}\begin{bmatrix} \hat{\sigma}_{11} I & \hat{\sigma}_{12} I & \dots & \hat{\sigma}_{1M} I \\ \hat{\sigma}_{21} I & \hat{\sigma}_{22} I & \dots & \hat{\sigma}_{2M} I \\ & \dots & \dots & \\ \hat{\sigma}_{M1} I & \hat{\sigma}_{M2} I & \dots & \hat{\sigma}_{MM} I \end{bmatrix}$$

Therefore, when the error matrix is unknown, the FGLS estimator using this iterative approach is

$$\tilde{\beta} = (X'\hat{\Omega}^{-1}X)^{-1}X'\hat{\Omega}^{-1}y$$

Assuming appropriate model specification, this FGLS estimator is unbiased consistent, and exhibit asymptotic efficiency. The asymptotic efficiency of FGLS estimators may not carry over to small samples because of the variability introduced by the estimated Ω (Greene 2003).

The problem with this system estimation is that specification error in one equation would contaminate the whole system, and all parameter estimates in the other equations. Single equation estimation method confines such error in one equation only. Therefore, it is important to ensure the correlation of contemporaneous disturbance terms and the specification before pursuit of system estimation.

3.4 Maximum Likelihood Estimation

Maximum likelihood estimation estimates the parameters so that the possibility of getting the realized sample is maximized. The maximum likelihood estimators are obtained by maximizing the likelihood function. The likelihood function is defined as: $L = P(Y_1)P(Y_2)...P(Y_n)$, where Y is independently normally distributed.

The likelihood function for seemingly unrelated regression is:

$$L = \frac{-MT}{2}\log(2\prod) - \frac{1}{2}\log|\Omega| - \frac{1}{2}(y - X\beta)'\Omega^{-1}(y - X\beta)$$

The maximum likelihood estimators are:

$$\breve{\beta} = (X'\breve{\Omega}^{-1}X)^{-1}X'\breve{\Omega}^{-1}y$$

$$\breve{\Omega} = \begin{bmatrix} \breve{\sigma}_{11}I & \breve{\sigma}_{12}I & \dots & \breve{\sigma}_{1M}I \\ \breve{\sigma}_{21}I & \breve{\sigma}_{22}I & \dots & \breve{\sigma}_{2M}I \\ & \dots & \dots & \\ \breve{\sigma}_{M1}I & \breve{\sigma}_{M2}I & \dots & \breve{\sigma}_{MM}I \end{bmatrix}$$

If the model is correctly specified, and the disturbances are normally distributed, maximum likelihood estimators are consistent and asymptotically efficient. These estimators have the same asymptotic properties as the iterative generalized least squares method.

There are cases where OLS and SUR give equivalent estimation results. As shown by Dwivedi and Srivastava (1978), the first case is when system equations are not seemingly but actually unrelated, which makes the covariance matrix diagonal. The second case is when explanatory variables are identical.

CHAPTER 4 DATA OVERVIEW, MODEL SPECIFICATION AND EMPIRICAL RESULTS

4.1 Data Overview

GDP data is available from World development indicator. It is measured in constant local currency. The unit used in this study is billion yuan. Consumption (million tons of coal equivalent) of coal, oil, gas, and hydro electricity can be obtained from China statistical yearbook from 1960 to 2003. Nominal prices (yuan/ton) for coal and oil are available from 1960 to 2001. Consumer price index (CPI) is available from 1960 to 2003. Nominal energy (coal and oil) prices divided by CPI gives the real energy prices. The quantity of labor force is available from database of World Development Indicator, the World Bank Group. Ex-factory price indices of industrial products by sector, including coal industry, oil industry, and electricity industry are available from 1980 to 2003. Unfortunately, price for capital, price of gas and price for electricity from 1960 to 1979, cannot be found. Gross capital formation (constant local currency) is available in World Development Indicator by World Bank Group. The unit used in this study is 10 billion local currency.

In order to make the units of energy consumption and energy prices compatible, real energy prices are used to compute the growth rate of energy prices. Using the energy price indices in 1980 as the starting point, year 1979-1960's price indices are calculated backwards using the real growth rate of energy prices. After this calculation, energy price indices from 1960 to 2003 are available. Normalize the 1960 price index to be 1.

China is under state-controlled economy for decades. It has to be noted that the quality of data at early years like 1960s is questionable. In 1960s, China experienced several influential political issues, "Great Leap Forward", "Cultural Revolution", and also suffered from "The Three Year Natural Disaster", all of which severely influenced Chinese economy, government functioning, and quality of statistics. However, the data in the early years is not abandoned because we want the sample size to be larger, and do the best as possible to understand Chinese economic

structure, even if it may not be consistent with standard economic theory. It also has to be noted that before the reform of 1978, most economic activities are "planned". There were quota of consumption, targeted production, and controlled prices. The series of reforming policies from 1978 greatly stimulated market economy, economic activity started to be more on the track and statistics started to be more reliable.

4.2 Model Specification

Given the aforementioned points discussed in duality theory, direct estimation of production function is not utilized and cost function is preferred. Also can be seen from Chapter 2, traditional Cobb-Douglas and CES functional forms put prior restrictions on elasticity of substitution, which is arbitrary and not suitable for studying substitution possibilities. Furthermore, as discussed in Chapter 3, system estimation of the cost function and share equations would give more efficient estimates. Therefore, Translog cost function with system estimation method is chosen for this research. The software used in estimating system equations and calculating relevant variables is TSP4.5.

4.2.1 Translog Specification

To form a full, or long run, Translog cost function, prices for all inputs have to be available. However, capital price and labor price data are not available. Nevertheless, this problem is not insurmountable. The short run model, in the sense that capital and labor are fixed inputs and coal and oil are variable inputs, is formed according to Brown and Christensen (1981). In other words, it is still interesting to see how coal and oil substitute for each other when holding capital and labor fixed. In order to uncover the relationship between fixed and variable inputs, elasticity of intensity can be calculated.

To be more specific, the general production function with single output and multiple inputs is: $Y=F(X1,\ldots,Xn,t)$. As discussed before, if the production function has convex isoquants, and if for any level of output, the cost-minimizing input mix is employed, then there exists a total cost function that is dual to the original production

function, which can be described as: CT=G(Y, P1,...,Pn, t), CT=$\sum_{i=1}^{n} PiXi$.

If cost is minimized with respect to a subset of the factor inputs conditional on the level of output and the remaining inputs, then there exists a variable cost function that is dual to the original production function: CV=H(Y, P1,...,Pl,Z1,...,Zm, t), where Zi represent the subset of the Xi not necessarily in static equilibrium, l+m=n, and CV=$\sum_{i=1}^{l} P_i X_i$

The most general translog short-run cost function for the four inputs (Coal (C), Oil (O), labor (L), capital (K)) with labor and capital fixed can be expressed as:

$$\begin{aligned}
\ln C &= R_0 + R_Y \ln Y + \sum_i^l R_i \ln P_i + \sum_i^m R_i \ln Z_i + 0.5 R_{YY} (\ln Y)^2 \\
&+ 0.5 \sum_i^l \sum_j^l \gamma_{ij} \ln P_i \ln P_j + 0.5 \sum_i^m \sum_j^m \delta_{ij} \ln Z_i \ln Z_j \\
&+ \sum_i^l R_{iQ} \ln Y \ln P_i + \sum_i^l \sum_j^m \rho_{ij} \ln P_i \ln Z_j \\
&+ \sum_i^m R_{Yi} \ln Y \ln Z_i + R_T t + 0.5 R_{TT} t^2 \\
&+ R_{TY} t \ln Y + \sum_i^l R_{iT} t \ln P_i \\
&+ \sum_i^m R_{iT} t \ln Z_i + \varepsilon_c
\end{aligned} \tag{4.1}$$

where l=C,O, and m=L,K

C is the total variable cost, which is calculated as: $C = P_c * Q_c + P_O * Q_O$

Y is the output. Pi's are the prices of the variable factor inputs, coal and oil. T is an indicator of time. The terms containing T are technology terms.

Symmetry can be imposed on γ_{ij} and δ_{ij}, and theoretical restriction of homogeneity of degree one in input prices is required as follows:

$$\sum_i^l R_i = 1, \tag{4.2}$$

$$\sum_i^l \gamma_{ij} = \sum_i^l \gamma_{ji} = \sum_i^l R_{iQ} = \sum_i^l \rho_{ij} = \sum_i^l R_{iT} = 0, \quad \forall j. \tag{4.3}$$

Logarithmically differentiating (4.1) with respect to input prices yields

$$\frac{\partial \ln C}{\partial \ln Pi} = \frac{PiXi}{Ci} = S_i \tag{4.4}$$

where *Si* indicates the cost share of the *i*th input factor. The translog cost function thus yields the cost share equations:

$$S_i = R_i + R_{iQ} \ln Y + \sum_j^l \gamma_{ij} \ln P_j + \sum_j^m \rho_{ij} \ln Z_j + R_{iT} t + \varepsilon_i \tag{4.5}$$

where i=C, O.

The parameters of the cost function can be estimated by performing multivariate regression of the system (4.1) and (4.5). Efficient estimates can be obtained using Zellner's Approach to Generalized Least Squares. Since the cost share equations sum to unity, and the estimated covariance matrix is singular, one of the share equations should be deleted when conducting system estimation. The estimates obtained are independent of which equation is deleted, and the estimates are asymptotically equivalent to maximum likelihood estimates.

The translog cost function given by (4.1) does not constrain the production structure nor does it impose restrictions on the elasticities of substitution. However, these restrictions can be tested statistically. The elasticities of substitution can be restricted to unity by eliminating the second-order terms in the prices from the translog cost function. Hence, restrictions for the unitary elasticity of substitution are $\gamma_{ij} = 0$. If these restrictions are valid, the model reduces to a Cobb-Douglas functional form. The restriction of no technological change can be tested as $R_t = R_{tt} = R_{it} = R_{tY} = 0$. The validity of restrictions can be tested using likelihood ratio test.

4.2.2 Model Selection and Testing for Restrictions: Likelihood Ratio Test

Likelihood ratio test (LRT) is used to test whether one simple model is nested in another more complicated, or general model. The term "Nested" is used in the sense that the more complex model must differ from the simple model only by the

addition of one or more parameters. It is known that adding additional parameters will always result in a higher likelihood score. However, there comes a point when adding additional parameters is no longer justified in terms of significant improvement in fit of a model to a particular dataset. The LRT provides one objective criterion for selecting among possible models. The LRT begins with a comparison of the likelihood scores of the two models:

LR = 2*(lnL1-lnL2)

Where lnL1 is log likelihood of model 1 and lnL2 is log likelihood of model 2. This LRT statistic approximately follows a chi-square distribution. To determine if the difference in likelihood scores of the two models is statistically significant, degrees of freedom must be considered. In the LRT, degrees of freedom is equal to the number of additional parameters in the more complex model. Using this information we can then determine the critical value of the test statistic from standard statistical tables

Therefore, starting from the unrestricted model specified as (4.1), the restrictions of unitary elasticity of substitution, no technical change, can be tested using likelihood ratio test. After the tests of these restrictions, combination of restrictions can be constructed.

4.2.3 Elasticity of Substitution

Uzawa (1962) has shown that the Allen elasticities of substitution can be computed from the partial derivatives of the cost function. For the cost function (4.1), it is computed as:

$$\sigma_{ij} = \frac{C * C_{ij}}{C_i * C_j}$$

where $C_i = \frac{\partial C}{\partial P_i}$.

Berndt and Wood (1975) have shown that Allen partial elasticities of substitution (AES) and price elasticities of demand for the input factors (ε), respectively, can be calculated as

$$\sigma_{ij}=\frac{\beta_{ij}+S_iS_j}{S_iS_j} \text{ for } i \neq j$$

and

$$\sigma ii=\frac{\beta_{ii}+S_i^2-S_i}{S_i^2}$$

and

$\varepsilon ij = Sj\sigma ij$
$\varepsilon ii = Si\sigma ii$

Blackorby and Russell (1989) demonstrated that the Morishima (1967) elasticity of substitution (MES) preserves the salient features of Hicksian concept in the multifactor context. In particular, it measures the ease of substitution, and it is a sufficient statistic for assessing the effects of changes in price or quantity ratios on relative factor shares. In addition, MES is a logarithmic derivative of a quantity ratio with respect to a marginal rate of substitution or a price ratio. According to Blackorby and Russell (1989), the MES can be calculated as $Mij = \varepsilon ji - \varepsilon ii$ and $Mji = \varepsilon ij - \varepsilon jj$.

4.2.4 Elasticity of Intensity

Allen Partial Elasticity of Substitution gives the substitutability between coal and oil, but it's impossible to calculate it for the fixed inputs-labor and capital. In order to uncover the substitution relationship between variable and fixed inputs, elasticity of intensity defined by Diewert (1974) can be calculated. Elasticity of intensity is given as:

$$e_{lm}=\frac{\partial \ln X_l}{\partial \ln Z_m}$$

e_{lm} measures the percentage increase of input l due to a percentage increase of input m. It can be calculated from the cost function as: $e_{lm}=\frac{\partial^2 C}{\partial P_l \partial Z_m}.\frac{Z_m}{X_L}$.

Therefore, the substitutability between variable and fixed inputs can be calculated.

4.3 Empirical Results

4.3.1 Model Selection

Likelihood Ratio Test is conducted to test for validity of restricted models. The results are summarized in table 4.1.

Table 4.1 Likelihood Ratio Test Results

Model	Log Likelihood	No. of Restrictions	Calculated Chi-square	Critical Value
General	206.662			
No Technical Change	145.413	6	122.498	12.59
Unitary Elasticity of Substitution	95.8545	6	221.615	12.59

The restrictions of no technological change and unitary elasticity of substitution are rejected. Therefore the Cobb-Douglas functional form is not accepted and there is technological change. The model specified in (4.1) is used for empirical modeling.

4.3.2 Results of System Estimation

As shown before, it is possible to estimate the parameters of the cost function using ordinary least squares (OLS), but that neglects the information contained in the cost share equations. An alternative estimation procedure is to estimate the cost shares as a multivariate regression system (Berndt and Wood 1975). However, in this approach several parameters (R_t, R_{tt}), which are found only in the cost function, cannot be estimated. The optimal procedure is to jointly estimate the cost function with the cost share equations as a multivariate regression system. Inclusion of the cost share equations in the estimation results in more degrees of freedom without adding any unrestricted regression coefficients, and results in more efficient parameter estimates. The system estimation results are shown in Table 4.2. All the coefficients are significant.

Table 4. 2 Translog Cost Function Estimation Result

Parameter	Estimate	Standard Error	t-statistic	P-value
R_0	-1544.18	362.494	-4.25989	[.000]
R_1	-2.73614	.963391	-2.84011	[.005]
R_3	504.446	123.605	4.08112	[.000]
R_4	51.8900	11.3682	4.56451	[.000]
R_{YY}	4.99128	1.63148	3.05935	[.002]
γ_{11}	.177243	.025869	6.85142	[.000]
δ_{33}	-79.6757	21.7484	-3.66352	[.000]
δ_{34}	-13.5192	3.92350	-3.44570	[.001]
δ_{34}	1.13723	.257288	4.42007	[.000]
R_{1Q}	.414053	.046999	8.80977	[.000]
ρ_{13}	.285429	.169507	1.68388	[.092]
ρ_{14}	-.127073	.018052	-7.03936	[.000]
R_{Y3}	-3.56008	1.32803	-2.68072	[.007]
R_{Y4}	-2.82626	.531055	-5.32198	[.000]
R_T	-13.6052	2.74145	-4.96275	[.000]
R_{TT}	-.055145	.012085	-4.56325	[.000]
R_{TY}	-.139221	.065904	-2.11247	[.035]
R_{1T}	-.030487	.384645E-02	-7.92594	[.000]
R_{3T}	2.32751	.442107	5.26458	[.000]
R_{4T}	.293434	.043377	6.76469	[.000]

4.3.3 Regularity Conditions

As discussed before, regularity conditions are crucial in the validity of duality of cost and production. Positivity condition is satisfied since the cost at every sample point is positive. Monotonicity condition can be checked by fitted cost share. The fitted share at every sample point is positive, so monotonicity condition is also satisfied. Details are in Appendix A1. The most important regularity remain is the curvature condition.

Curvature condition can be checked by the negative semidefinteness of Hessian matrix of the cost function. As can be seen in Appendix A2, maximum eigen values for years 1960-1999 are positive and maximum eigen values are 0 from year 2000 on, which means the eigen value from 2000-2003 are either negative or zero in the last few observation points and satisfies the curvature condition. This is a very interesting result. It is not surprising to discover that Chinese economy is not consistent with what economic theory would require. In the early years of the sample, Chinese economy is under strict state control. The economic assumptions such as minimization of cost may be violated, which results in the empirical results not consistent with economic theory. On the other hand, curvature condition, or concavity in prices, is based on the theoretical belief that increased input price increases total cost, but at a decreasing rate as substitution occurs. More specifically in this case, increased price of coal would lead to a substitution from coal to oil, and vice versa. Violation of curvature condition indicates that Chinese economy does not behave according to this belief. For most of the sample points (1960-1999) under study, this substitution does not occur. This counter-intuitive result may be attributed to the economic and infrastructure structure, and the substitution between coal and oil is not free and easy.

However, it is significant to note that the curvature condition is satisfied for the years 2000-2003. Although it is a small portion of the total sample, it is likely to be an indication that Chinese economy is getting more close to standard economic theory. Some important economic reforms started in 1978, and after years of transition, Chinese economy is getting on the way.

The economic reforms commencing 1978 start to stimulate economy, introduce market economy and competition. Although there are still state-owned organizations, policy encourage private-owned companies, foreign investors. Facing competition from various new sources, state-owned organizations have to reform and find their new ways of operation. After some transition period since 1978, it is reasonable to believe that the recent years' Chinese economy is more inclined to comply with economic theory. In terms of energy sector, no substantial substitution between coal and oil was discovered before 1999, and substitution starts to occur afterwards. This energy sector change may be a part of the whole picture of economic reform.

Therefore, although most sample points violate curvature condition, it is still significant to know that at the end of the sample points, curvature condition is satisfied. It gives us a better picture of Chinese economic reform, economic structure, and energy substitution.

4.3.4 Elasticity of Substitution and Price Elasticity

Qualitatively, we have learned that until recently, Chinese economy may not be consistent with standard economic theory. After the year 2000, all economic assumptions of a well-behaved cost function are satisfied. It is not just a coincidence, but a consequence of economic reform and years of transition period.

Quantitatively, it is significant to calculate the ease of substitution and other economic relevant information. AES, price elasticity between coal and oil, and own price elasticity are calculated at every sample point, as can be seen in Appendix A3 and A4. The points that satisfy the curvature condition is further reported in the following tables.

Table 4.3 AES between coal and oil from 2000 to 2003

Year	AES
2000	0.10005
2001	0.10423
2002	0.066236
2003	0.12712

From Table 4.3, it can be seen that AES between coal and oil from 2000 to 2003 are positive and range from 0.066 to 0.127. The positivity of AES indicates that coal and oil are substitutes. Quantitatively, the quantity ratio of coal to oil would be changed 0.066 to 0.127 units when price ratio (or marginal rate of substitution) changes one unit. The relative low value indicates the AES between coal and oil is inelastic, and is not easily substituted for each other.

Table 4.4 Price Elasticities between Coal and Oil

Year	Coal-Coal	Oil-Oil
2000	-0.026980	-0.073069
2001	-0.028315	-0.075911
2002	-0.016869	-0.049367
2003	-0.036018	-0.091104

As indicated in Table 4.4, from 2000 to 2003, own price elasticity for coal ranges from -0.01 to -0.03. Own price elasticity for oil ranges from -0.04 to -0.07. Own price elasticity for oil is more elastic than coal.

Table 4.5 Elasticities of Intensity

	Coal-Labor	Coal-Capital	Oil-Labor	Oil-Capital
2000	1.05920	0.83258	0.83966	1.45341
2001	1.05473	0.85606	0.85328	1.38589
2002	1.04910	0.87377	0.85631	1.36941
2003	1.04990	0.88657	0.87377	1.28692

Table 4.5 displays the elasticity of intensity between fixed inputs-labor, capital, and variable inputs-coal, oil. Elasticity of intensity between coal and labor is elastic, so is it between oil and capital. In contrast, elasticity of intensity between coal and capital, oil and labor is inelastic. Elasticities of intensity between coal and labor, coal and capital, oil and labor, oil and capital are around 1.05, 0.8, 0.8, and 1.4 respectively.

To sum up, empirical results indicate that the hypothesis of unitary elasticity of substitution and no technical changes are rejected. Regularity conditions are all satisfied except that the curvature condition. Interestingly, the curvature condition is satisfied at the end of the sample. AES between coal and oil ranges from 0.066 to 0.127 in the years 2000 to 2003. Own price elasticities of coal ranges from -0.01 to

-0.03. Own price elasticities of oil ranges from -0.04 to -0.07. In general the demand of energy is not very sensitive to the price changes, which is in consistency with some other research (A. Austin, 2005). As Austin and some other research argued, the Chinese energy market may create inefficient outcomes, and "price signals have negligible effect on consumer behavior and investment". In this research, the violation of the curvature condition may also be an indication of inefficient market mechanism, or at least, the market is not behaving as the standard economic theory would expect. The satisfaction of all regularity conditions for recent years may be an indicator of the adjusting process, or a more efficient market. In fact, the Chinese government has been dedicating to develop a more efficient energy market, those efforts can be seen by the advocacy of market economy, enforcement of Energy Conservation Law, experiments with marketised pricing for electricity, and industrial reform.

CHAPTER 5 MULTIEQUATION TIME-SERIES ANALYSIS

5.1 Motivation for Multiequation Time-Series Analysis

From chapter 2 to chapter 4, theoretical-empirical approach is used. Duality theory and system estimation are utilized for translog cost function. Counter-intuitive result, such as positive own price elasticity, is not uncommon in the literature. As argued, the counter-intuitive result might because of the violations of economic assumptions, or policy-related issues. All the reasons are entangled and embedded in the data available.

Therefore, why not let the dataset speak for itself? Aside from economic theory and assumptions, what information does the data tell? Vector Autoregression (VAR) modeling strategy can be used for this purpose. Error Correction Model (ECM) modeling strategy can be used if there is evidence of cointegration. VAR or ECM modeling strategy can uncover more dynamic relationship between variables, which is not studied in translog cost function. More detailed description of the method and the model used in this chapter can be found in Walter Enders (2004).

Another motivation is due to the limited study in electricity sector. As is described in the data section, due to the lack of observations in electricity sector, it is not included in the translog cost function. Attempt to study the electricity sector is made in this chapter. All estimations are produced by software RATS 5.

5.2 The Stochastic Properties of the Data

Before proceeding to the VAR or ECM modeling, it is important to uncover the stochastic properties of the data. All variables are tested for unit roots using standard Dickey-Fuller tests, where the null hypothesis (H_o) is a unit root (I(1)), tested against the following alternative hypotheses (H_A): 1) I(0) with drift; 2) I(0) around a linear trend; 3) I(0) around a linear trend with a (one time) break in the intercept. Detailed results are presented in Appendix A6. Table 5.1 summarizes the results.

The testing for unit root is important because the econometric consequences of nonstationarity can be quite severe, leading to least square estimators, test statistics and predictors unreliable.

Table 5.1 Results for Unit Root Test

H_o: unit root (I(1))			
Variable	H_A: I(0) with drift	H_A: I(0) with trend	H_A: I(0) with break
G	1.5172	-5.5697*	-5.67412*
D_c	-1.2773	-3.1586	-5.31329**
P_c	0.5347	-1.4741	-6.12039*
D_o	-2.3744	-2.2228	-6.68053*
P_o	0.6158	-0.9386	-3.91042
D_E	-0.3744	-3.4681**	-4.28017
P_E	-3.2420*	-3.1476	-8.42857*

*Indicates rejection of null at 5% significance level

** Indicates rejection of null at 10% significance level

From Table 5.1, it can be concluded that the hypothesis of unit root is rejected for GDP, price of oil, demand of oil, and price of electricity at 5% significance level. Unit root is rejected for coal demand and electricity demand at 10% significance level. Unit root cannot be rejected for oil price. However, it has to be noted that the power of the various Dickey-Fuller tests can be very low. As such, it will too often indicate that a series contains a unit root (Walter Enders, 2004). Generally speaking, there is no overwhelming evidence of unit root in the date series.

5.3 VAR, Impulse Response Function, and Dynamic Correlation

A Vector Augoregression (VAR) analysis is "a-theoretical" in the sense that other than in choosing which variables should be used, there is very little in the way

of economic theory that guides a VAR and all variables are considered to be endogenous and treated systematically. For the purpose of this study, the variables of energy demand (D), GDP (Y) and energy price (P) are used for each energy type (that is, coal, oil, and electricity), and a three variable VAR system can be written as:

$$\begin{aligned} D_t &= \alpha_1 + \sum_{i=1}^{k} \beta_{1i} D_{t-i} + \sum_{i=1}^{k} \beta_{2i} Y_{t-i} + \sum_{i=1}^{k} \beta_{3i} P_{t-i} + \varepsilon_{1t} \\ Y_t &= \alpha_2 + \sum_{i=1}^{k} \beta_{4i} D_{t-i} + \sum_{i=1}^{k} \beta_{5i} Y_{t-i} + \sum_{i=1}^{k} \beta_{6i} P_{t-i} + \varepsilon_{2t} \\ P_t &= \alpha_3 + \sum_{i=1} \beta_{7i} D_{t-i} + \sum_{i=1} \beta_{8i} Y_{t-i} + \sum_{i=1}^{k} \beta_{9i} P_{t-i} + \varepsilon_{3t} \end{aligned} \tag{5.1}$$

where k is the lag length, which can be determined by the use of some systems model selection criterion, such as the Aikaike Information Criteria (AIC) or Schwarz information criteria (SIC). The error terms are assumed to be serially uncorrelated. One of the major advantages of a VAR is that right hand side variables are pre-determined, or lagged values, simultaneous equation bias problem can be avoided. Thus, efficient estimates can be obtained by using ordinary least squares estimation method for each equation.

As can be seen in system of equations (5.1), a VAR has a very rich set of dynamics, which is not captured in previous chapters. After some mathematical transformation of (5.1), the VAR have a moving-average representation as follows (Walter Enders, 2004):

$$\begin{bmatrix} D_t \\ Y_t \\ P_t \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} + \sum_{i=0}^{\infty} \begin{bmatrix} \phi_{11}(i) & \phi_{12}(i) & \phi_{13}(i) \\ \phi_{21}(i) & \phi_{22}(i) & \phi_{23}(i) \\ \phi_{31}(i) & \phi_{32}(i) & \phi_{33}(i) \end{bmatrix} \begin{bmatrix} \varepsilon_{1,t-i} \\ \varepsilon_{2,t-i} \\ \varepsilon_{3,t-i} \end{bmatrix} \tag{5.2}$$

The nine sets of coefficients $\phi_{11}(i)$, etc, are the impulse response functions. The dynamics are most often uncovered through impulse response functions, which give the effects of the shocks, ε_{1t}, ε_{2t} and ε_{3t} on the entire time paths of the entire sequences of $\{D_t\}$ $\{Y_t\}$ and $\{P_t\}$. After n periods, the cumulative sum of the effects of, say, ε_{3t} (which can be thought of as a shock to P_t) on D_t is given by

$$\sum_{i=0}^{n}\phi_{13}(i) \qquad (5.3)$$

The elasticity of response of one variable to another can be calculated as the dynamic correlation. Dynamic path of the elasticity of response of demand to a shock in P_t is given by the dynamic correlation

$$\rho(i)=\frac{\sum_{i=0}^{n}\phi_{13}(i)}{\sum_{i=0}^{n}\phi_{33}(i)} \qquad (5.4)$$

This dynamic correlation gives the price elasticity of demand over time.

Dynamic correlation can be calculated from impulse response functions. Detailed calculated impulse response function can be found in Appendix A7.

Figure 5.1 Dynamic Correlation for Coal:

Response of Demand to a Shock in Price

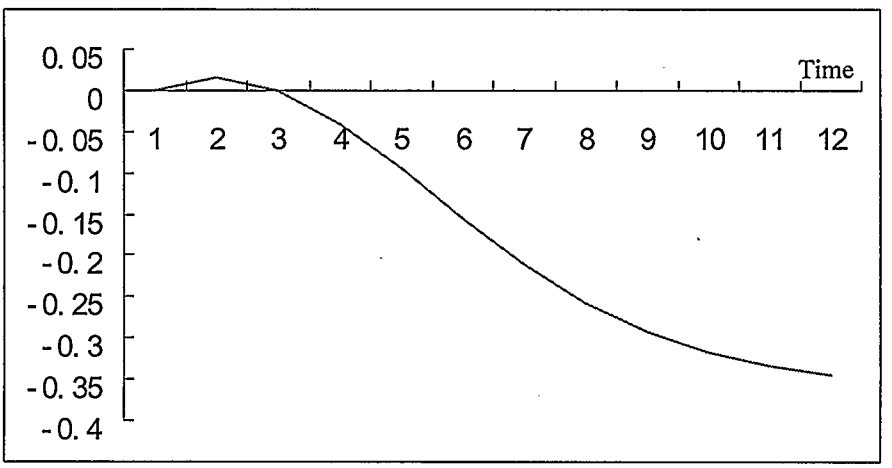


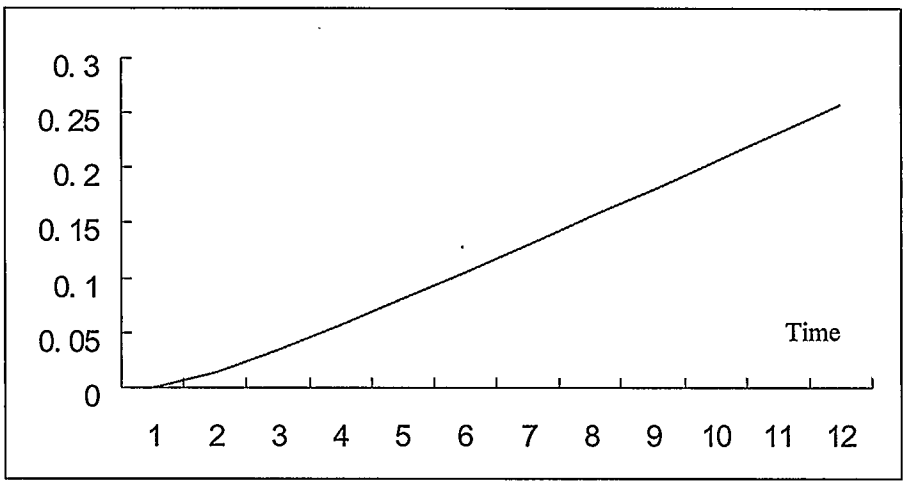
Figure 5.1 indicates that price elasticity of demand is initially slightly positive. As time goes by, the cumulative price elasticity of demand is negative and the quantity is around –0.35. It may imply that coal demand is not very sensitive to coal price at the beginning but the cumulative effect would be negative as time goes by.

Figure 5.2 indicates that demand response of oil to a shock in price is positive and increasing over time. This nature of the is consistent with the violation of curvature in chapter 4, as can be seen, the own price elasticity of demand for oil is

positive in Translog cost function for most sample points. Although at the end of the sample points, the properties discovered by VAR estimation are still dominated by the most sample points, which is inconsistent with economic theory. The consistent result of positive own-price elasticity may result from the possibility of exogenous oil prices. Also, as argued before, it might be an indication of violation of standard economic theory. The demand for oil is also not sensitive to oil price.

Figure 5.2 Dynamic Correlation for Oil:

Response of Demand to a Shock in Price

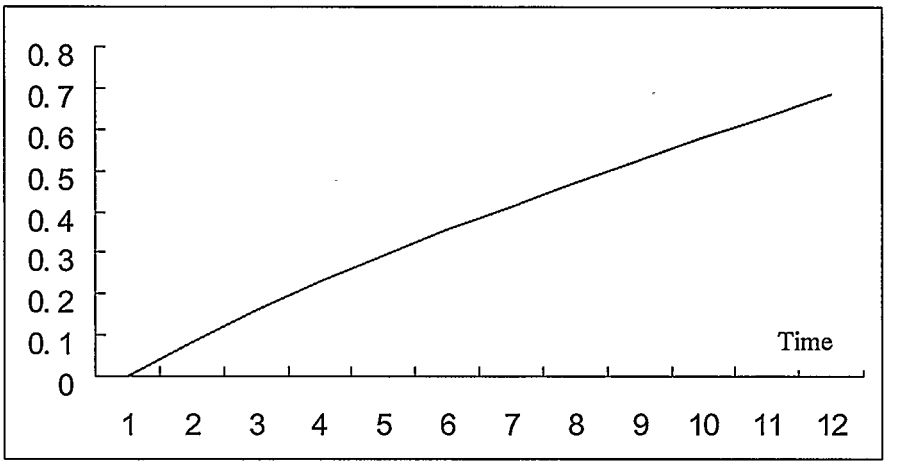


Electricity is not studied in Chapter 4 due to the unavailability of data. Using the electricity data from 1980 to 2003, VAR model is formed and dynamic correlation is shown in Figure 5.3

Figure 5.3 indicates that price demand response of electricity to a shock in price is positive and increasing over time. It may as well be attributed to the insensitivity of demand to price. However, Chinese government has started reforms in power sector also, future directions of power market would be largely influenced by the new policies.

Figure 5.3 Dynamic Correlation for Electricity:

Response of Demand to a Shock in Price



Overall, the findings of VAR estimation strategy are consistent with the findings by using duality and system estimation. After all, for majority time of the study period, it is very possible that standard economic theory is not satisfied. But it does not render the research pointless. For one thing, we understand the unique property of Chinese economy. For another thing, combined with previous finding, it's very important to know that China may have positive own-price elasticity for energy before, but the trend starts to change. That is, when applying for the insensitivity of energy demand to energy price, it has to be noted that negative own price elasticity was present in the recent years.

CHAPTER 6 CONCLUSION

6.1 Summary of the Research

With rapid economic growth, China is the second largest energy consumer in the world. Recently, more and more attention has been focused on China's energy issues. However, no research has been done in the topic of energy substitution in terms of using econometric analysis and flexible functional form. It is important to know the ease of substitution between different types of energy and input factors, both for China's energy policy and for international energy markets. This motivation initiates the research.

Two approaches have been used in this research. In the first approach, the theory of duality between cost and production is used due to: 1) Drawbacks in estimating production function and; 2) Advantages in estimating cost function and using system estimation. Translog functional form is tested against Cobb-Douglas and is shown to be superior. The advantage of Tranlog functional form is that it is flexible, and therefore do not restrict the substitutability between input factors to be one. In the second approach, the modeling strategy of Vector Autoregression(VAR) is used. The purpose to form a VAR is to let the data speak, and uncover the dynamic properties that are not covered in translog cost function.

Empirical results of the Translog cost function show four major results: first, the curvature condition is violated at the beginning of the sample points, but satisfied in the last few years of the sample, from 2000 to 2003. This result is significant since the Chinese government has been devoted to economic reform since 1980s. A more open economy, market power, and competition were introduced since then. Increasing attention has been put on the energy sector since the late 1990s. It is very possible that, after decades of economic reform and adjusting, and especially recent years' reform in the energy sector, the Chinese economy and energy sector is moving toward a more marketised direction and is behaving more consistently with standard economic theory.

Second, coal and oil were not substitutes until recently. The Estimated Allen

Partial Elasticity of Substitution ranges from 0.066 to 0.127 in the last few years. The ease of substitution of the major fuels, coal and oil, is still low. In earlier years, this substitutability was not discovered, which may be caused by the economic structure and technological infrastructure.

Third, in recent years, own price elasticities of coal ranges from -0.01 to -0.03. Own price elasticities of oil ranges from -0.04 to -0.07. The demand of energy is not very sensitive to price changes.

Fourth, elasticities of intensity between coal and labor, coal and capital, oil and labor, and oil and capital are around 1.05, 0.8, 0.8, and 1.4 respectively.

The strategy of VAR is more dynamic in the modeling process and gives the following results: Price elasticity of coal demand is initially positive and then became negative and increasing over time. But price elasticity of oil is positive and increasing over time. These results are slightly different from, but not inconsistent with, the Translog estimation. After all, for most of the data points, it is very possible that the standard economic assumptions are not satisfied. Therefore the positive own-price elasticity is not a surprise.

6.2 Limitations and Additional Research Opportunities

It has to be acknowledged that the research still has some limitations. The first limitation is that capital price data is not available, so the long-run Tranlog model cannot be specified. But the Translog model with labor and capital fixed also gives valuable information in the topic of interest. The second limitation is that Translog is a locally flexible functional form. It may be interesting to use some other flexible functional forms, such as Normalize Quadratic, to see whether results differ by switching functional forms. The third limitation is the degrees of freedom. Although system estimation method is used, the number of parameters for estimation and observation points may still imply degrees of freedom to be a concern.

CHAPTER 7 BIBLIOGRAPHY

Allen, R.G.D. (1938). Mathematical Analysis for Economists. Macmillan, London.

Austin, A. (2005). Energy and Power in China: Domestic Regulation and Foreign Policy. The Foreign Policy Centre, UK.

Berndt, E.R. and Wood, D.O. (1975). Technology, prices, and the derived demand for energy. Review of Econonomics and Statistics 57(3).

Blackorby Charles and Russell R. Robert. (1989). Will the Real Elasticity of Substitution Please Stand Up? (A Comparison of Allen/Uzawa and Morishima Elasticities). American Economic Review 79(4)

Bong Chin Kim and Walter C. Labys. (1988). Application of the translog model of energy substitution to developing countries: The case of Korea. Energy Economics, Oct.

Berndt, E.R., and Wood, D.O. (1975). Technology, prices, and the derived demand for energy. Review of Economics and Statistics 57(3).

Brown, R. S., and Christensen, L. R. 1981. Estimating Elasticities of Substitution in a Model of Partial Static Equilibrium, in Modelling and Measuring Natural Resource Substitution, ed. by E. R. Berndt, and B. C. Field. Cambridge, MA: MIT Press.

Martinello, F. (1985). Factor substitution, technical change, and returns to scale in Canadian forest industries. Canadian Journal of Forest Research

Christensen, L.R., and Greene, W.H. (1976). Economies of scale in U.S. electric power generation. Journal of Political Economy 84(4).

Christensen, L.R., Jorgenson, D.W, and Lau, L.J. (1971). Conjugate Duality and the Transcendental Logarithmic Production Function. Econometrica, 39.

Christensen, L.R., Jorgenson, D.W., and Lau, L.J. (1973). Transcendental logarithmic production frontier. Review of Econonomics and Statistics 55(Feb).

Considine T. J. (1989). Separability, Functional form, and Regulatory Policy in Models of Interfuel Substitution. Energy Economics(April).

Diewert, W. E. (1974). Functional Forms for Revenue and Factor Requirements Functions. International Economic Review 15.

Diewert, W. Erwin and Terence J. Wales. (1987). Flexible Functional Forms and Global Curvature Conditions. Econometrica 55.

Diewert, W. Erwin. (1986). A Brief Survey of Duality Theory. Discussion Paper No. 86-13

Diewert, W. Erwin and Terence J. Wales. (1988). Normalized Quadratic Systems of Consumer Demand Functions. Journal of Business and Economic Statistics 6.

Dwivedi, T and K. Srivastava. (1978). Optimality of Least Squares in the SUR Model. Journal of Econometrics 7.

Ernst R. Berndt (1973). The Translog Function and The Substitution of Equipment, Structures, and Labor in US Manuracturing. Journal of Econometrics 1 81-l 14.

Eric D. Larson, Wu Zongxin, Pat DeLaquil, Chen Weiying, Gao Pengfei. (2003). Future Implications of China's Energy Technology Choices. Energy Policy

Fuss, M.A. (1977). The Demand for Energy in Canadian Manufacturing: an Example of the Estimation of Production Structures with Many Inputs. Journal of Econometrics 5.

Gan Lin (1998). Energy Development and Environmental Constraints in China. Energy Policy. Vol. 26. No. 2

Garbaccio R. F., Ho M. S. and Jorgenson D. W. (1999) Why has the Energy-Output Ratio Fallen in China? The Energy Journal Vol 20, No. 3

Gordon, S (1996). Using Mixture of Flexible Functional Forms to Estimate Factor Demand Elasticities. The Canadian Journal of Economics 29(3)

Greene, William H. (2003). Econometric Analysis. Prentice Hall.

Hicks J.R. (1932). Theory of Wages.MacMillan, London.

Jae Wan Chung. (1986). On the Estimation of Factor Substitution in the Translog Model. The Review of Economics and Statistics.

James H. Grant. (1993). The Translog Approximate Function: Substitution among Inputs in Manufacturing Evaluated at Sample Means. Economics Letters (41).

Jorgensen, D. W., and B. Fraumeni. (1981). "Relative Prices and Technical Change." In E. R. Berndt, and B. C. Field (eds.), Modelling and Measuring Natural Resource Substitution. Cambridge, Mass.: MIT Press.

Lau, L.J. (1978). Testing and Imposing Monotonicity, Convexity, and Quasi-Convexity Constraints. In Production Economics: A Dual Approach to Theory and Applications (Vol. 1) M. Fuss and D. McFadden (eds.), Amsterdam: North

Holland, (1978), pp. 409-453.

Lerner Abba. (1933). The Diagrammatical Representation of Elasticity of Substitution. Review of Economic Studies.

Morishima, M. (1967). A Few Suggestions on the Theory of Elasticity. In Japanese. Keizai Hyoron. Economic Review(16).

Mundlak, Y. (1996). Production Function Estimation: Reviving the Primal. Econometrica 64 (2)

Giovanni Urga, Chris Walters. (2003). Dynamic Translog and Linear Logit Models: a Factor Demand Analysis of Interfuel Substitution in US Industrial Energy Demand. Energy Economics 25 (1-21).

Hicks. (1932). On the relationship between the Hicks and Robinson Definitions.

Hing Lin Chan, Shu Kam Lee. (1997). Modelling and Forecsting the Demand for Coal in China. Energy Economics (19).

J. Sinton, M. Levine, D. Fridley, F. Yang, and J. Lin. (1999). Status Report on Energy Efficiency Policy and Prorams in China. Lawrence Berkeley National Laboratory.

Lin Gan. (1998). Energy development and environmental constraints in China. Energy Policy 26(2).

Nautiyal, J.C., and Singh, B.K. (1985). Production Structure and Derived Demand for Factors Inputs in the Canadian lumber industry. Forest Science (31).

Paul Crompton, Yanrui Wu. (2004) Energy Consumption in China: Past Trends and

Future Directions. Energy Economics

Ryan, David L. and Terence J. Wales. (1998). A Simple Method for Imposing Local Curvature in Some Flexible Consumer-Demand Systems. Journal of Business and Economic Statistics (16).

Ryan L. David and Terence J. Wales. (2000) Imposing local concavity in the translog and generalized Leontief cost functions. Economics Letters 67(3).

Shashi Kant and J.C. Nautiyal. (1997). Production Structure, Factor Substitution, Technical Change, and Total Factor Productivity in the Canadian Logging Industry. Can. J. For. Res. (27).

Uzawa Hirofumi. (1962). Production Functions with Constant Elasticities of Substitution. Review of Economic Studies Oct.(29).

Varian, H.R. (1978). Microeconomic Analysis. New York: Norton and Company.

Solveig Glomsrod, Wei Taoyuan. (2003). Coal Cleaning: A Viable Strategy for Reduced Carbon Emissions and Improved Environment in China? Energy Policy.

Smil, Vaclav. (1990). A Report Prepared for the Office of Technology Assessment, U.S. Congress. China's Energy.

Smil, Vaclav. (1988). Energy in China's Modernization. New York: M. E. Sharpe

Tomitate, Takao. (1989). Economic Development and Energy Problems in China. Energy in Japan (95).

Richard F. Garbaccio, Mun S. Ho and Dale W. Jrgenson. (1999). Why Has the

Energy-output Ratio Fallen in China. The Energy Journal 20(3).

Walter Enders. (2004) Applied Econometric Time Series. Second Edition

Wang, Bangcheng. (1990). Energy Conservation in China. Energy Exploration & Exploitation 8(4)

Wei Zhihong. (2004). Economic Development and Energy Issues in China. Working paper prepared for 1st KEIO-UNU-JFIR Panel Meeting

Wu. Yanrui. (2002). Deregulation and Growth in China's Energy Sector: A Review of Recent Development. Energy Policy.

Zellner, A. (1962). An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias. Journal of American Statistical Association (67).

Zhang, Z.X. (2000). Can China afford to commit itself an emissions cap? An economic and political analysis. Energy Econmics (22).

Zhang Z. X. (1995). Energy Conservation in China: An International Perspective. Energy Policy 23(2).

Zhang Z. X. (2003). Why Did the Energy Intensity Fall in China's Industrial Sector in the 1990s? The Relative Importance of Structural Change and Intensity Change. Energy Economics (25).

APPENDIX: COMPLETE EMPIRICAL RESULTS

* Index 1, 2, 3, 4 represent Coal, Oil, Labor, Capital respectively

A1 Checking Regularity Condition "Monotonicity": Fitted Share Equations

	S1	S2
1	0.95152	0.048481
2	0.94378	0.056224
3	0.96893	0.031070
4	0.91548	0.084516
5	0.92685	0.073148
6	0.94170	0.058302
7	0.93167	0.068332
8	0.91201	0.087990
9	0.86342	0.13658
10	0.88850	0.11150
11	0.88431	0.11569
12	0.90825	0.091747
13	0.92096	0.079039
14	0.91969	0.080309
15	0.89899	0.10101
16	0.90324	0.096756
17	0.87838	0.12162
18	0.87567	0.12433
19	0.87026	0.12974
20	0.87892	0.12108
21	0.87962	0.12038
22	0.88586	0.11414
23	0.88952	0.11048
24	0.88658	0.11342
25	0.88300	0.11700
26	0.89548	0.10452
27	0.88379	0.11621
28	0.89605	0.10395
29	0.90869	0.091311
30	0.90584	0.094160
31	0.89577	0.10423
32	0.88742	0.11258
33	0.90173	0.098274
34	0.86308	0.13692
35	0.83226	0.16774
36	0.81333	0.18667
37	0.82907	0.17093
38	0.83134	0.16866

39	0.83377	0.16623
40	0.80117	0.19883
41	0.73033	0.26967
42	0.72833	0.27167
43	0.74533	0.25467
44	0.71667	0.28333

A2 Checking Curvature Condition: Maximum Eigen-value of Hessian Matrix of Cost Function at Every Observation Point

Year	Maximum Eigen-Value	Year	Maximum Eigen-Value
1960	0.26223	1982	0.15794
1961	0.24836	1983	0.15337
1962	0.29428	1984	0.14787
1963	0.19974	1985	0.1673
1964	0.21889	1986	0.14907
1965	0.24468	1987	0.1682
1966	0.22716	1988	0.18854
1967	0.19399	1989	0.1839
1968	0.11863	1990	0.16775
1969	0.15635	1991	0.15467
1970	0.14988	1992	0.17725
1971	0.18783	1993	0.11814
1972	0.2089	1994	0.075279
1973	0.20677	1995	0.050832
1974	0.17287	1996	0.071054
1975	0.1797	1997	0.074055
1976	0.14084	1998	0.077286
1977	0.13675	1999	0.035893
1978	0.12867	2000	0
1979	0.14165	2001	0
1980	0.1427	2002	0
1981	0.15226	2003	0

A3 Allen Partial Elasticities of Substitution (AES) between Coal and Oil

1960	-2.84223	1982	-0.80358
1961	-2.34023	1983	-0.76259
1962	-4.88749	1984	-0.71567
1963	-1.29074	1985	-0.89374
1964	-1.6143	1986	-0.72569
1965	-2.22829	1987	-0.90294
1966	-1.78407	1988	-1.13614
1967	-1.2087	1989	-1.07801
1968	-0.50297	1990	-0.89837
1969	-0.78913	1991	-0.77408
1970	-0.7325	1992	-1.00012
1971	-1.12701	1993	-0.49984
1972	-1.43492	1994	-0.26962
1973	-1.39972	1995	-0.1674
1974	-0.95185	1996	-0.25069
1975	-1.02807	1997	-0.26408
1976	-0.65919	1998	-0.27881
1977	-0.62803	1999	-0.11266
1978	-0.56977	2000	0.10005
1979	-0.6655	2001	0.10423
1980	-0.67382	2002	0.066236
1981	-0.75295	2003	0.12712

A4 Price Elasticities: Own Price Elasticities and Cross Price Elasticities

Year	E11	E22	E21	E12
1960	0.13779	2.70444	-2.70444	-0.13779
1961	0.13158	2.20865	-2.20865	-0.13158
1962	0.15186	4.73563	-4.73563	-0.15186
1963	0.10909	1.18165	-1.18165	-0.10909
1964	0.11808	1.49621	-1.49621	-0.11808
1965	0.12991	2.09837	-2.09837	-0.12991
1966	0.12191	1.66216	-1.66216	-0.12191
1967	0.10635	1.10234	-1.10234	-0.10635
1968	0.068697	0.43427	-0.43427	-0.068697
1969	0.087987	0.70115	-0.70115	-0.087987
1970	0.084742	0.64776	-0.64776	-0.084742
1971	0.1034	1.02361	-1.02361	-0.1034
1972	0.11341	1.3215	-1.3215	-0.11341
1973	0.11241	1.28731	-1.28731	-0.11241
1974	0.096147	0.8557	-0.8557	-0.096147
1975	0.099473	0.9286	-0.9286	-0.099473
1976	0.080167	0.57902	-0.57902	-0.080167
1977	0.078081	0.54995	-0.54995	-0.078081
1978	0.073924	0.49585	-0.49585	-0.073924
1979	0.080579	0.58492	-0.58492	-0.080579
1980	0.081117	0.5927	-0.5927	-0.081117
1981	0.085941	0.66701	-0.66701	-0.085941
1982	0.088778	0.71481	-0.71481	-0.088778
1983	0.086495	0.6761	-0.6761	-0.086495
1984	0.083731	0.63194	-0.63194	-0.083731
1985	0.093412	0.80033	-0.80033	-0.093412
1986	0.084335	0.64135	-0.64135	-0.084335
1987	0.093857	0.80909	-0.80909	-0.093857
1988	0.10374	1.0324	-1.0324	-0.10374
1989	0.10151	0.97651	-0.97651	-0.10151
1990	0.093637	0.80473	-0.80473	-0.093637
1991	0.087147	0.68693	-0.68693	-0.087147
1992	0.098285	0.90183	-0.90183	-0.098285
1993	0.068439	0.4314	-0.4314	-0.068439
1994	0.045226	0.22439	-0.22439	-0.045226
1995	0.031249	0.13615	-0.13615	-0.031249
1996	0.042852	0.20784	-0.20784	-0.042852
1997	0.04454	0.21954	-0.21954	-0.04454

1998	0.046348	0.23246	-0.23246	-0.046348
1999	0.0224	0.09026	-0.09026	-0.0224
2000	-0.02698	-0.073069	0.073069	0.02698
2001	-0.028315	-0.075911	0.075911	0.028315
2002	-0.016869	-0.049367	0.049367	0.016869
2003	-0.036018	-0.091104	0.091104	0.036018

A5 Elasticities of Intensity

EI13	EI14	EI23	EI24
1.09221	1.18026	-0.80973	-2.53801
1.0216	1.24291	0.63749	-3.07749
1.01443	1.1531	0.55005	-3.77456
1.01837	1.36197	0.80103	-2.92091
1.02109	1.57005	0.73278	-6.22301
1.02476	2.06036	0.60014	-16.12702
1.02882	4.39726	0.60705	-45.31932
1.02385	3.90888	0.75277	-29.15042
1.02491	0.45389	0.84254	4.45223
1.02643	-0.48658	0.78937	12.8462
1.0362	0.29084	0.72331	6.42081
1.04055	0.10195	0.59854	9.89024
1.03976	-0.84852	0.5367	22.53883
1.04684	-0.32994	0.46364	16.23031
1.04882	0.25132	0.56554	7.66316
1.05388	0.3269	0.49699	7.28356
1.04644	0.55049	0.66461	4.24666
1.04741	0.61292	0.66609	3.72632
1.05713	0.69358	0.61677	3.05535
1.05376	0.6884	0.60976	3.26192
1.05187	0.69837	0.62098	3.20399
1.04599	0.68134	0.64309	3.47318
1.048	0.67619	0.6135	3.60716
1.05215	0.65832	0.59235	3.67075
1.06258	0.62348	0.5277	3.8417
1.0861	0.68421	0.2623	3.70559
1.09293	0.67502	0.29331	3.47141
1.10061	0.5765	0.13268	4.6507
1.12474	0.51701	-0.2414	5.80649
1.11095	0.5706	-0.067399	5.13086
1.09624	0.61275	0.17293	4.32812
1.0907	0.64402	0.28507	3.80602
1.09147	0.61118	0.16071	4.56767
1.1218	0.6975	0.23225	2.90679
1.13508	0.72031	0.32979	2.38774
1.14022	0.75818	0.38909	2.05361
1.11034	0.77496	0.46485	2.09148
1.08706	0.78701	0.5709	2.04983

1.07167	0.80821	0.64053	1.96196
1.06362	0.82626	0.74363	1.70009
1.0592	0.83258	0.83966	1.45341
1.05473	0.85606	0.85328	1.38589
1.0491	0.87377	0.85631	1.36941
1.0499	0.88657	0.87377	1.28692

A6 Complete Results for Unit Root Tests

Coal:

```
************************************************************
* TESTING THE NULL HYPOTHESIS OF A UNIT ROOT IN D          *
*          Using data from 1960:01 to 2003:01              *
*    Choosing the optimal lag length for the ADF regression *
*     between 0 and  11  lags.                             *
*    using the BIC model selection criterion.              *
************************************************************

Model Selection Criteria
   Minimum AIC at lag:    1
   Minimum BIC at lag:    1

************************************************************
* Augmented Dickey-Fuller t-test with 1 lags:     2.7153   *
*              1%        5%        10%                     *
*            -2.62     -1.95     -1.61                     *
*                                                          *
* Augmented Dickey-Fuller Z-test with 1 lags:     0.3732   *
*              1%        5%        10%                     *
*           -12.9     -7.7      -5.5                      *
************************************************************



************************************************************
* TESTING THE NULL HYPOTHESIS OF A UNIT ROOT IN D          *
*          Using data from 1960:01 to 2003:01              *
*    Choosing the optimal lag length for the ADF regression *
*     between 0 and  11  lags.                             *
*    using the BIC model selection criterion.              *
************************************************************

Model Selection Criteria
   Minimum AIC at lag:    1
   Minimum BIC at lag:    1

************************************************************
* Augmented Dickey-Fuller t-test with 1 lags:    -1.2773   *
*              1%        5%        10%                     *
*            -3.58     -2.93     -2.60                     *
*                                                          *
```

```
* Augmented Dickey-Fuller Z-test with 1 lags:        -1.6225      *
*                    1%          5%          10%                      *
*                -18.9       -13.3       -10.7                        *
*                                                                     *
* Coefficient and T-Statistic on the Constant:                       *
*                    0.13662          1.5931                          *
*                                                                     *
* Joint test of a unit root and no constant:          5.0973      *
*                    1%          5%          10%                      *
*                   7.06        4.86        3.94                      *
***********************************************************************


***********************************************************************
* TESTING THE NULL HYPOTHESIS OF A UNIT ROOT IN D                     *
*              Using data from 1960:01 to 2003:01                     *
*      Choosing the optimal lag length for the ADF regression         *
*       between 0 and   11   lags.                                    *
*      using the BIC model selection criterion.                       *
***********************************************************************

Model Selection Criteria
    Minimum AIC at lag:      1
    Minimum BIC at lag:      1

***********************************************************************
* Augmented Dickey-Fuller t-test with 1 lags:        -3.1586      *
*                    1%          5%          10%                      *
*                -4.15       -3.50       -3.18                        *
*                                                                     *
* Augmented Dickey-Fuller Z-test with 1 lags:       -16.5222      *
*                    1%          5%          10%                      *
*                -25.7       -19.8       -16.8                        *
*                                                                     *
* Coefficient and T-Statistic on the Constant:                       *
*                    0.92570          3.2912                          *
* Coefficient and T-Statistic on the Linear Trend:                   *
*                    0.00920          2.9216                          *
*                                                                     *
* Joint test of a unit root and no linear trend       5.2411      *
*                    1%          5%          10%                      *
*                   9.31        6.73        5.61                      *
***********************************************************************
```

```
*************************************************************
* TESTING THE NULL HYPOTHESIS OF A UNIT ROOT IN R           *
*          Using data from 1960:01 to 2003:01               *
*    Choosing the optimal lag length for the ADF regression *
*     between 0 and  11  lags.                              *
*    using the BIC model selection criterion.               *
*************************************************************

Model Selection Criteria
   Minimum AIC at lag:    1
   Minimum BIC at lag:    1

*************************************************************
* Augmented Dickey-Fuller t-test with 1 lags:      1.2626   *
*              1%        5%        10%                      *
*            -2.62     -1.95     -1.61                      *
*                                                           *
* Augmented Dickey-Fuller Z-test with 1 lags:      1.8957   *
*              1%        5%        10%                      *
*            -12.9     -7.7      -5.5                       *
*************************************************************



*************************************************************
* TESTING THE NULL HYPOTHESIS OF A UNIT ROOT IN R           *
*          Using data from 1960:01 to 2003:01               *
*    Choosing the optimal lag length for the ADF regression *
*     between 0 and  11  lags.                              *
*    using the BIC model selection criterion.               *
*************************************************************

Model Selection Criteria
   Minimum AIC at lag:    8
   Minimum BIC at lag:    1

*************************************************************
* Augmented Dickey-Fuller t-test with 1 lags:      0.5347   *
*              1%        5%        10%                      *
*            -3.58     -2.93     -2.60                      *
*                                                           *
```

```
* Augmented Dickey-Fuller Z-test with 1 lags:          0.7844      *
*                    1%          5%          10%                        *
*                -18.9       -13.3       -10.7                      *
*                                                                      *
* Coefficient and T-Statistic on the Constant:                  *
*                    0.01854          1.3678                        *
*                                                                      *
* Joint test of a unit root and no constant:          1.7500      *
*                    1%          5%          10%                        *
*                   7.06        4.86        3.94                      *
************************************************************************


************************************************************************
* TESTING THE NULL HYPOTHESIS OF A UNIT ROOT IN R                      *
*              Using data from 1960:01 to 2003:01                  *
*      Choosing the optimal lag length for the ADF regression      *
*       between 0 and   11   lags.                                 *
*      using the BIC model selection criterion.                    *
************************************************************************

Model Selection Criteria
    Minimum AIC at lag:      8
    Minimum BIC at lag:      1

************************************************************************
* Augmented Dickey-Fuller t-test with 1 lags:         -1.4741      *
*                    1%          5%          10%                        *
*                -4.15       -3.50       -3.18                      *
*                                                                      *
* Augmented Dickey-Fuller Z-test with 1 lags:         -4.2143      *
*                    1%          5%          10%                        *
*                -25.7       -19.8       -16.8                      *
*                                                                      *
* Coefficient and T-Statistic on the Constant:                  *
*                   -0.03540         -1.1807                        *
* Coefficient and T-Statistic on the Linear Trend:              *
*                    0.00354          1.9988                        *
*                                                                      *
* Joint test of a unit root and no linear trend       2.1515      *
*                    1%          5%          10%                        *
*                   9.31        6.73        5.61                      *
************************************************************************
```

```
*************************************************************
* TESTING THE NULL HYPOTHESIS OF A UNIT ROOT IN G           *
*            Using data from 1960:01 to 2003:01             *
*     Choosing the optimal lag length for the ADF regression *
*      between 0 and  11  lags.                              *
*     using the BIC model selection criterion.               *
*************************************************************

Model Selection Criteria
   Minimum AIC at lag:    1
   Minimum BIC at lag:    0

*************************************************************
* Augmented Dickey-Fuller t-test with 0 lags:      6.0967   *
*                1%        5%        10%                    *
*              -2.62     -1.95     -1.61                     *
*                                                           *
* Augmented Dickey-Fuller Z-test with 0 lags:      0.4508   *
*                1%        5%        10%                    *
*             -12.9      -7.7      -5.5                      *
*************************************************************



*************************************************************
* TESTING THE NULL HYPOTHESIS OF A UNIT ROOT IN G           *
*            Using data from 1960:01 to 2003:01             *
*     Choosing the optimal lag length for the ADF regression *
*      between 0 and  11  lags.                              *
*     using the BIC model selection criterion.               *
*************************************************************

Model Selection Criteria
   Minimum AIC at lag:    1
   Minimum BIC at lag:    0

*************************************************************
* Augmented Dickey-Fuller t-test with 0 lags:      1.5172   *
*                1%        5%        10%                    *
*              -3.58     -2.93     -2.60                     *
*                                                           *
```

```
* Augmented Dickey-Fuller Z-test with 0 lags:          0.7790      *
*                    1%          5%          10%                      *
*                -18.9       -13.3       -10.7                        *
*                                                                     *
* Coefficient and T-Statistic on the Constant:                       *
*                  -0.05353          -0.6460                          *
*                                                                     *
* Joint test of a unit root and no constant:          18.5358      *
*                    1%          5%          10%                      *
*                   7.06        4.86        3.94                      *
***********************************************************************


***********************************************************************
* TESTING THE NULL HYPOTHESIS OF A UNIT ROOT IN G                     *
*              Using data from 1960:01 to 2003:01                     *
*      Choosing the optimal lag length for the ADF regression         *
*       between 0 and   11   lags.                                    *
*      using the BIC model selection criterion.                       *
***********************************************************************

Model Selection Criteria
    Minimum AIC at lag:      1
    Minimum BIC at lag:      0

***********************************************************************
* Augmented Dickey-Fuller t-test with 0 lags:          -5.5697      *
*                    1%          5%          10%                      *
*                -4.15       -3.50       -3.18                        *
*                                                                     *
* Augmented Dickey-Fuller Z-test with 0 lags:         -17.5599      *
*                    1%          5%          10%                      *
*                -25.7       -19.8       -16.8                        *
*                                                                     *
* Coefficient and T-Statistic on the Constant:                       *
*                  2.07360           5.6320                          *
* Coefficient and T-Statistic on the Linear Trend:                   *
*                  0.03485           5.8598                          *
*                                                                     *
* Joint test of a unit root and no linear trend       19.2556      *
*                    1%          5%          10%                      *
*                  9.31        6.73        5.61                       *
***********************************************************************
```

---

break date TB = 1991:01 statistic t(alpha=1) -5.31329

critical values at	1%	5%	10%	50%	90%	95%	99%
for 70 obs.	-6.32	-5.59	-5.29	-4.24	-3.32	-3.08	-2.67

---

number of lag retained : 6

---

explained variable : D

---

	coefficient	student
CONSTANT	3.99504	5.42853
DU	1.12032	3.83842
D(Tb)	-0.05067	-0.99263
TIME	0.05341	5.12046
DT	-0.03350	-4.01837
D {1}	0.13423	0.82381

---

break date TB = 1984:01 statistic t(alpha=1) -6.12039

critical values at	1%	5%	10%	50%	90%	95%	99%
for 70 obs.	-6.32	-5.59	-5.29	-4.24	-3.32	-3.08	-2.67

---

number of lag retained : 4

---

explained variable : R

---

	coefficient	student
CONSTANT	-0.10057	-3.10448
DU	-1.70474	-5.63212
D(Tb)	0.22027	4.11725
TIME	0.00820	3.90977
DT	0.06126	5.77013
R {1}	0.33038	3.01968

---

break date TB = 1976:01 statistic t(alpha=1) -5.67412

critical values at	1%	5%	10%	50%	90%	95%	99%
for 70 obs.	-6.32	-5.59	-5.29	-4.24	-3.32	-3.08	-2.67

number of lag retained : 1

explained variable : G

	coefficient	student
CONSTANT	3.36302	5.84121
DU	-0.36500	-5.17623
D(Tb)	0.04691	1.21830
TIME	0.04066	4.85836
DT	0.01852	5.45671
G {1}	0.36092	3.20448

Oil

```
*************************************************************
* TESTING THE NULL HYPOTHESIS OF A UNIT ROOT IN D           *
*          Using data from 1960:01 to 2003:01               *
*    Choosing the optimal lag length for the ADF regression *
*     between 0 and  11  lags.                              *
*    using the BIC model selection criterion.               *
*************************************************************

Model Selection Criteria
   Minimum AIC at lag:    2
   Minimum BIC at lag:    1

*************************************************************
* Augmented Dickey-Fuller t-test with 1 lags:      1.8353   *
*              1%          5%          10%                  *
*            -2.62       -1.95       -1.61                  *
*                                                           *
* Augmented Dickey-Fuller Z-test with 1 lags:      0.6674   *
*              1%          5%          10%                  *
*            -12.9       -7.7        -5.5                   *
*************************************************************



*************************************************************
* TESTING THE NULL HYPOTHESIS OF A UNIT ROOT IN D           *
*          Using data from 1960:01 to 2003:01               *
*    Choosing the optimal lag length for the ADF regression *
*     between 0 and  11  lags.                              *
*    using the BIC model selection criterion.               *
*************************************************************

Model Selection Criteria
   Minimum AIC at lag:    5
   Minimum BIC at lag:    1

*************************************************************
* Augmented Dickey-Fuller t-test with 1 lags:     -2.3744   *
*              1%          5%          10%                  *
```

```
*                -3.58       -2.93       -2.60                          *
*                                                                       *
* Augmented Dickey-Fuller Z-test with 1 lags:       -2.2474      *
*                   1%          5%          10%                         *
*                -18.9       -13.3       -10.7                          *
*                                                                       *
* Coefficient and T-Statistic on the Constant:                  *
*                   0.17052          3.0166                           *
*                                                                       *
* Joint test of a unit root and no constant:          6.5752      *
*                   1%          5%          10%                         *
*                  7.06        4.86        3.94                         *
*************************************************************************


*************************************************************************
* TESTING THE NULL HYPOTHESIS OF A UNIT ROOT IN D                       *
*            Using data from 1960:01 to 2003:01                 *
*      Choosing the optimal lag length for the ADF regression     *
*       between 0 and   11   lags.                                      *
*      using the BIC model selection criterion.                  *
*************************************************************************

Model Selection Criteria
    Minimum AIC at lag:     5
    Minimum BIC at lag:     1

*************************************************************************
* Augmented Dickey-Fuller t-test with 1 lags:       -2.2228      *
*                   1%          5%          10%                         *
*                -4.15       -3.50       -3.18                          *
*                                                                       *
* Augmented Dickey-Fuller Z-test with 1 lags:       -6.6748      *
*                   1%          5%          10%                         *
*                -25.7       -19.8       -16.8                          *
*                                                                       *
* Coefficient and T-Statistic on the Constant:                  *
*                   0.28288          3.0305                           *
* Coefficient and T-Statistic on the Linear Trend:              *
*                   0.00436          1.4992                           *
*                                                                       *
* Joint test of a unit root and no linear trend       4.0329      *
*                   1%          5%          10%                         *
```

```
*                  9.31         6.73         5.61                          *
************************************************************


************************************************************
* TESTING THE NULL HYPOTHESIS OF A UNIT ROOT IN R                  *
*              Using data from 1960:01 to 2003:01                  *
*      Choosing the optimal lag length for the ADF regression      *
*        between 0 and   11   lags.                                *
*      using the BIC model selection criterion.                    *
************************************************************

Model Selection Criteria
    Minimum AIC at lag:      1
    Minimum BIC at lag:      1

************************************************************
* Augmented Dickey-Fuller t-test with 1 lags:          1.4571      *
*                  1%          5%          10%                     *
*                -2.62       -1.95       -1.61                     *
*                                                                   *
* Augmented Dickey-Fuller Z-test with 1 lags:          2.1934      *
*                  1%          5%          10%                     *
*                -12.9       -7.7        -5.5                      *
************************************************************


************************************************************
* TESTING THE NULL HYPOTHESIS OF A UNIT ROOT IN R                  *
*              Using data from 1960:01 to 2003:01                  *
*      Choosing the optimal lag length for the ADF regression      *
*        between 0 and   11   lags.                                *
*      using the BIC model selection criterion.                    *
************************************************************

Model Selection Criteria
    Minimum AIC at lag:     11
    Minimum BIC at lag:      1

************************************************************
* Augmented Dickey-Fuller t-test with 1 lags:          0.6158      *
*                  1%          5%          10%                     *
```

```
*                  -3.58       -2.93       -2.60                          *
*                                                                         *
* Augmented Dickey-Fuller Z-test with 1 lags:          1.1694      *
*                     1%          5%          10%                         *
*                  -18.9       -13.3       -10.7                          *
*                                                                         *
* Coefficient and T-Statistic on the Constant:                  *
*                     0.02046          0.8991                           *
*                                                                         *
* Joint test of a unit root and no constant:           1.4607      *
*                     1%          5%          10%                         *
*                    7.06        4.86        3.94                         *
***************************************************************

***************************************************************
* TESTING THE NULL HYPOTHESIS OF A UNIT ROOT IN R                    *
*              Using data from 1960:01 to 2003:01                 *
*       Choosing the optimal lag length for the ADF regression    *
*        between 0 and   11   lags.                                *
*       using the BIC model selection criterion.                  *
***************************************************************

Model Selection Criteria
    Minimum AIC at lag:    11
    Minimum BIC at lag:     1

***************************************************************
* Augmented Dickey-Fuller t-test with 1 lags:         -0.9386      *
*                     1%          5%          10%                         *
*                  -4.15       -3.50       -3.18                          *
*                                                                         *
* Augmented Dickey-Fuller Z-test with 1 lags:         -2.0499      *
*                     1%          5%          10%                         *
*                  -25.7       -19.8       -16.8                          *
*                                                                         *
* Coefficient and T-Statistic on the Constant:                  *
*                    -0.04687          -1.1676                          *
* Coefficient and T-Statistic on the Linear Trend:              *
*                     0.00437          2.0021                           *
*                                                                         *
* Joint test of a unit root and no linear trend        2.2085      *
*                     1%          5%          10%                         *
```

```
*                  9.31         6.73         5.61                          *
************************************************************

************************************************************
* TESTING THE NULL HYPOTHESIS OF A UNIT ROOT IN G                  *
*              Using data from 1960:01 to 2003:01                  *
*      Choosing the optimal lag length for the ADF regression      *
*       between 0 and   11   lags.                                 *
*      using the BIC model selection criterion.                    *
************************************************************

Model Selection Criteria
    Minimum AIC at lag:      1
    Minimum BIC at lag:      0

************************************************************
* Augmented Dickey-Fuller t-test with 0 lags:          6.0967      *
*                   1%          5%          10%                    *
*                -2.62       -1.95       -1.61                     *
*                                                                  *
* Augmented Dickey-Fuller Z-test with 0 lags:          0.4508      *
*                   1%          5%          10%                    *
*                -12.9       -7.7        -5.5                      *
************************************************************

************************************************************
* TESTING THE NULL HYPOTHESIS OF A UNIT ROOT IN G                  *
*              Using data from 1960:01 to 2003:01                  *
*      Choosing the optimal lag length for the ADF regression      *
*       between 0 and   11   lags.                                 *
*      using the BIC model selection criterion.                    *
************************************************************

Model Selection Criteria
    Minimum AIC at lag:      1
    Minimum BIC at lag:      0

************************************************************
* Augmented Dickey-Fuller t-test with 0 lags:          1.5172      *
*                   1%          5%          10%                    *
```

```
*                  -3.58       -2.93       -2.60                          *
*                                                                         *
* Augmented Dickey-Fuller Z-test with 0 lags:          0.7790      *
*                    1%          5%          10%                          *
*                 -18.9       -13.3       -10.7                          *
*                                                                         *
* Coefficient and T-Statistic on the Constant:                    *
*                  -0.05353          -0.6460                          *
*                                                                         *
* Joint test of a unit root and no constant:          18.5358      *
*                    1%          5%          10%                          *
*                  7.06        4.86        3.94                          *
***************************************************************


***************************************************************
* TESTING THE NULL HYPOTHESIS OF A UNIT ROOT IN G                  *
*              Using data from 1960:01 to 2003:01                 *
*      Choosing the optimal lag length for the ADF regression     *
*       between 0 and   11   lags.                                *
*      using the BIC model selection criterion.                   *
***************************************************************

Model Selection Criteria
    Minimum AIC at lag:     1
    Minimum BIC at lag:     0

***************************************************************
* Augmented Dickey-Fuller t-test with 0 lags:         -5.5697      *
*                    1%          5%          10%                          *
*                 -4.15       -3.50       -3.18                          *
*                                                                         *
* Augmented Dickey-Fuller Z-test with 0 lags:        -17.5599      *
*                    1%          5%          10%                          *
*                 -25.7       -19.8       -16.8                          *
*                                                                         *
* Coefficient and T-Statistic on the Constant:                    *
*                  2.07360           5.6320                          *
* Coefficient and T-Statistic on the Linear Trend:                *
*                  0.03485           5.8598                          *
*                                                                         *
* Joint test of a unit root and no linear trend       19.2556      *
*                    1%          5%          10%                          *
```

```
*                    9.31         6.73         5.61                          *
************************************************************************

------------------------------------------------------------------------
break date TB = 1967:01   statistic t(alpha==1)        -6.68053
critical values at 1%        5%          10%        50%        90%        95%
99%
  for 70 obs.     -6.32     -5.59     -5.29     -4.24     -3.32     -3.08     -2.67
------------------------------------------------------------------------
     number of lag retained : 4
-----------------------------------------------------
explained variable :     D
-----------------------------------------------------
                              coefficient              student
-----------------------------------------------------
CONSTANT                        2.12731                7.41391
DU                             -1.19676               -4.31854
D(Tb)                          -0.18235               -2.33940
TIME                           -0.19015               -4.89272
DT                              0.20331                5.22459
D {1}                           0.74611               19.63266
-----------------------------------------------------
------------------------------------------------------------------------
break date TB = 1985:01   statistic t(alpha==1)        -3.91042
critical values at 1%        5%          10%        50%        90%        95%
99%
  for 70 obs.     -6.32     -5.59     -5.29     -4.24     -3.32     -3.08     -2.67
------------------------------------------------------------------------
     number of lag retained : 6
-----------------------------------------------------
explained variable :     R
-----------------------------------------------------
                              coefficient              student
-----------------------------------------------------
CONSTANT                        0.35226                2.90256
DU                             -3.28161               -3.47782
D(Tb)                           0.19095                1.50842
TIME                           -0.01158               -2.21197
DT                              0.11909                3.59166
R {1}                           0.25022                1.30499
-----------------------------------------------------
------------------------------------------------------------------------
break date TB = 1976:01   statistic t(alpha==1)        -5.67412
critical values at 1%        5%          10%        50%        90%        95%
```

99%

for 70	obs.	-6.32	-5.59	-5.29	-4.24	-3.32	-3.08	-2.67

number of lag retained : 1

explained variable : G

	coefficient	student
CONSTANT	3.36302	5.84121
DU	-0.36500	-5.17623
D(Tb)	0.04691	1.21830
TIME	0.04066	4.85836
DT	0.01852	5.45671
G {1}	0.36092	3.20448

Electricity:

```
************************************************************
* TESTING THE NULL HYPOTHESIS OF A UNIT ROOT IN D          *
*          Using data from 1980:01 to 2003:01              *
*    Choosing the optimal lag length for the ADF regression *
*     between 0 and  6  lags.                               *
*    using the BIC model selection criterion.              *
************************************************************
```

Model Selection Criteria
Minimum AIC at lag: 0
Minimum BIC at lag: 0

```
************************************************************
* Augmented Dickey-Fuller t-test with 0 lags:      5.6143  *
*                1%        5%        10%                   *
*              -2.66     -1.95     -1.60                    *
*                                                          *
* Augmented Dickey-Fuller Z-test with 0 lags:      0.4016  *
*                1%        5%        10%                   *
*             -11.9      -7.3      -5.3                     *
************************************************************
```

```
************************************************************
* TESTING THE NULL HYPOTHESIS OF A UNIT ROOT IN D          *
*          Using data from 1980:01 to 2003:01              *
*    Choosing the optimal lag length for the ADF regression *
*     between 0 and  6  lags.                               *
*    using the BIC model selection criterion.              *
************************************************************
```

Model Selection Criteria
Minimum AIC at lag: 1
Minimum BIC at lag: 1

```
************************************************************
* Augmented Dickey-Fuller t-test with 1 lags:        -0.3744      *
*                    1%          5%          10%                  *
*                 -3.75       -3.00       -2.63                   *
*                                                                 *
* Augmented Dickey-Fuller Z-test with 1 lags:        -0.2650      *
*                    1%          5%          10%                  *
*                -17.2       -12.5       -10.2                    *
*                                                                 *
* Coefficient and T-Statistic on the Constant:                    *
*                    0.11581          0.9024                      *
*                                                                 *
* Joint test of a unit root and no constant:          5.4226      *
*                    1%          5%          10%                  *
*                  7.88        5.18        4.12                   *
************************************************************


************************************************************
* TESTING THE NULL HYPOTHESIS OF A UNIT ROOT IN D                 *
*              Using data from 1980:01 to 2003:01                 *
*      Choosing the optimal lag length for the ADF regression     *
*        between 0 and   6   lags.                                *
*      using the BIC model selection criterion.                   *
************************************************************

Model Selection Criteria
    Minimum AIC at lag:     1
    Minimum BIC at lag:     1

************************************************************
* Augmented Dickey-Fuller t-test with 1 lags:        -3.4681      *
*                    1%          5%          10%                  *
*                 -4.38       -3.60       -3.24                   *
*                                                                 *
* Augmented Dickey-Fuller Z-test with 1 lags:       -27.5514      *
*                    1%          5%          10%                  *
*                -22.5       -17.9       -15.6                    *
*                                                                 *
* Coefficient and T-Statistic on the Constant:                    *
*                    2.47273          3.5665                     *
* Coefficient and T-Statistic on the Linear Trend:                *
*                    0.05099          3.4372                     *
```

```
*                                                                    *
* Joint test of a unit root and no linear trend       6.0170       *
*                    1%          5%          10%                    *
*                 10.61        7.24        5.91                     *
**********************************************************************


**********************************************************************
* TESTING THE NULL HYPOTHESIS OF A UNIT ROOT IN R                     *
*              Using data from 1980:01 to 2003:01                     *
*      Choosing the optimal lag length for the ADF regression         *
*       between 0 and   6   lags.                                     *
*      using the BIC model selection criterion.                       *
**********************************************************************

Model Selection Criteria
    Minimum AIC at lag:      3
    Minimum BIC at lag:      2

**********************************************************************
* Augmented Dickey-Fuller t-test with 2 lags:        -0.7738       *
*                    1%          5%          10%                    *
*                -2.66        -1.95        -1.60                    *
*                                                                    *
* Augmented Dickey-Fuller Z-test with 2 lags:        -1.4615       *
*                    1%          5%          10%                    *
*                -11.9        -7.3        -5.3                     *
**********************************************************************


**********************************************************************
* TESTING THE NULL HYPOTHESIS OF A UNIT ROOT IN R                     *
*              Using data from 1980:01 to 2003:01                     *
*      Choosing the optimal lag length for the ADF regression         *
*       between 0 and   6   lags.                                     *
*      using the BIC model selection criterion.                       *
**********************************************************************

Model Selection Criteria
    Minimum AIC at lag:      3
    Minimum BIC at lag:      0
```

```
*************************************************************
* Augmented Dickey-Fuller t-test with 0 lags:        -3.2420      *
*                    1%          5%          10%                  *
*                 -3.75       -3.00       -2.63                   *
*                                                                 *
* Augmented Dickey-Fuller Z-test with 0 lags:       -15.1011      *
*                    1%          5%          10%                  *
*                -17.2       -12.5       -10.2                    *
*                                                                 *
* Coefficient and T-Statistic on the Constant:                    *
*                 0.08622          2.0385                         *
*                                                                 *
* Joint test of a unit root and no constant:          5.2581      *
*                    1%          5%          10%                  *
*                  7.88        5.18        4.12                   *
*************************************************************


*************************************************************
* TESTING THE NULL HYPOTHESIS OF A UNIT ROOT IN R                 *
*              Using data from 1980:01 to 2003:01                 *
*      Choosing the optimal lag length for the ADF regression     *
*       between 0 and   6   lags.                                 *
*      using the BIC model selection criterion.                   *
*************************************************************

Model Selection Criteria
   Minimum AIC at lag:     3
   Minimum BIC at lag:     0

*************************************************************
* Augmented Dickey-Fuller t-test with 0 lags:        -3.1476      *
*                    1%          5%          10%                  *
*                 -4.38       -3.60       -3.24                   *
*                                                                 *
* Augmented Dickey-Fuller Z-test with 0 lags:       -15.5288      *
*                    1%          5%          10%                  *
*                -22.5       -17.9       -15.6                    *
*                                                                 *
* Coefficient and T-Statistic on the Constant:                    *
*                 0.06567          0.8661                         *
* Coefficient and T-Statistic on the Linear Trend:                *
*                 0.00176          0.3298                         *
```

```
*                                                                    *
* Joint test of a unit root and no linear trend        5.0868       *
*                    1%          5%          10%                     *
*                 10.61        7.24        5.91                     *
**********************************************************************


**********************************************************************
* TESTING THE NULL HYPOTHESIS OF A UNIT ROOT IN G                     *
*              Using data from 1980:01 to 2003:01                     *
*      Choosing the optimal lag length for the ADF regression         *
*       between 0 and   6   lags.                                     *
*      using the BIC model selection criterion.                       *
**********************************************************************

Model Selection Criteria
    Minimum AIC at lag:      4
    Minimum BIC at lag:      4

**********************************************************************
* Augmented Dickey-Fuller t-test with 4 lags:          2.6197        *
*                    1%          5%          10%                     *
*                -2.66        -1.95        -1.60                     *
*                                                                    *
* Augmented Dickey-Fuller Z-test with 4 lags:          0.2125        *
*                    1%          5%          10%                     *
*                -11.9        -7.3        -5.3                      *
**********************************************************************


**********************************************************************
* TESTING THE NULL HYPOTHESIS OF A UNIT ROOT IN G                     *
*              Using data from 1980:01 to 2003:01                     *
*      Choosing the optimal lag length for the ADF regression         *
*       between 0 and   6   lags.                                     *
*      using the BIC model selection criterion.                       *
**********************************************************************

Model Selection Criteria
    Minimum AIC at lag:      3
    Minimum BIC at lag:      3
```

```
************************************************************
* Augmented Dickey-Fuller t-test with 3 lags:        -1.4114     *
*                    1%          5%          10%                  *
*                 -3.75       -3.00       -2.63                   *
*                                                                 *
* Augmented Dickey-Fuller Z-test with 3 lags:        -0.3573     *
*                    1%          5%          10%                  *
*                -17.2       -12.5       -10.2                   *
*                                                                 *
* Coefficient and T-Statistic on the Constant:                    *
*                 0.17894          2.0398                         *
*                                                                 *
* Joint test of a unit root and no constant:          3.9920      *
*                    1%          5%          10%                  *
*                 7.88        5.18        4.12                    *
************************************************************


************************************************************
* TESTING THE NULL HYPOTHESIS OF A UNIT ROOT IN G                 *
*              Using data from 1980:01 to 2003:01                 *
*      Choosing the optimal lag length for the ADF regression     *
*       between 0 and   6   lags.                                 *
*      using the BIC model selection criterion.                   *
************************************************************

Model Selection Criteria
    Minimum AIC at lag:     3
    Minimum BIC at lag:     3

************************************************************
* Augmented Dickey-Fuller t-test with 3 lags:        -3.3858     *
*                    1%          5%          10%                  *
*                 -4.38       -3.60       -3.24                   *
*                                                                 *
* Augmented Dickey-Fuller Z-test with 3 lags:        64.2151     *
*                    1%          5%          10%                  *
*                -22.5       -17.9       -15.6                   *
*                                                                 *
* Coefficient and T-Statistic on the Constant:                    *
*                 3.97304          3.4631                         *
* Coefficient and T-Statistic on the Linear Trend:                *
*                 0.05661          3.3130                         *
```

```
*                                                                    *
* Joint test of a unit root and no linear trend       7.1463       *
*                  1%          5%          10%                      *
*                10.61       7.24        5.91                       *
**********************************************************************

----------------------------------------------------------------------
break date TB = 1996:01   statistic t(alpha==1)        -4.28017
critical values at 1%        5%          10%        50%        90%        95%
99%
  for 70  obs.     -6.32     -5.59     -5.29     -4.24     -3.32     -3.08     -2.67
----------------------------------------------------------------------
     number of lag retained : 1
-----------------------------------------------------
explained variable :    D
-----------------------------------------------------
                              coefficient              student
-----------------------------------------------------
CONSTANT                        3.02198                4.38887
DU                             -0.12211               -0.49965
D(Tb)                           0.16591                2.66896
TIME                            0.06216                4.13886
DT                              0.00570                0.49380
D {1}                           0.04419                0.19787
-----------------------------------------------------
----------------------------------------------------------------------
break date TB = 1993:01   statistic t(alpha==1)        -8.42857
critical values at 1%        5%          10%        50%        90%        95%
99%
  for 70  obs.     -6.32     -5.59     -5.29     -4.24     -3.32     -3.08     -2.67
----------------------------------------------------------------------
     number of lag retained : 0
-----------------------------------------------------
explained variable :    R
-----------------------------------------------------
                              coefficient              student
-----------------------------------------------------
CONSTANT                       -0.08131               -1.55884
DU                              1.09249                3.56980
D(Tb)                           0.47759                4.84384
TIME                            0.02581                4.12288
DT                             -0.06810               -4.14387
R {1}                          -0.33347               -2.10777
-----------------------------------------------------
```

---

break date TB = 1991:01 statistic t(alpha==1) -7.16647

critical values at	1%	5%	10%	50%	90%	95%	99%
for 70 obs.	-6.32	-5.59	-5.29	-4.24	-3.32	-3.08	-2.67

---

number of lag retained : 6

---

explained variable : G

---

	coefficient	student
CONSTANT	8.18209	7.50888
DU	-0.04394	-2.67001
D(Tb)	-0.05085	-4.35135
TIME	0.09490	5.91637
DT	0.01374	7.39782
G {1}	-0.25409	-1.45199

A7 Impulse Response Functions

Impulse Response Function for Coal: Response of Demand to a Shock in Price

	D	P	G
1	0	0.056571	0.003295
2	0.002089	0.0818	0.007942
3	-0.00232	0.087994	0.009761
4	-0.01241	0.084563	0.009619
5	-0.02428	0.077848	0.008922
6	-0.03413	0.071213	0.008436
7	-0.03968	0.06569	0.008248
8	-0.04045	0.060913	0.008051
9	-0.03735	0.056023	0.007466
10	-0.03201	0.050308	0.00625
11	-0.02609	0.043512	0.004378
12	-0.02083	0.035856	0.002016

Impulse Response Function for Oil: Response of Demand to a Shock in Price

	D	R	G
1	0	0.094301	0.005998
2	0.002951	0.118671	0.013482
3	0.008093	0.11494	0.016156
4	0.013242	0.102278	0.015635
5	0.017441	0.087724	0.013682
6	0.020515	0.07348	0.011235
7	0.022574	0.060139	0.008697
8	0.023781	0.047839	0.006223
9	0.024285	0.036605	0.003878
10	0.024212	0.026431	0.001687
11	0.023666	0.017301	-0.00034
12	0.022738	0.009182	-0.00218

Impulse Response Function for Electricity: Response of Demand to a Shock in Price

	D	P	G
1	0	0.151716	0.008127
2	0.017122	0.051965	0.014
3	0.017974	0.014947	0.01693
4	0.015989	0.003245	0.018069
5	0.014384	0.000355	0.018363
6	0.01347	1.83E-05	0.018319
7	0.012995	0.000202	0.018159
8	0.012732	0.000381	0.017969

9	0.012556	0.000479	0.017773
10	0.012411	0.00052	0.01758
11	0.012276	0.000532	0.01739
12	0.012145	0.000533	0.017203