Whenever the production process increases slowly the usual production function i.e., having multiplicative error term do not behave well for forecasting purpose. In this situation, Cobb-Douglas type production function with additive errors may be applicable. In this regard, the estimation procedure is to be complicated. Hence, non-linear estimation procedures are applicable. However, both Cobb-Douglas type function with multiplicative and additive errors are to be estimated here. Different model selection criteria show that Cobb-Douglas type production function with additive errors is more suitable for forecasting the industrial production in Bangladesh.

Currently, Md. Moyazzem Hossain working as an Assistant Professor in the department of Statistics at Jahangirnagar University, Bangladesh. He has published more than 45 research articles in national and international journals. In addition, he is now serving as an editorial board member of four international journals.

978-3-659-77719-6

Production Behaviors of the Industry

Hossain, Majumder

Md. Moyazzem Hossain
Ajit Kumar Majumder

Production Behaviors of Manufacturing Industries in Bangladesh

LAP LAMBERT Academic Publishing

Md. Moyazzem Hossain
Ajit Kumar Majumder

Production Behaviors of Manufacturing Industries in Bangladesh

Md. Moyazzem Hossain
Ajit Kumar Majumder

Production Behaviors of Manufacturing Industries in Bangladesh

LAP LAMBERT Academic Publishing

Impressum / Imprint
Bibliografische Information der Deutschen Nationalbibliothek: Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.d-nb.de abrufbar.

Bibliographic information published by the Deutsche Nationalbibliothek: The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at http://dnb.d-nb.de.

Coverbild / Cover image: www.ingimage.com

Verlag / Publisher:
LAP LAMBERT Academic Publishing
ist ein Imprint der / is a trademark of
ICS Morebooks! Marketing SRL
4, Industriala street, 3100 Balti, Republic of Moldova
Email: info@omniscriptum.com

Herstellung: siehe letzte Seite /
Printed at: see last page
ISBN: 978-3-659-77719-6

Analysis of Production Behaviors of Manufacturing Industries in Bangladesh by Intrinsically Nonlinear Cobb-Douglas Production Function

Contents

List of the Tables

List of Figures

Abstract

This thesis considers the Cobb-Douglas type production functions, which are nonlinear regression model. To estimate the production of a product these nonlinear models are to be fit. But the nonlinear estimation problems are not simple and straight forward; hence some related computational issues to over come this problem which are discussed in this study. Thus appropriate nonlinear production function models with additive error terms are selected by model selection criteria for this study. Time series data for the period 1982-83 through 1997-98 for some selected manufacturing industries are used and the data are collected from Bangladesh Bureau of Statistics (BBS) and CMI. Fitting the model then we have deal with the problem of econometric analysis such as: Autocorrelation and Multicollinearity. We have detected the problems separately and if there exists, the presence of any of these problems, we have tried to remove them with appropriate techniques. Finally, the results obtained in this study have been analyzed and comments are made with the aim of meeting the objectives of the study.

A brief summery of the contents of this thesis is as follows: Chapter 1 furnishes the background; Chapter 2 discusses the relevant literature for this thesis. In Chapter 3 discusses different technique of nonlinear estimation. Chapter 3 also discusses the computational issues that we face to estimate the intrinsically nonlinear production function.

In Chapter 4, we estimate the Cobb-Douglas production function. Chapter 4 also discusses comparison of intrinsically linear and intrinsically non linear model. Hence we observed that our suggested model is better for this study. In chapter 5 we test the violation of autocorrelation for the proposed fitted model and finally we remove the autocorrelation by suitable method. The existence of multicollinearity is observed in chapter 6. In Chapter 7, some conclusions have been drawn about intrinsically nonlinear model for manufacturing production in Bangladesh.

Chapter One

Introduction

1.1 Background

A developing country like Bangladesh which is facing enormous problems so far as industrialization policy is concerned does not follow the policy of Marxian Economy, neither thus it strictly follow the policy of a Capitalist country. The economy of Bangladesh actually turned out to be a mixed economy since a long time. But it could not stand either Marxian economy nor on mixed economy for different situation that occurred in national and international politics. However, the economy of Bangladesh more or less now biased to Capitalism. In these circumstances, it is essential for Bangladesh to go for mass industrialization to strengthen the economy of Bangladesh for this purpose, of course our policy for industrialization must be well-planned, well-defined and well-thoughtful. Strictly speaking, the development of economy is solely dependent on the industrial policies of the country. The factors of production that would be used in a production function will inevitably depend on the import policy of spare parts, intermediate goods and capital goods, etc. A production function gives us indication about the nature of the production inputs used in the production function.

Naturally, it is clear that different types of industries require different levels of factors of production. This depends on the nature, size and pattern of the industry concerned. There is a good number of manufacturing industries in Bangladesh with respect to its assets and demands. Some of the industries, as for instance, garments, textiles, construction, industrial chemical etc. may be profitable and some of them like petroleum refinery, jute, electrical machinery etc. may incur losses. The industry sector was severely damaged during the war of liberation in 1971. Replacement and rehabilitation costs estimated for the industries were estimated at Tk 291 million, of which Tk 223 million was estimated for public sector enterprises. The public sector started in 1972 with 72 Jute mills (with production capacity of 79,200 tons), 44 Textile mills (13.4 million pounds), 15 Sugar mills (169,000 tons), 2 Fertilizer factories (446,000 tons), one Steel mill (350,000 tons), one Diesel engine unit (3,000) and one Shipbuilding yard. Mills and factories in the public sector however, soon became losing concerns largely because of mismanagement and leakage of resources. The government had to quickly review its policy of dominating the public sector. Although it continued to exercise control over industries, it soon raised the allowable ceilings of private investment. However, this did not bring much improvement.

After a series of adjustments and temporary changes in state policy, the government finally adopted a new industrial policy in 1982, following which 1,076 state-owned enterprises were handed over to private owners. Unfortunately, denationalization created a new problem of industries. They started getting sick because of failures of the inexperienced owners. Many of them were more interested in getting ready cash from selling of the cheaply acquired property than in sustaining and developing the industries. The result was that industrial sickness affected 50% of industries in Food manufacturing, 70% of them in Textile, 100% in Jute, 60% in Paper and Paper board, 90% in Leather and Rubber products, 50% in Chemicals and Pharmaceuticals, 65% in Glass and Ceramics, and 80% in Engineering industries. The largest group of industries in Bangladesh falls in the category of small and cottage industries and their number in 1984 was 932,200 units, of which 20.7% were in Handlooms, 15.4% in Bamboo and Cane work, 8.1% in Carpentry, 6.1% in products from Jute and Cotton yarn, 3.4% in Pottery, 0.3% in Oil crushing, 3.2% Blacksmiths, 0.8% in Bronze casting, and the rest in other types of crafts.

In 1984, Bangladesh had 58 textile mills with 6,000 looms and 1,025,000 spindles. The annual production of the mills was 106.2 million pounds of yarn and 63 million metres of cloth. Textile is a public sector dominated industry in Bangladesh and like most other sectors, Textile also incurred losses, which amounted to Tk 353.4 million in 1984. Problems in the sector include poor management as well as difficulties in developing skilled workers and shortage in supply of raw material and power. Bangladesh had 70 Jute mills with 23,700 spindles in 1984. These employed 168,000 workers and 27,000 other staff and used 545,000 tons of raw Jute. But their production was less than the 561,000-tons figure of 1969, when the country had 55

Jute mills with 21,508 spindles. The three major centres of Jute industry in Bangladesh are Dhaka, Chittagong and Khulna. The Jute industry in the country has been declining in the face of competition from India and in an international situation, where Jute goods are being replaced by cheap and durable Plastic products.

Development of new industries like Sulphuric acid, Chemicals, Paper, Caustic soda, Glass, Fertilizer, Ceramics, Cement, Steel and Engineering in Bangladesh was slow in the period before 1985. There were only two plants for production of Sulphuric acid in the country in 1985 and their total production was 5,995 MT, while the production of this important ingredient for industries like Soap, Paper, Cast iron and Steel was 6,466 MT in 1970. Production of Caustic soda in 1985 was 6,787 MT. The soda was used almost entirely in Paper mills. Because of availability of Sand, Salt and Limestone within Bangladesh, the country has a good prospect in developing its glass industry. Dhaka and Chittagong are the two major centres for this industry. The automatic Glass factory at Kalurghat of Chittagong produced 12.9 million sq ft of sheet glass in 1985.

The fertilizer industry in the country uses Natural gas as the main raw material. The fertilizer factories produced a total of 808,660 MT in 1985. 741,463 MT was urea, 9,634 MT Ammonium Sulphate, and 57,563 MT Triple Super Phosphate. The three major factories were at Fenchuganj, Ghorasal and Ashuganj. The total production of cement in the country in 1985 was 292,000 MT. The major industries were at Chhatak and Chittagong. Pakshi of Pabna and Chandraghona of Chittagong were the main locations for the Paper industry in Bangladesh. The total production of Paper in 1985 was about 7,500 MT. In 1985, Khulna had a Newsprint mill with a production capacity of 55,000 MT and a Hardboard mill that produced 1,621 sq metre of hardboard. Around this time Bangladesh also had some mills for production of Particle boards and Partex. The country also achieved self-sufficiency in producing matches; major centres of match production were Dhaka, Khulna, Khepupara, Chittagong, Sylhet, Bogra and Rajshahi. The total production was 1.30 gross boxes in 1985. That year Bangladesh had 8 sugar mills with a total annual production of 87,000 tons. The Sugar mill at Darshana (Ishwardi) produced Sugar as well as

Alcohol, Methilated Spirit and rectified spirit. The Iron and Steel mills in Bangladesh were mostly under the Steel and Engineering Corporation and were concentrated in Chittagong and Dhaka, although there were some Steel and Ironwork enterprises in Khulna, Kushtia and Bogra. Industries marked by notable development in Bangladesh in the mid-1980s include Shipbuilding, Automobiles (assembly), Oil refinery, insulators and Sanitary wares, Telephone equipment, Electrical goods, Televisions (assembly), Cigarette, and Vegetable Oil. The country achieved a significant success in developing Garment industry in this decade. The government followed a strategy of planned growth blended with 'free play' of market forces. The manufacturing sector showed some growth in the 1990s. The share of the manufacturing sector in the country's GDP rose to 11% in 1996. Investment in the sector was Tk 57.8 billion in 1997 as compared to Tk 22.5 billion in 1991. The share of the public sector in the total investment in the country's industries fell from 37.03% in 1991 to 8.63% in 1997. The basic industrial statistics as adopted from the latest census of manufacturing industries in Bangladesh are presented in a table (the census did not cover the cottage industry sector).

Table1.1 Summary statistics of different Manufacturing Industries of Bangladesh in 1999-2000.

	Number		**Value (Million Tk)**
Establishments (by admin. divisions) All	24752	Fixed Assets	243805
Dhaka	11588	Products and by-products total	590865511
Chittagong	4235	Finished products	581177103
Rajshahi	6570	by-products	8011880
Khulna	2314	industrial wastes	1676528
Barisal	45		
Sylhet	404		
Workers Both sexes	2613564	Taxes paid	32222
Male	1699897		
Female	913667	Gross output	639220
All employees Both sexes	2259717	Gross value added	235443
Male	1421734	Value added at factor cost	155820
Female	837983		

Source: Bangladesh Bureau of Statistics

The government continues to implement a Privatization programme to hand over public sector enterprises to private owners. Simultaneously, the government implements a programme of rehabilitating industries identified as sick because of various reasons. Industries identified for rehabilitation under the programme in 2000 included one Cement factory (annual production capacity 0.15 million tons), one Paper mill (30,000 tons), one Newsprint mill (52,000 tons), 6 Cigarette factories (630 million sticks), 8 Oil mills (934,818 tons), 2 Food processing units (950,400 tons), 2 Fish processing units (6.9 million tons), 2 Cold storages (5.9 million lbs), one Beverage producing unit (4.3 million bottles), 3 Chemical industry units (26,100 tons), one Glass factory (7.5 million feet) and 12 Pharmaceutical units.

The fifth Five-Year Plan for the period 1997-2002 stipulated a total outlay of Tk 8.95 billion in industry including Tk 1.39 billion in the private sector. In 2000, the total employment in industries was estimated at 0.6 million, of which the private sector employed 0.5 million. Industrialisation efforts of the government during the 1990s included investment in balancing, modernisation and reconstruction, creation of new industrial estates and export processing zones, promotion of private investment, and attraction of foreign direct investment. The policy changes have been in line with trends in the international market, recommendations of donor countries and agencies for liberalization of trade and investment and Structural Adjustment Programmes.

Almost at regular intervals of 4 to 6 years after 1982, the government adopted new industrial policies with increased incentives for private investors from both home and abroad. These policies have some common aspects such as incentives to promote industrialization in rural and remote areas and to encourage entrepreneurs to use local raw materials, and the efforts towards development of a system that would help in transfer of technology.

Manufacturing industry in Bangladesh achieved respectable growth during 1990s (Table 2). The contribution of manufacturing to GDP increased from 12.9 per cent in 1990-91 to 15.4 percent in 1999-00. However, the sector's current share in GDP appears rather modest for it to spearhead sustained high growth of the economy. The

growth of Bangladesh's manufacturing sector has also been rather narrowly based with readymade garments accounting for nearly a quarter of the sectoral growth.
Other important export industries contributing to sectoral growth are Fish & Seafood, and Leather tanning. Major import substituting industries experiencing significant growth during this period include Pharmaceutical, Indigenous cigarettes (*bidi*), Job printing and Re-rolling mills.

Table 1.2 Growths of Manufacturing and Share in GDP.

Year	Yearly Growth (%)	Share in GDP (%)
1990-91	6.4	12.9
1991-92	7.4	13.3
1992-93	8.6	13.8
1993-94	8.1	14.4
1994-95	10.5	15.1
1995-96	6.4	15.4
1996-97	5.0	15.4
1997-98	8.5	15.9
1998-99	3.2	15.6
1999-00	4.8	15.4
2000-01	9.1	15.8

Source: Bangladesh Bureau of Statistics

Weavers work in almost all parts of Bangladesh but their major concentration is in areas like Narsingdi, Baburhat, Homna, Bancharampur, Bajitpur, Tangail, Shahjadpur and Jessore. The Silk industry has flourished in Rajshahi and Bholarhat. Other places earning reputation in cottage industries during the 1980s in Bangladesh include Chapai Nawabganj and Islampur (bronze casting), Sylhet (mat and cane furniture), Comilla (pottery and bamboo work), Cox's Bazar (cigar), Barisal (choir) and Rangpur (checkered carpet).

From the Figure 1.1, we see that, most of the industries of Bangladesh are situated at Dhaka, Chittagong, and Rajshahi based on some facilities.

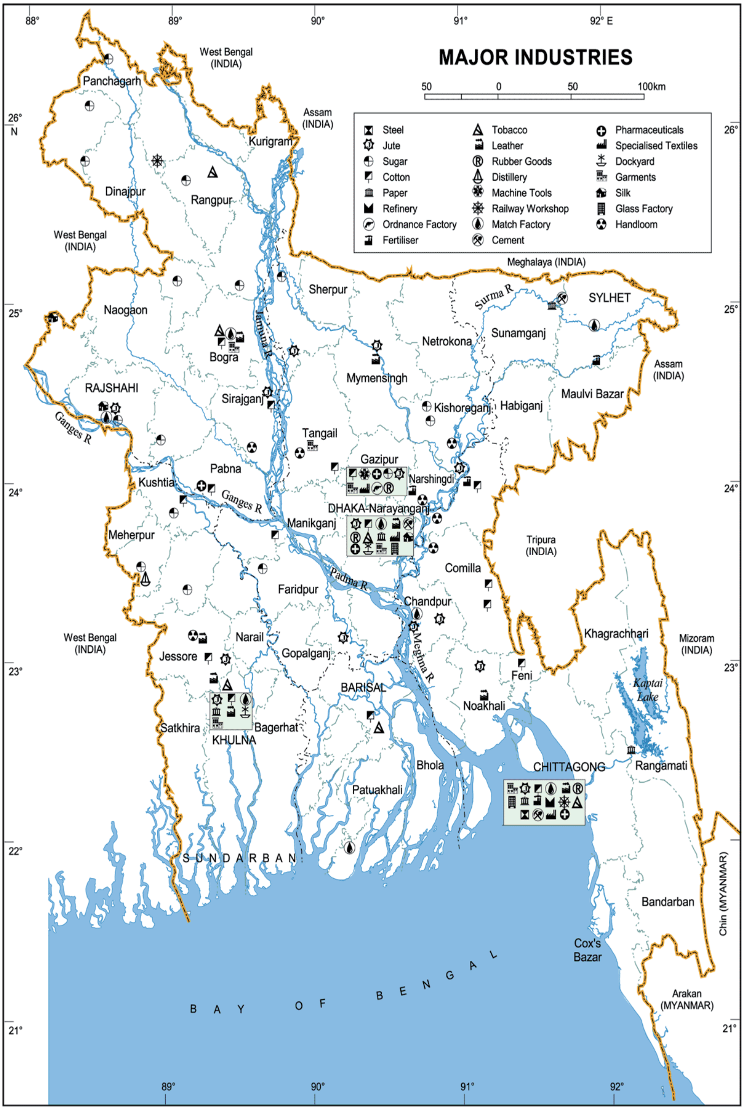

Figure1.1 The position of manufacturing industries of Bangladesh.

1.2 Problems and Motivation

Bangladesh is industially a less developed country among the developing countries of South Asia. Industrial sector contributes to the economic development. Out of these industries, 60% are agro-based, such as jute, textile, paper, sugar etc. The population of Bangladesh has increased but the land for cultivation has decreased in course of time. That is why it is the crutial time to find the alternatives.

Every industrialist tries to produce goods with maximum profit but with minimum cost. In order to do this, it has to be decided what to produce, how much to produce and how to produce. The industries need various inputs such as labor, raw material, machines etc. to produces goods.

An industy's production cost depend on the quantities of inputs it buy and on the prices of each input. Thus an industry needs to select the optimal combination of inputs, that is, the combination that enables it to produce the desired level of output with minimum cost and hence with maximum profitability.

In developing counties, efficiency of economic development has determined by the analysis of industrial production. An examination of the characteristic of industial sector is an essential aspect of growth studies. The most of the developed countries are highly industrialized as they brief "The more industrialization, the more development".

For proper industrialization and industrial development we have to study industial input-output relationship that leads to production analysis. For a number of reasons econometrician's belief that industrial production is the most important component of economic development because

- ❖ If domestic industrial production increases, GDP will increase.
- ❖ If elasticity of labor is higher, implement rates will increase.
- ❖ Investment will increase if elasticity of capital is higher.

In the present times, production takes place by the combination forces of various factors of production such as land, labor, capital etc. In this connection, Socialist countries are using different patterns of level of factors of production for their respective industrialization policy according to the taste, demand and nature of their country-wide population, its size, location and environment. Bangladesh are a developing country. It is essential for Bangladesh to go for mass industrialization to strengthen the economy of Bangladesh for this purpose, of course our policy for industrialization must be well planned, well defined and well thoughtful.

It is obvious that the development of economy is solely dependent on the industrial polices of the country. By using production function we can get industrial policies especially indication about the nature of the production inputs used in the production function. The production function may be intrinsically linear and intrinsically nonlinear. Also we know that the parameters are restricted in the production function.

In this thesis, we develop a model which gives optimal combination of inputs, that is, the combination that enables it to produce the desired level of output with minimum cost and hence with maximum profitability.

1.3 Objectives of the Study

The main objectives of this study are:

i) to identify the industrial policies espicially identify the nature of the production inputs.

ii) to estimate the labor elasticity of substitution as well as the capital elasticity of substitution for each of the selected manufacturing industries of Bangladesh.

iii) to indentify the appropriate Cobb-Douglas production function.

iv) to discuss the computational issues of estimating non-linear production function.

1.4 Data Collection

In recent publications of "Statistical Yearbook of Bangladesh" and "Report on Bangladesh Census of Manufacturing Industries (CMI)" published by BBS, we get the published secondary data for the major manufacturing industries of Bangladesh.

1.5 Variables Considered in the Study

Perhaps the most important consideration in productivity measurement is the measure of inputs and output. The reliability of performance measures of economic agents hinges on accuracy of measures of output and inputs.

Subject to the availability of data, we consider three variables in this study. The Gross output consider as the dependent variable and the two independent variables are

i. Total fixed assets
ii. Total person engaged

1.5.1 Gross output

It's the total yearly output of the particular Manufacturing Industry in Bangladesh measured in terms of money, usually in thousand taka.

1.5.2 Independent variables

- ***Total fixed assets***

It refers to the assets of a particular Manufacturing industry of Bangladesh, whether obtained from other enterprises or produced by the out of its own resources for its own use, which are expected to have a productive life more than one year. It includes land, building, other constructions, machineries, tools and equipments, transport etc.

- ***Total person engaged***

It refers to the persons both employee and workers who are engaged in the production of an industry in Bangladesh.

1.6 Limitation of the Study

Due to unavailability of the data we could not developed a model by using appropriate variables. We only use production, total fixed assets and total person engaged to develop nonlinear model. On the basis of the available of the data we collect data over the period 1982-83 to 1997-98. Thus we can precisely identify the limitation of our study as follows:

i) Due to unavailable information in Bangladesh Bureau of Statistics (BBS) we cannot use the recent data of the manufacturing industries.

i) Estimation of nonlinear model is a serious problem that we observe in the study.

1.7 Outline of the Study

This research is organized into seven chapters. Chapter 2 consists of different types of production function with their properties and limitations. Chapter 3 discusses different nonlinear regression model and nonlinear estimation methods. This chapter also contains the computational issues of our analysis. In chapter 4, we estimate the parameters of Cobb-Douglas production function. This chapter also contains different model selection criteria and based on them, we select the best model for the data under study.

Chapter 5 contains brief description of autocorrelation. This chapter also discusses the tests of detecting autocorrelation and the remedial measures. In chapter 6, we discuss briefly about multicollinearity. A summary with some concluding remarks and some suggestions for further research is contained in the final chapter.

1.8 Computation

When the model is intrinsically non-linear, it is very difficult to estimate the parameters by ordinary software. In some cases SPSS, SAS does not provide the best solution. The optimal solution depends on the initial value. By trial and error method we find the appropriate initial value and it is time consuming. It's a great problem. In this situation we write program and we can handle this problem easily. All the programs used to estimate the results in this thesis are written in GAUSS System Version 3.2.18. We used Gauss built in optimization routine to estimate optimal parameters to estimate the production.

Chapter Two

Different Types of Production Function

2.1 Introduction

Productive activities are as diverse as life itself. An agriculture farm takes fertilizer, seed, land and labor and turns them into wheat or corn. Modern factories take inputs such as energy, raw materials, computerized machinery and labor and use them to produce tractors, TV's or tubes of toothpaste. An accounting firm takes pencils, computers, paper, office space and labor and produces audits or tax returns for its clients. The agriculture firm; factory and accounting firm always attempt to produce the maximum level of output for a given dose of inputs avoiding waste whenever possible. Now, if we discuss about inputs like land, labor and capital and also outputs like wheat, TV's and profits; a question may be arise, if we have a fixed amount of inputs, how much output can we get?

In order to get the answer of above question, firstly, we have to establish the relation between inputs and outputs. In the view of economists, the relationship between the amount of input required and the amount of output that can be obtained is called the production function. We know, there are a number of production functions developed by the economists. The most frequently used production functions are Cobb-Douglas production function, Constant Elasticity of Substitution (CES) production function and Variable Elasticity of Substitution (VES) production function. There are some other types of production functions; however, they are seldom used in econometric research. Secondly, we have to choose or select a production function and estimate it and thus for a fixed amount of input, we can easily get estimated output.

The organization of the present chapter is as follows: Section 2.2 represents the definition of production with its drawbacks. Section 2.3 explores different types of model. In Section 2.4, we describe different types of production function with its merits and limitations.

2.2 Production Function

Businessmen as well as industrialists concerning the theory of the firm make decisions about what, how much to produce, and how to produce it. Such decisions are related to costs of production, the kind of market in which the firm operates, and the internal organization of the firm.

The starting point for the theory of the firm usually assumes profit-maximizing business operating in a competitive market and producing only a single good (Hurbury, p-37, 1968).The firm is the basic production unit, producing goods (and services, such as transporting, financing, whole-selling and retailing) using certain inputs called "factors of production" such as labor and capital (Intriligator, M.D., p-251, 1980).

Production functions are one of the most basic economic relationships, and ideally they express the technical relationships between physical quantities of inputs to the production process, and the physical quantity of output produced. This description should be applied at the micro-economic level, for a single firm or a single homogenous product, and a few homogeneous inputs. Production function, in this sense, shows up a one-way relationship in the causal flow.

It is, therefore, important in theories of economic growth and in theories of distribution. At the micro-economic level it is of interest because of its usefulness in the analysis of such problems as the degree of which substitution between the various factors of production is possible and extent to which firms experience decreasing or increasing returns to scale as output expands. At both the macro and micro-levels, the production of any increase in output over time can be attributed to, firstly, increases in the inputs of factors of production; secondly, to the existence of increasing returns to scale; and thirdly, to what is commonly referred to as 'technical progress' (Thomas, R. L., p-208, 1985).

The concept of a production function plays an important role in both micro and macro economics. Production functions are one of the most basic econometric relationships, and ideally they express the technical relationships between physical quantities of inputs to the production process, and the physical quantity of output produced. Mathematically, such a basic relationship can be expressed as:

$$P = f(A, B, C, ...)$$

where, P is the output (dependent variable); A is the labor (independent variable); B is the capital (independent variable); C is the land (independent variable) and other natural.

2.2.1 Drawbacks of production function

The Econometricians try to estimate the form of this theme, it would have many welcome features: it would be in physical units, so shifting monetary values would not affect it; it would be unambiguous one-way in the causal flow. It is expressed-from given inputs to the resultant output; it would stand independent of other economic relationships such as input supply functions and output demand functions. Its parameters, when estimated, would describe the technological features of a well-defined productive process, and would be given concrete interpretation.

We regret to speak that there is no such production function which ensures this theme. Further, no-one is really interested in the production function for a single firm or a single, simple process. They usually relate to a whole industry or even a whole economy, and thus move to the micro-economic level. Then we are really aggregating over a set of different production functions and different techniques, and the procedure will become more and more meaningless as we aggregate further and further, till the result is so much of a bland 'average' that it bears no economic meaning any more, in the sense of describing the technical features of the transformation of inputs into output (Hebden, J., pp-87-88, 1983).

Through aggregation the causal direction has gone from one extreme (input → output) to the other (output → inputs). However, the truth lies between these, and makes added complications: output and inputs are usually decided upon together by the producer, or at least some degree of feedback from output to inputs occurs, either directly or via the ancillary relationships of supply and demand functions (Hebden, J., pp-87-88, 1983).

This feedback makes it incorrect to regard output as the dependent variable and inputs as exogenous variables in the production function, at whatever level of aggregation. So when production functions are estimated in isolation by simultaneous equation least squares methods, their estimates are all subject to bias because the simultaneous context of the production function has been ignored. However, this does not appear to have deterred many distinguished researchers (Hebden, J., pp-87-88, 1983).

2.3 Different Types of Model

In statistical analysis to solve the problem it is necessary to know the type of model. The model is broadly classified into two types:

- Linear regression model and
- Non-linear regression model

In the linear regression model the parameter are in linear. The simple linear regression model, classical normal linear regression models are the examples of linear regression models.

In statistical analysis linear regression models are very commonly used technique. But in real field most of the situation arises where variable follows non-linear model. For example, the most popular population growth model is nothing but a non-linear model. Non-linear models tend to be used either when they are suggested by theoretical consideration or to build known non-linear behavior into a model. Even when a linear approximation works well, a non-linear model may still be used to retain a clear interpretation of the parameters (see Seber and Wild, (1989), pp.-5). In non-linear model the parameters are in non-linear. The polynomial regression model, logit model, probit model, tobit models are the non-linear regression models.

Since the main concern is to fit a production function for this study, so we classify the non-linear model according to this production function. Here we can split the structure of non-linear model into three categories, which are discussed in below.

2.3.1 Intrinsically linear model

When the non-linear model can be converted in a linear model then the model is known as intrinsically linear model. The form of this model is as,

$$y = \alpha X^{\beta} u \tag{2.1}$$

$$y = \alpha X^{\beta} e^{-\lambda t} u \tag{2.2}$$

Generally, in such model errors are multiplicative in nature and we get converted linear model by taking log transformation.

2.3.2 Intrinsically non-linear model

When the non-linear model can be converted in a linear model then the model is known as intrinsically linear model. The form of this model is as,

$$y = \alpha X^{\beta} + u \tag{2.3}$$

$$y = \alpha X^{\beta} e^{-\lambda t} + u \tag{2.4}$$

2.3.3 Partially non-linear model

When the values of the parameter(s) is given, which are in non-linear, then the non-linear model turns to a linear model and this type of model is known as partially non-linear model. For example, if the value of β, which is in non-linear, is known then the model (2.3) is written as,

$$y = \alpha Z + u \tag{2.5}$$

Where, $Z = X^{\beta}$, is known, then the model (2.8) is treated as linear model and known as a partially non-linear model.

And also for the model (2.7), when the value of β and λ are given then the model reduce to a linear model as,

$$y = \alpha Z V + u \tag{2.6}$$

Where, $Z = X^{\beta}$ and $V = e^{-\lambda t}$, are known then the model is partially non-linear model. Cobb-Douglas-type production function may take also this feature.

2.4 Different Types of Production Function

In Econometrics we use the following production function:

- Cobb-Douglas production function.
- C.E.S production function.
- Variable Elasticity of Substitution (VES) production function.
- Nerlove-Ringstand production function.
- Zellner-Revankar production function.

2.4.1 Cobb-Douglas production function

The Cobb-Douglas production function is the widely used function in Econometrics. A famous case is the well-known Cobb-Douglas production function introduced by Charles W. Cobb and Paul H. Douglas (1928), although anticipated by Knut Wicksell (1901: p.128, 1923) and, some have argued, J. H. von Thünen (1863). They have estimated it after studying different industries in the world, for this it is used as a fairly universal law of production.

The Cobb-Douglas production function with multiplicative error term can be represented as,

$$P = AL^{\alpha}K^{\beta}U \qquad (2.7)$$

Where, P is the output; L is the Labor input; K is the Capital input; A is a constant; U is the random error term. α and β are positive parameters and $\alpha > 0$, $\beta > 0$, $L > 0$ and $\alpha + \beta = 1$.

The Cobb-Douglas production function with additive error term can be represented as,

$$P = AL^{\alpha}K^{\beta} + U \qquad (2.8)$$

Where, P is the output; L is the Labor input; K is the Capital input; A is a constant; U is the random error term. α and β are positive parameters and $\alpha > 0$, $\beta > 0$, $L > 0$ and $\alpha + \beta = 1$.

2.4.1.1 Economic significance of the Cobb-Douglas production function

The common economic significance of the Cobb-Douglas production function is as follows:

i) It provides important information regarding industrial and agricultural sectors through which we can make different policies (Singh, S.P., p-276, 1977).

ii) Since we can determine the marginal productivity of capital as well as labor, it is helpful in wage determination principles (Singh, S.P, p-276, 1977).

iii) The estimated elasticity of co-efficient may be helpful in international or inter-sectoral comparisons.

iv) It helps us in the study of the different laws of returns to scale.

v) It provides us information regarding the substitutability of the factors of production.

vi) It is used to determine the degree of homogeneity (Singh, S.P., p-276, 1977).

2.4.1.2 Drawbacks of the Cobb-Douglas production function

The main drawbacks of the Cobb-Douglas production function are discussed below:

i) Why have we chosen for a multiplicative, or log linear form? Does this ensure correctness and applicability? A linear form might fit the data equally well. When our data do not include extreme values, a linear and a by-linear form are often hard to distinguish. However, the multiplicative form is not a guess, but result of prior theory as how the productive process works; linear form would be easy for regression, but would imply nonessential features like ever-increasing or ever-decreasing marginal returns to input (Hebden, J., pp.-97-100, 1983).

ii) Since we have taken only two inputs instead of multiple inputs, then the estimates of the parameters, A, α, β will be distorted, biased, and the distribution of the residual may be distorted too. This bias in the estimates may lead one to conclude $\alpha+\beta>1$, i.e., increasing returns to scale exist, while actually $\alpha+\beta=1$. Griliches (1964) has included more than just K and L as inputs and thereby lessened the possibility of biased estimates and reduced the role of A (Hebden, J., pp.-97-100, 1983).

iii) If returns to scale are constant, then the Cobb-Douglas function cold is simplified to:

$$Q = AK^{\alpha} L^{1-\alpha} u \qquad \text{Where } \alpha+\beta=1$$

Or, $$\frac{Q}{L} = A\left(\frac{L}{K}\right)^{\alpha}$$

This form is a simple one rather than a multiple regression, with only A and α to be estimated. However if we impose constant returns from the outset, we are forcing the data into a form they may not fit. It would be much better to put no constraint on the sum $\alpha+\beta$; after fitting the regression.

$$\log Q = \log A + \alpha \log K + \beta \log L$$

We would test the hypothesis of constant returns with a student's t-test, thus:

$$H_0 : \alpha+\beta = 1 \qquad \text{against} \qquad H_1 : \alpha+\beta > 1$$

If the alternative hypothesis refers to increasing returns to scale.

Test statistic: $$t_{(n-3)} = \frac{(\alpha+\beta)-(\alpha+\beta)}{\sqrt{\text{var}(\alpha+\beta)}}$$

$$= \frac{(\alpha+\beta)-1}{\sqrt{\text{var}(\alpha)+\text{var}(\beta)+2\,\text{cov}(\alpha,\beta)}}$$

If H_0 is true, so a simple significance test can be made.

iv) The trouble with the regression form

$$\ln Q = \ln A + \alpha \ln K + \beta \ln L$$

is that we may encounter severe multicollinearity between L and K, especially with cross-section (rather than time series) data and especially if our observations are for firms in a fairly homogeneous industry.

v) If we assume the Cobb-Douglas function to apply at the level of a single firm, it will not apply to an aggregate of whole firms; while, conversely, if it applies to an industry, it will not apply to the firms that make up that industry.

vi) Should output be measured gross or net? Our choice was for the treatment of raw materials. If we measure the output as 'value added', then our production function is really like the model given bellow:

$$Q - M = AK^{\alpha}L^{\beta}$$

Where, $Q - M =$ value added, Q =gross output, M = raw materials and fuel Implying that

$$Q = AK^{\alpha}L^{\beta} + M$$

Which in turn implies that materials enter the production function additively? But why should it not be:

$$Q = AK^{\alpha}L^{\beta}M^{\gamma}$$

i.e., materials entering multiplicatively and thus contributing to the degree of outweighing to scale? This is more realistic, since at different bends of production there may be economies or diseconomies in the use of raw materials and fuel, rather than the constant relationship of M to Q that the value added form implies.

vii) In the Cobb-Douglas function, α and β are by definition constants, regardless of the levels of output or input. But they represent input elasticities and factor shares, both of which might vary with output or input levels. Similarly, returns to scale are fixed ($\alpha + \beta$) for all output levels-which is unrealistic. Moreover by definition we know that elasticity of substitution may be positive or negative, if positive it can be greater than 1 or less than 1. We also get various values of α and β for various industries.

viii) The Cobb-Douglas function provides a single equation system. Though we may determine the supply of inputs and demands for output which would provide a simultaneous equation system. Due to lack of this, we get a biased estimate.

2.4.1.3 Common criticisms about the restriction Cobb-Douglas production function

The common criticisms about the restriction Cobb-Douglas production function forms are:

- It cannot handle a large number of inputs.
- The function is based on restrictive assumptions of perfect competition in the factor and product markets.
- It assumes constant returns to scale (CRS).
- Serial correlation and heteroscedasticity are common problems that beset this function too.
- Labor and capital, are correlated and the estimates are bound to be biased.
- Unitary elasticity of substitution is unrealistic.
- It is inflexible in form.
- Single equation estimates are bound to be inconsistent.
- Other criticisms relate to the level of aggregation and nature of technology.
- It cannot measure technical efficiency levels and growth very effectively.

2.4.2 C.E.S production function

A production function which has a constant elasticity of Substitution (not necessary equal to unity) is known as the constant Elasticity of substitution (CES) production function. This production function was derived independently by two different groups of econometricians. One consisting of K.J. Arrow, Chenery, B.C. Minhas (in 1961) and R. M. Solow and the other group consisting of M. Brown and De Cani. It can be expressed as

$$P = \gamma[\delta C^{-\alpha} + (1-\delta)N^{-\alpha}]^{-\frac{v}{\alpha}}; \gamma > 0; 0 < \delta < 1; \alpha > 1 \qquad (2.9)$$

Where, P = Out put, C = Capital input, N = Labor Input, α = Substitution Parameter, γ =Technical efficiency coefficient, δ = Coefficient of capital intensity, $1-\delta$ = Labor intensity coefficient and V = Degree of homogeneity.

The other forms of CES production function which give the same results are:

i) $P=\gamma[\delta C^{-\alpha}+(1-\delta)N^{-\alpha}]^{-\frac{v}{\alpha}};\alpha=-(1-\frac{1}{\sigma})$

ii) $P=\gamma[\delta_1 C^{-\alpha+\delta^2}N^{-\alpha}]^{-\frac{v}{\alpha}};\alpha=-(1-\frac{1}{\sigma})$

iii) $P=\gamma[\delta C^{\alpha'}+(1-\delta)N^{\alpha'}]^{\frac{v}{\alpha'}};\alpha'=(1-\frac{1}{\sigma})$

iv) $P=\gamma[\frac{N^{\alpha}C^{\alpha}}{\delta N^{\alpha}+(1-\delta)C^{\alpha}}]^{\frac{v}{\alpha}};\alpha=(1-\frac{1}{\sigma})$

Where, $\sigma=$ Elasticity of Substitution.

2.4.2.1 Properties of C.E.S. production function

The main properties of C.E.S production function are given below:

i) If the production function is linear and homogeneous then the elasticity of substitution is $\sigma=\frac{1}{1+\alpha}$, if and only if the function is

$$P=\gamma[\delta C^{-\alpha}+(1-\delta)N^{-\alpha}]^{-\frac{v}{\alpha}} \text{ such that } \gamma>0,\ 0<\delta<1,\ \alpha>-1$$

ii) The marginal product of C.E.S production function is zero.

iii) In the C.E.S production function, the marginal product curves are falling downwards i.e.,

$$\frac{\delta^2 P}{\delta N^2}<0 \text{ and } \frac{\delta^2 P}{\delta C^2}<0.$$

2.4.2.2 Advantage of C.E.S production function over Cobb-Douglas production function

The main advantages of C.E.S production function over Cobb-Douglas production function are given below

i) The C.E.S production function represents the more general form of production technique than Cobb-Douglas production function. In the C.E.S production function the elasticity of substitution is constant and not necessarily equal to unity i.e. $\sigma\neq 1$.

ii) The C.E.S production function has more important parameters than Cobb-Douglas. Therefore it has wider scope, substitutability and efficiency.

iii) The Cobb-Douglas production function is a special case of C.E.S relation. If we put $\alpha = 0$ in the C.E.S relation we shall get the situation of Cobb-Douglas production function.

iv) The C.E.S function is easy to estimate the parameters. It also removes all the difficulties and unrealistic assumptions of Cobb-Douglas production function.

2.4.2.3 Limitations of C.E.S production function

Although C.E.S production function has removed all difficulties and unrealistic assumptions of Cobb-Douglas function, and has wide scope in Econometrics, it also has some limitations. The main limitations are as follows:

i) The C.E.S production function does consider only two factors i.e. N and C. It is not applicable for more factors of production. Professor H. Uzawa has concluded that it is difficult to generalize it to n-factors of production.

ii) The C.E.S production function contains one parameter, namely v which is affected by the scale of operation and technological change. These two forces may affect the degree of returns to scale but can't distinguish them separately.

iii) It is assumed in this function that elasticity of substitution (σ) changes in response to technology only and factor proportions do not affect it. While the empirical study shows that the elasticity of substitution also changes due to changed factor proportions. The function has ignored this very important fact.

iv) Lastly, δ (parameter) of C.E.S function is not dimensionless. Besides it, there is also the difficulty of fitting the data to this function.

2.4.3 Variable Elasticity of Substitution (VES) production function

Once the assumption of a unitary elasticity of substitution σ implicitly assumed in the Cobb-Douglas function had been superseded by the merely constant σ of the more general CES function, it was clear that the next stage would be the developments of variable elasticity of substitution (VES) production functions.

Reasons why the elasticity of substitution should vary are not hard to find. For example, σ might be expected to vary with the capital-labor ratio $K-L$. The greater is the ratio, i.e., the greater the capital intensity of production, the harder it is likely to be to substitute further capital for labor and the lower σ is likely to be. Alternatively, even with a constant $K-L$ ratio, the elasticity of substitution may simply change over time if technical progress affects the case with which factors may be substituted for each other (Thomas, R. L., p-241, 1985).

Christensen, Jorgenson and Lau (1973) finally presented a general form for VES functions. The problem had always been that of finding a functional form that not only allowed for a variable elasticity of substitution but also was easily estimable and could be considered an efficiently close approximation to whatever the underlying productive process actually was. The 'transcendental logarithmic' or translog production function is of the form

$$Q = AK^{\alpha}L^{\beta}e^{\alpha' K+\beta' L}; A>0, \alpha', \beta' \leq 0 \qquad (2.10)$$

This case reduces to the Cobb-Douglas if α' and β' vanish.

Taking logarithms we obtain

$$\ln Q = \ln A + \alpha \ln K + \beta \ln L + \alpha' K + \beta' K$$

So $\ln Q$ is a linear function of the inputs K and L, as well as the logarithms of the input in $\ln K$ and $\ln L$. For this reason it is possible for eventually falling. This function also permits variable elasticity of substitution over the range of inputs.

2.4.4 Nerlove-Ringstand production function

Nerlove (1971) and Ringstand developed a production function which is a general generalization of the Cobb-Douglas production function , of the form ,

$$Q^{1+c\ln Q} = AK^{\alpha}L^{\beta}, \ c \geq 0 \qquad (2.11)$$

This case reduced to the Cobb-Douglas form if $C=0$.

2.4.5 Zellner-Revankar production function

Revankar (1971) developed a model in which a linear function of the capital/labor ratio was used and Sato and Hossman (1968) presented a series of forms-one similar to Revankar's and another in which σ varied over time. Zellner-Revankar production function is one of the generalizations of the Cobb-Douglas production functions , written in the form

$$Qe^{cQ} = AK^{\alpha}L^{\beta}, \quad c \geq 0 \tag{2.12}$$

where, P is the output; L is the Labor input; K is the Capital input; A is a constant; U is the random error term. α and β are positive parameters and $\alpha > 0$, $\beta > 0$, $L > 0$ and $\alpha + \beta = 1$.

This case reduces to the Cobb-Douglas form if $c = 0$. Taking logarithms,

$$\ln Q + cQ = \ln A + \alpha \ln k + \beta \ln L$$

$$= a + \alpha \ln k + \beta \ln L \tag{2.13}$$

where, $a = \ln A$.

In the transcendental case inputs in this case output and the logarithms of input enter on the left-hand side (Intriligator, M. D., p-279, 1980).

2.5 Concluding Remarks

In this chapter we discuss five types of production functions. From among these types, we consider Cobb-Douglas production function for some of its well-known properties.

Chapter 3

Estimation of non-linear production function

3.1 Introduction

Nonlinear Estimation will compute the relationship between a set of independent variables and a dependent variable. For example, we may want to compute the relationship between the inputs and output of a firm, the dose of a drug and its effectiveness, the relationship between training and subsequent performance on a task, the relationship between the price of a house and the time it takes to sell it, etc. We may recognize research issues in these examples that are commonly addressed by such techniques as multiple regression or analysis of variance. In fact, we may think of Nonlinear Estimation as a generalization of those methods. Specifically, multiple regression (and ANOVA) assumes that the relationship between the independent variable(s) and the dependent variable is linear in nature. Nonlinear Estimation leaves it up to us to specify the nature of the relationship; for example, we may specify the dependent variable to be a logarithmic function of the independent variable(s), an exponential function, a function of some complex ratio of independent measures, etc.

When allowing for any type of relationship between the independent variables and the dependent variable, two issues raise their heads. First, what types of relationships "make sense", that is, are interpretable in a meaningful manner? Note that the simple linear relationship is very convenient in that it allows us to make such straightforward interpretations as "the more of x (e.g., the higher the price of a house), the more there is of y (the longer it takes to sell it); and given a particular increase in x, a proportional increase in y can be expected." Nonlinear relationships cannot usually be interpreted and verbalized in such a simple manner. The second issue that needs to be addressed is how to exactly compute the relationship, that is, how to arrive at results that allow us to say whether or not there is a nonlinear relationship as predicted.

This chapter is organized into the following sections. Section 3.2 represents the estimation procedure of non linear model. In section 3.3, contains some computational issues.

3.2 Estimating Nonlinear Models

Technically speaking, Nonlinear Estimation is a general fitting procedure that will estimate any kind of relationship between a dependent (or response variable), and a list of independent variables. In general, all regression models may be stated as:

$$y = F\left(x_1, x_2, ..., x_n\right) \tag{3.1}$$

In most general terms, we are interested in whether and how a dependent variable is related to a list of independent variables; the term $F(...)$ in the expression above means that y, the dependent or response variable, is a function of the X's, that is, the independent variables.

An example of this type of model would be the linear multiple regression model. For this model, we assume the dependent variable to be a linear function of the independent variables, that is:

$$y = a + b_1 X_1 + b_2 X_2 + ... + b_n X_n \tag{3.2}$$

Nonlinear Estimation allows us to specify essentially any type of continuous or discontinuous regression model. Some of the most common nonlinear models are probit, logit, exponential growth, and breakpoint regression. However, we can also define any type of regression equation to fit to our data. Moreover, we can specify either standard least squares estimation, maximum likelihood estimation (where appropriate), or, again, define our own "loss function".

To estimate the production function we need to know different types of non-linear estimation. In non-linear model it is not possible to give a closed form expression for the estimates as a function of the sample values, i.e., the likelihood function or sum of squares cannot be transformed so that the normal equations are linear. The idea of using estimates that minimize the sum squared errors is a data-analytic idea, not a statistical idea; it does not depend on the statistical properties of the observations. The other properties of the estimates do depend on the statistical model (see Christensen (2000), pp.-223).

The idea of least squares estimation can be extended to very general non-linear situations.
The method which are used to estimate the parameters in non-linear system are:

- Non-linear least squares
- Linear approximation
- Numerical method
  - Gauss-Newton Algorithm
  - Newton-Raphson method
  - Steepest descent method
  - Marquardt method
  - Quadratic hill climbing method
  - Rank one correlation method
  - Davidon-Fletcher-Powell method
  - Method of scoring
  - Brown-Dennis method
  - Direct search method

3.2.1 Non-linear least squares

Suppose that we have n observations (x_i, y_i), $i = 1, 2, ..., n$, from a fixed regressor non-linear model,

$$y_i = f(x_i;\theta) + \varepsilon_i \qquad (3.3)$$

Where, $E[\varepsilon_i] = 0$, x_i is a $k \times 1$ vector, and the least squares estimate of θ, denoted by $\hat{\theta}$, minimize the error sum squares,

$$S(\theta) = \sum_{i=1}^{n} \left[y_i - f(x_i;\theta) \right]^2 \qquad (3.4)$$

over $\theta \in \Theta$, a subset of R^p. Unlike the linear least squares situations, $S(\theta)$ may several relative minima in addition to the absolute minimum $\hat{\theta}$. We cannot be sure that non-linear least squares is the most efficient estimator, however except in the case of normally distributed disturbance (see Seber and Wild (1989), pp.-20-22). If we assume that the ε_i are normally distributed, then $\hat{\theta}$ is also the maximum likelihood estimator.

When each $f(x_i;\theta)$ is differentiable with respect to θ, and $\hat{\theta}$ is in the interior of θ, $\hat{\theta}$ will satisfy

$$\left.\frac{\partial S(\theta)}{\partial \theta_r}\right|_{\hat{\theta}} = 0; r = 1, 2, ..., p. \tag{3.5}$$

The equations $\frac{\partial f(x_i;\theta)}{\partial \theta'}$ are called the normal equations for the non-linear model. For most non-linear models they cannot be solved analytically, so that iterative methods are necessary.

3.2.2 Linear approximation

The linear approximation (or linearization) method uses the results of linear least squares in a succession of stages. We begin by first noting that in a small neighborhood of θ^*, the true value of θ, we have the linear Taylor expansion,

$$f_i(\theta) \approx f_i(\theta^*) + \sum_{r=1}^{p} \left.\frac{\partial f_i}{\partial \theta_r}\right|_{\theta^*} (\theta_r - \theta_r^*)$$

or,

$$f(\theta) \approx f(\theta^*) + F.(\theta_r - \theta_r^*) \tag{3.6}$$

where $F. = F.(\theta) = \left(\frac{\partial f(\theta)}{\partial \theta'}\right)$. When n is large we get $\hat{\theta} - \theta^* \approx \theta - \theta^*$ and $\hat{\theta} - \theta^* = (F.'F.)^{-1} F.'\varepsilon$ (see Seber and Wild (1989), pp.-23).

Hence we get the sum squares from (3.6) as,

$$
\begin{aligned}
S(\theta) &= \left\| y - f(\theta) \right\|^2 \\
&\approx \left\| y - f\left(\theta^*\right) - F.(\theta - \theta^*) \right\|^2 \\
&= \left\| \varepsilon - F.\left(\theta - \theta^*\right) \right\|^2
\end{aligned}
$$

Since, $y - f\left(\theta^*\right) = \varepsilon$. And the estimated sum squares is take the form as,

$$
\begin{aligned}
(n-p)s^2 &= S\left(\hat{\theta}\right) \\
&= \left\| y - f\left(\hat{\theta}\right) \right\|^2 \\
&\approx \left\| \left(I_n - P_F\right)\varepsilon \right\|^2 \\
&= \varepsilon'\left(I_n - P_F\right)\varepsilon .
\end{aligned}
$$

Thus we get,

$$
\begin{aligned}
S\left(\theta^*\right) - S\left(\hat{\theta}\right) &= \left\| y - f\left(\theta^*\right) \right\|^2 - \left\| y - f\left(\hat{\theta}\right) \right\|^2 \\
&\approx \left\| \varepsilon \right\|^2 - \left\| \left(I_n - P_F\right)\varepsilon \right\|^2 \\
&= \varepsilon'\varepsilon - \varepsilon'\left(I_n - P_F\right)\varepsilon ,
\end{aligned}
$$

where, $P_F = F.\left(F.'F.\right)^{-1} F.'$ and $I_n - P_F$ are symmetric and idempotent.

$$
\begin{aligned}
S\left(\theta^*\right) - S\left(\hat{\theta}\right) &= \varepsilon' P_F \varepsilon \\
&\approx \left(\hat{\theta} - \theta^*\right)' F.'F.\left(\hat{\theta} - \theta^*\right).
\end{aligned}
$$

3.2.3 Numerical methods

In the non-linear estimation problems it is most convenient to write down the normal equations (3.5) and develop an iterative technique for solving them. Estimating the parameters of such statistical model requires optimizing some kind of objective function, i.e., least squares estimates are obtained by minimizing a sum of squares and maximum likelihood estimation is done by maximizing the likelihood function. Most of the minimization methods are iterative. Some of these methods are discuss in the following section.

3.2.3.1 Gauss-Newton algorithm

This method is also known as GAUSS method. The Gauss-Newton algorithm is a method for finding least squares estimates in non-linear problems. The algorithm consists of obtaining a sequence of linear least squares estimates that converge to the least squares estimate in the non-linear problem. This procedure requires as initial guess (estimate) for θ, say $\theta^{(1)}$ and defines a series of estimates $\theta^{(r)}$ that converge to the least squares estimate $\hat{\theta}$.

Suppose $\theta^{(a)}$ is an approximation to the least square estimate $\hat{\theta}$ of a non-linear model. For θ close to $\theta^{(a)}$, we use a linear Taylor expansion

$$f(\theta) \approx f\left(\theta^{(a)}\right) + F.^{(a)}\left(\theta - \theta^{(a)}\right) \tag{3.7}$$

Where $F.^{(a)} = F.\left(\theta^{(a)}\right)$. Appling this to the residual vector $r(\theta)$, we have

$$\begin{aligned} r(\theta) &= y - f(\theta) \\ &\approx r\left(\theta^{(a)}\right) - F.^{(a)}\left(\theta - \theta^{(a)}\right). \end{aligned}$$

Substituting in $S(\theta) = r'(\theta) r(\theta)$ leads to

$$S(\theta) \approx r'\left(\theta^{(a)}\right) r\left(\theta^{(a)}\right) - 2r'\left(\theta^{(a)}\right) F.^{(a)}\left(\theta - \theta^{(a)}\right) + \left(\theta - \theta^{(a)}\right)' F.^{(a)\prime} F.^{(a)}\left(\theta - \theta^{(a)}\right) \tag{3.8}$$

The right hand side is minimized with respect to θ when

$$\begin{aligned} \theta - \theta^{(a)} &= \left(F.^{(a)\prime} F.^{(a)}\right)^{-1} F.^{(a)\prime} r\left(\theta^{(a)}\right) \\ &= \delta^{(a)}, \text{ (say)}. \end{aligned}$$

(3.9)

This suggest that, given a current approximation $\theta^{(a)}$, the next approximation should be

$$\theta^{(a+1)} = \theta^{(a)} + \delta^{(a)} \tag{3.10}$$

This provides an iterative scheme for obtaining $\hat{\theta}$. The approximation of $S(\theta)$ by the quadratic (2.4), and the resulting updating formulae (2.5) and (2.6), are usually referred to as the Gauss-Newton method. It forms the basis of a number of least squares algorithms. The Gauss-Newton algorithm is convergent, i.e., $\theta^{(a)} \to \hat{\theta}$ as $a \to \infty$, provided that $\theta^{(1)}$ is close to enough to θ^{*} (true value of θ) and n is large enough.

3.2.3.2 Newton-Raphson method

Newton-Raphson is one of the popular Gradient methods of estimation. In Newton-Raphson method we find the values of β_j that maximize a twice differentiable concave function, the objective function $g(\beta)$. In this method we approximate $g(\beta)$ at β^t by Taylor series expansion up to the quadratic terms

$$g(\beta) \approx g(\beta^t) + G(\beta^t)(\beta - \beta^t) + \frac{1}{2}(\beta - \beta^t)' H(\beta^t)(\beta - \beta^t)$$

Where, $G(\beta^t) = \left[\frac{\partial g}{\partial \beta_i}\right]_{\beta^t}$ is the gradient vector and $H(\beta^t) = \left[\frac{\partial^2 g}{\partial \beta_i \partial \beta_k}\right]_{\beta^t}$ is the Hessian matrix. This Hessian matrix is positive definite, the maximum of the approximation $g(\beta)$ occurs when its derivative is zero.

$$G(\beta^t) + H(\beta^t)(\beta - \beta^t) = 0$$

$$\beta = \beta^t - \left[H(\beta^t)\right]^{-1} G(\beta^t) \tag{3.11}$$

This gives us a way to compute β^{t+1}, the next value in iterations is,

$$\beta^{t+1} = \beta^t - \left[H(\beta^t)\right]^{-1} G(\beta^t)$$

The iteration procedures continue until convergence is achieved. Near the maximum the rate of convergence is quadratic as define by

$$\left|\beta_i^{t+1} - \hat{\beta}_i\right| \le c \left|\beta_i^t - \hat{\beta}_i\right|^2$$

for some $c \geq 0$ when β_i^t is near $\hat{\beta}_i$ for all i. Thus we get estimates $\hat{\beta}_i$ by Newton-Raphson methods.

3.2.3.3 Steepest Descent method

Steepest descent is well known Gradient method. In the Gradient method the direction vector δ is usually chosen to be,

$$\delta = -P_n \gamma_n$$

Where, γ_n is the gradient vector of the objective function (3.4). In the steepest descent initially we obtained by choosing,

$$P_n = I_k \tag{3.12}$$

in all iteration, where k is the dimension of the parameter space.

Although this method is very simple, its use cannot be recommended in most cases, since it may converge very slowly if the minimum is in a long and narrow valley, that is, if the objective function is ill-conditioned. It is clear that using the same direction matrix P_n in each iteration does not allow a flexible adjustment to different shapes of the objective function surface. However, the steepest descent method can be valuable if it combined with Marquardt algorithm.

3.2.3.4 Marquardt method

The Marquardt method developed by Marquardt in 1963, sometimes referred to as the Marquardt-Levenberg method, can be used to modify procedures that do not guarantee a positive definite direction matrix P_n.

This utilizes the fact that

$$P_n + \lambda_n \overline{P}_n \tag{3.13}$$

is always positive definite if $\overline{P}_n$ is positive definite and the scalar λ_n is sufficiently large. A possible choice for $\overline{P}_n$ is the identity matrix. Typically this method is used in combination with the Gauss algorithm and $Z(\theta_n)' Z(\theta_n)$ is modified rather than its inverse. Thus the new direction matrix is given by

$$P_n = \left[Z(\theta_n)' Z(\theta_n) + \lambda_n \bar{P}_n \right]^{-1} \tag{3.14}$$

Where, I_k can be used as $\bar{P}_n$. For a λ_n close to zero, this method is equivalent to the Gauss algorithm.

Marquardt's method appears to perform very well in practice even if the initial parameters vector θ_1 is not close to the minimum of the objective function.

3.2.3.5 Quadratic Hill Climbing method

The Quadratic hill climbing method had developed by S. M. Goldfeld, R. E. Quandt and H.F. Trotter in 1966. In this method we modify the Hessian of the objective and use as the direction matrix, i.e.,

$$P_n = \left[H_n + \lambda_n I_k \right]^{-1} \tag{3.15}$$

This is a modification of Newton method, where the essence is to take that step at each iteration that maximizes a quadratic approximation to the function (3.4) on a sphere of suitable radius (see Goldfeld and Quandt (1973), pp.-5).

3.2.3.6 Rank one Correlation method

In this algorithms approximate the inverse of the Hessian of the objective function in each iteration by adding a correction matrix to the approximate inverse of the Hessian used, i.e.,

$$P_{n+1} = P_n + M_n \tag{3.16}$$

Where, M_n is the correction matrix and P_n is an approximation to H_n^{-1}. In the $(n+1)$st step P_{n+1} is used as the direction matrix. The first order approximation of the gradient,

$$\gamma_n \approx \gamma_{n+1} + H_{n+1}(\theta_n - \theta_{n+1})$$

which leads to

$$H_{n+1}^{-1}\left(\gamma_{n+1}-\gamma_{n}\right)\approx\left(\theta_{n+1}-\theta_{n}\right)$$

if the Hessian is non-singular. Replacing H_{n+1}^{-1} in this equation by (3.16), we get,

$$M_{n}\left(\gamma_{n+1}-\gamma_{n}\right)=\eta_{n} \tag{3.17}$$

Where, $\eta_{n}=\left(\theta_{n+1}-\theta_{n}\right)-P_{n}\left(\gamma_{n+1}-\gamma_{n}\right)$. Hence we get different correction matrices fulfilling (3.17), i.e.,

$$M_{n}=\frac{\eta_{nn}\eta_{n}^{\prime}}{\eta_{n}^{\prime}\left(\gamma_{n+1}-\gamma_{n}\right)}.$$

This is the only symmetric matrix of rank one which meets the requirements of (3.17), the resulting algorithms is therefore called the rank one correction method. This method developed by Broyden in 1965.

3.2.3.7 Davidon-Fletcher-Powell method

Davidon-Fletcher-Powell method is developed by Davidon in 1959, and Flecher and Powell in 1963. Where the correction matrix define by,

$$M_{n}=\frac{\zeta_{n}\zeta_{n}^{\prime}}{\zeta_{n}\left(\gamma_{n+1}-\gamma_{n}\right)}-\frac{P_{n}\left(\gamma_{n+1}-\gamma_{n}\right)\left(\gamma_{n+1}-\gamma_{n}\right)^{\prime}P_{n}}{\left(\gamma_{n+1}-\gamma_{n}\right)^{\prime}P_{n}\left(\gamma_{n+1}-\gamma_{n}\right)}$$

If the step length t_{n} for each step $\zeta_{n}=-t_{n}P_{n}\gamma_{n}$ is selected so as to minimize $H\left(\theta_{n}+\zeta_{n}\right)$ for the given θ_{n}, γ_{n} and P_{n}. Then $P_{n+1}=P_{n}+M_{n}$ will always be positive definite. Therefore, choosing P_{n} as the step direction in the nth iteration guarantees an acceptable step.

3.2.3.8 Method of scoring

This is a feature of Gauss algorithm, which can be used for maximum likelihood estimation. For this method the direction matrix is given by,

$$P_{n}=-\left[E\left.\frac{\partial^{2}\ln l}{\partial\theta\partial\theta}\right|_{\theta_{n}}\right]^{-1}$$

Where, l is the likelihood function. Thus, using the negative log likelihood function as the objective function, the Hessian is approximated by its expected value.

3.2.3.9 Brown-Dennis method

Brown and Dennis developed the Brown-Dennis method in 1971, suggesting the possibility of combining the GAUSS algorithm with a Quasi-Newton method. Instead of the inverse Hessian of the objective function, we approximate the Hessian of $f_1(\theta)$ iteratively; that is, we choose

$$F_{t,n+1} = F_{t,n} + M_{t,n} \tag{3.18}$$

Where, $F_{t,n}$ is an approximation to

$$\left.\frac{\delta^2 f_t}{\delta\theta\delta\theta'}\right|_{\theta_n} \tag{3.19}$$

From

$$\left.\frac{\delta f_t}{\delta\theta}\right|_{\theta_n} \approx \left.\frac{\delta f_t}{\delta\theta}\right|_{\theta_n+1} + \left[\left.\frac{\delta^2 f_t}{\delta\theta\delta\theta'}\right|_{\theta_n+1}\right](\theta_\text{n} - \theta_{n+1}) \tag{3.20}$$

It follows that $M_{t,n}$ should satisfy

$$M_{t,n}(\theta_{n+1} - \theta_n) = \mu_{t,n} \tag{3.21}$$

With

$$\mu_{t,n} = \left.\frac{\delta f_t}{\delta\theta}\right|_{\theta_{n+1}} - \left.\frac{\delta f_t}{\delta\theta}\right|_{\theta_n} + F_{t,n}(\theta_{n+1} - \theta_n) \tag{3.22}$$

For instance, a correlation matrix of rank one such as

$$M_{t,n} = \frac{\mu_{t,n}(\theta_{n+1} - \theta_n)'}{(\theta_{n+1} - \theta_n)'(\theta_{n+1} - \theta_n)} \tag{3.23}$$

could be used. The direction matrix for this algorithm is

$$P_n = \left\{ Z(\theta_n)' Z(\theta_n) - \sum_{t,t'=1}^{T} \left[y_t - f_t(\theta_n) \right] F_{t',n} \right\}^{-1} \quad (3.24)$$

3.2.3.10 Direct search method

When the first partial derivatives of the objective function do not exist or are different to compute then Direct search method play a effective role.

Starting from some initial $(k \times 1)$ parameter vector θ_1, a search is performed in k directions $\delta_1, \delta_2, ..., \delta_k$, which are at least linearly independent but often orthogonal. A topical iteration is

$$\theta_{n+1} = \theta_n + t_n \delta_j \quad (3.25)$$

The step length t_n is chosen such that $H(\theta_{n+1}) \leq H(\theta_n)$. The methods differ in how they select step length and direction and when they apply a new set of direction vectors.

Non-linear estimation procedure yields complicated computation then the linear estimation procedure. In most situation non-linear estimation problem can be solved by minimizing the error sum square estimation method using any of the optimization method (see Goldfeld and Quandt (1973), pp.-16).

3.3 Computational Issues

The estimation of non-linear model is not straight forward, in this analysis we introduce computational procedure for estimating such non-linear models. We use the GAUSS (See Aptech, 1995), the SAS system version 8.0, SPSS version 12.0 and MATLAB version 13.0.

Non-linear module of SPSS results in faster solution but most of the time it fails to converge. To save computational time we use SPSS. The most sensitive element in SPSS is the starting point, which is useful to converge the estimates. SPSS cannot switch the optimization algorithm and affected by starting points. So, we consider the GAUSS and SAS constrained optimization module for our computations.

The GAUSS constrained optimization module is used to find the optimal estimated value of the parameters in the non-linear model (2.8). The most critical element in GAUSS optimization is also the starting point. Accuracy of the results and the number of iteration required to converge depend on the starting values. When the optimization did not work, we tried different starting points as there seem to be no general methods for using starting points. As a result, we force (by pressing Alt+C) to get the estimated value and this value is used further as a starting point. The GAUSS optimization module is also affected by local maxima.

GAUSS consider four line search method in optimization procedure, i.,e., Steplength, STEPBT, HALF and BRENT method. Default method is STEPBT, which is the best method. The Steplength method will generate fast iteration but slower convergence and the BENT method will generate slow iteration but faster convergence. The optimization module also considers grid search procedure, where the line is split into some grid and the searching is done on the basis of each of these grids.

The GAUSS optimization module uses Newton-Raphson method as a default algorithm. But the switching is done can be manually or allowing more time. When the algorithm convergence then the optimum conditions are satisfied and exit from the iteration procedure. The iteration may require huge time for too large value of the data. In our computation, some situation arises that the GAUSS module does not convergence and the optimum conditions are not satisfy. The GAUSS optimization module also fails to switch different algorithm quickly, so we use the SAS system version 8.0 in our computations. In some cases, all of these softwares failed to converge. We left them unsolved as they require extensive study on estimation of strictly non-linear methods and require more time.

The SAS system switching optimization algorithm is vary quickly, even when it is fail to convergence the system can gives the estimated value. By default the SAS system consider Gauss-Newton optimization algorithm. Besides it the system considers Marquardt methods and other methods discuss in section (3.2.3.4). Whenever convergence problem is encounter then it improves by switching the algorithms.

Chapter 4

Analysis

4.1 Introduction

In this chapter we discuss the analysis procedure and compare the restricted and intrinsically nonlinear model with intrinsically linear model. By using different model selection criteria we select the best model for different manufacturing industries of Bangladesh.

The organization of the present chapter is as follows. Section 4.2 discusses the Cobb-Douglas production function. Section 4.3 consist name of different manufacturing industry in Bangladesh that we consider in this study. Section 4.4 discusses the results and comments. In section 4.5, we discuss the model selection criteria and based on them we select the best model. Section 4.6 discusses the hypothesis testing of the best model. Some concluding remarks are discussed in section 4.7.

4.2 Cobb-Douglas Production Function

The Cobb-Douglas production function is the widely used function in Econometrics. A famous case is the well-known Cobb-Douglas production function introduced by Charles W. Cobb and Paul H. Douglas (1928), although anticipated by Knut Wicksell (1901: p.128, 1923) and, some have argued, J. H. von Thünen (1863). They have estimated it after studying different industries in the world, for this it is used as a fairly universal law of production.

The Cobb-Douglas production function with multiplicative error term can be represented as,

$$P = AL^{\alpha}K^{\beta}U \tag{4.1}$$

where, P is the output; L is the Labor input; K is the Capital input; A is a constant; U is the random error term. α and β are positive parameters and $\alpha > 0$, $\beta > 0$, $L > 0$ and $\alpha + \beta = 1$.

The Cobb-Douglas production function with additive error term can be represented as,

$$P = AL^{\alpha} K^{\beta} + U \tag{4.2}$$

Where, P is the output; L is the Labor input; K is the Capital input; A is a constant; U is the random error term. α and β are positive parameters and $\alpha > 0$, $\beta > 0$, $L > 0$ and $\alpha + \beta = 1$.

4.3 Selected Manufacturing Industries of Bangladesh for this Study

We are going to apply the production functions on the basis of some selected manufacturing industries of Bangladesh. When we have gone to different industries to collect necessary data and information required for preparing the thesis, we did not get all the necessary information for obvious reasons. For example, many industries do not supply their man-hours and capital input data possibly to keep these information secret. Consequently a good number of industries show their production cost much higher than the gross output. So we get in many cases a biased and unrealistic scenario. For this reason any study related to production function, these primary data cannot be used.

In recent publications of "Statistical Yearbook of Bangladesh" and "Report on Bangladesh Census of Manufacturing Industries (CMI)" published by BBS, we get the published secondary data for the major manufacturing industries of Bangladesh. We have chosen the following manufacturing industries for the ongoing analysis.

1) Food manufacturing
2) Manufacturing of Textile
3) Manufacturing of Wearing apparel except footwear
4) Manufacturing of Leather & Leather products
5) Manufacturing of Leather footwear
6) Manufacturing of Wood & cork products
7) Manufacturing of Furniture & fixtures (wooden)
8) Manufacturing of Paper & paper products
9) Manufacturing of Printing & publications
10) Manufacturing of Drugs & pharmaceuticals

11) Manufacturing of Industrial chemical
12) Manufacturing of Other chemical products
13) Manufacturing of plastic products
14) Manufacturing of Pottery & China-ware
15) Manufacturing of Glass & glass products
16) Manufacturing of Non-metalic mineral products
17) Manufacturing of Iron & steel basic industries
18) Manufacturing of Fabricated metal products
19) Manufacturing of Non-electrical machinery
20) Manufacturing of Electrical machinery
21) Manufacturing of transport equipment

4.4 Results and Discussion

In case of Cobb-Douglas production function with multiplicative error terms i. e., for intrinsically linear model, we get the following estimates:

Table 4.1 The estimates of intrinsically linear Cobb-Douglas production function.

Industry name	Intercept	Calculated t value	Capital elasticity (α)	Calculated t value	Labor elasticity (β)	Calculated t value	Return to scale ($\alpha+\beta$)	$\gamma=\frac{1}{\alpha+\beta}$	R^2
Chemical	2.477	2.293	0.489	10.366	0.717	3.765	1.206	0.830	0.955
Drugs	7.761	2.138	0.580	2.294	-0.174	-0.385	0.406	2.463	0.325
Electrical machinery	6.877	6.879	0.326	1.949	0.499	1.757	0.825	1.212	0.848
Food	-1.550	-2.069	1.230	7.739	-0.074	-0.327	1.155	0.865	0.982
Furniture	0.925	0.385	0.698	2.788	0.713	1.917	1.411	0.708	0.654
Glass	5.449	3.815	0.387	4.786	0.432	1.368	0.819	1.221	0.830
Iron	-1.110	-0.508	0.007	0.056	2.154	5.257	2.161	0.463	0.804
Leather footwear	4.301	3.683	0.575	3.330	0.369	2.083	0.944	1.060	0.969
Leather	6.470	9.316	0.313	2.147	0.660	3.254	0.973	1.028	0.968
Metal	-0.256	-0.120	0.439	3.186	1.083	5.998	1.522	0.657	0.822
Mineral	3.287	3.578	0.810	9.838	0.052	0.540	0.862	1.159	0.922

Industry name	Intercept	Calculated t value	Capital elasticity (α)	Calculated t value	Labor elasticity (β)	Calculated t value	Return to scale ($\alpha+\beta$)	$\gamma = \frac{1}{\alpha+\beta}$	R^2
Non Electrical machinery	8.162	2.413	0.122	0.559	0.547	2.28	0.669	1.494	0.301
Other Chemical	1.879	1.083	0.717	10.544	0.473	1.776	1.190	0.840	0.942
Paper	6.367	1.512	0.310	0.565	0.573	0.973	0.883	1.132	0.552
Plastic	4.550	2.183	0.664	1.693	0.090	0.094	0.754	1.326	0.697
Pottery	2.526	2.464	0.542	3.360	0.577	2.245	1.119	0.894	0.901
Printing	-1.188	-0.715	-0.025	-0.038	2.016	2.189	1.991	0.502	0.962
Textile	6.766	2.050	0.478	4.214	0.236	1.228	0.714	1.401	0.577
Transport	5.142	4.783	0.250	1.806	0.863	2.417	1.113	0.899	0.929
Wearing	3.572	3.244	0.236	0.874	0.926	3.373	1.162	0.861	0.990
Wood	2.016	1.364	0.728	5.705	0.317	5.120	1.045	0.957	0.900

The estimates are obtained by applying Ordinary Least Square (OLS) method. The 4th column represents the estimates of Capital elasticity; the 5th column represents the calculated value of t statistic; the 6th and 7th column refers the Labor elasticity and its calculated value of t statistic respectively. The 8th and 9th column of the table represents the returns of scale and the economy of an industry respectively. The last column of the table represents the value of R^2.

There are economies of scale in the manufacturing of Iron & steel basic industries, Printing & publication, Fabricated metal products, Furniture & fixtures (wooden), Industrial chemicals, Other chemical products, Wearing apparel except footwear, Food, Pottery & China-ware, Transport equipment and Wood & cork products since $\gamma < 1$. There are diseconomies of scale in the Leather footwear, Leather & leather products, Paper & paper products, Non-metalic mineral products, Electrical machinery, Glass & glass products, Plastic products, Textile, Non-electrical machinery and Drug & pharmaceutical industries.

Table 4.2 The estimates of Cobb-Douglas production function with additive error term (intrinsically nonlinear) and without any restriction on parameters of the industries under study.

Industry name	Intercept	Standard error	Capital elasticity (α)	Standard error	Labor elasticity (β)	Standard error	Returns to scale ($\alpha+\beta$)	R^2	$\gamma=\frac{1}{\alpha+\beta}$
Chemical	9.6485	12.3953	0.6143	0.0767	0.4524	0.1604	1.0667	0.949	0.937
Drugs	266409.3	1229527	0.3182	0.4522	-0.1119	0.8862	0.2063	0.083	4.847
Electrical machinery	2949.46	3249.9	-0.1346	0.2121	1.1920	0.3820	1.0574	0.873	0.946
Food	0.3123	0.4473	1.402	0.1306	-0.394	0.1372	1.008	0.972	0.992
Furniture	0.0457	0.1161	1.396	0.2620	0.106	0.3074	1.502	0.914	0.666
Glass	580.129	554.006	0.4204	0.1157	0.2198	0.2479	0.6402	0.85	1.562
Iron	0.0245	0.0697	0.151	0.1464	2.204	0.2580	2.355	0.874	0.425
Leather footwear	0.0085	0.0098	1.224	0.0733	0.269	0.0334	1.493	0.995	0.670
Leather	780.0625	611.9324	0.286	0.1044	0.684	0.1777	0.97	0.96	1.031
Metal	0.4375	1.1731	0.602	0.1870	0.88	0.2352	1.482	0.795	0.675
Mineral	75.1416	193.1749	0.809	0.17926	-0.067	0.0738	0.742	0.8	1.348
Non electrical machinery	3668.83	9771.5	-0.416	0.0992	1.604	0.2955	1.188	0.77	0.842
Other chemical	4.2748	8.1292	1.073	0.1323	-0.097	0.4043	0.976	0.965	1.025
Paper	1.0666	7.7612	0.802	0.8280	0.461	0.7636	1.263	0.611	0.792
Plastic	22444.34	39583.7	0.491	0.2749	-0.408	0.6431	0.083	0.527	12.048
Pottery	22.9491	44.6271	0.899	0.2729	-0.18	0.2763	0.719	0.883	1.391
Printing	23.8531	63.5103	-0.326	1.0426	2.001	1.5326	1.675	0.863	0.597
Textile	16730.75	54810.3	0.378	0.1235	0.132	0.1587	0.51	0.499	1.961
Transport	65.6149	135.1517	0.316	0.1343	0.86	0.2709	1.176	0.892	0.850

Industry name	Intercept	Standard error	Capital elasticity (α)	Standard error	Labor elasticity (β)	Standard error	Returns to scale ($\alpha+\beta$)	R^2	$\gamma=\frac{1}{\alpha+\beta}$
Wearing	1065.44	1572.4	1.043	0.4667	-0.449	0.5655	0.594	0.932	1.684
Wood	2320.71	4925.3	0.315	0.1736	0.275	0.0690	0.590	0.789	1.695

There are economies of scale in the manufacturing of Chemical, Electrical machinery, Food, Furniture & fixtures (wooden), Iron & steel basic, Leather footwear, Fabricated metal products, Non electrical machinery, Paper & paper products, Printing & publications, Transport equipment since $\gamma < 1$.

There are diseconomies of scale in the Drugs & pharmaceuticals, Glass & glass products, Leather & leather products, Non-metalic mineral products, Other chemical products, Plastic products, Pottery & China-ware, Textile, Wearing apparel except footwear, Wood & crock products industries.

In Cobb-Douglas production function, we know that, the Capital and Labor elasticity are positive, so we use the nonlinear estimation procedures. Under this restriction we get the following estimates of the manufacturing industries under study.

Table 4.3 The estimates of Cobb-Douglas production function with additive error term (intrinsically nonlinear) and putting restriction on parameters of the industries under study.

Name of Industry	Intercept	Capital elasticity (α)	Labor elasticity (β)	Returns to scale ($\alpha+\beta$)	$\gamma=\frac{1}{\alpha+\beta}$
Chemical	9.64825	0.61429	0.45239	1.06668	0.93749
Drugs	1.15421	1.06239	0	1.06239	0.94127
Electrical machinery	2.85917	0	1.70804	1.70804	0.58547
Food	0.01789	1.33397	0	1.33397	0.74964
Furniture	0.04564	1.39531	0.10634	1.50165	0.66593
Glass	580.1568	0.42038	0.21982	0.6402	1.56201

Name of Industry	Intercept	Capital elasticity (α)	Labor elasticity (β)	Returns to scale ($\alpha+\beta$)	$\gamma=\frac{1}{\alpha+\beta}$
Iron	0.02452	0.15062	2.20455	2.35517	0.42460
Leather footwear	0.00851	1.22373	0.26919	1.49292	0.66983
Leather	780.16	0.28655	0.68451	0.97106	1.02980
Metal	0.43754	0.60164	0.88004	1.48168	0.67491
Mineral	98.91466	0.75564	0	0.75564	1.32338
Non-electrical machinery	3.27594	0	1.78538	1.78538	0.56010
Other chemical	2.92618	1.04445	0	1.04445	0.95744
Paper	1.06585	0.80162	0.46125	1.26287	0.79185
Plastic	13084.06	0.32799	0	0.32799	3.04887
Pottery	60.7963	0.73152	0	0.73152	1.36702
Printing	8.87988	0	1.58091	1.58091	0.63255
Textile	16743.19	0.37752	0.13172	0.50924	1.96371
Transport	65.66222	0.31633	0.85981	1.17614	0.85024
Wearing	1749.324	0.67873	0	0.67873	1.47334
Wood	2321.932	0.31543	0.27466	0.59009	1.69466

There are economies of scale in the manufacturing of Chemical, Drugs & pharmaceuticals, Electrical machinery, Food, Furniture & fixtures (wooden), Iron & steel basic, Leather footwear, Fabricated metal products, Non-electrical machinery, Other chemical products, Paper & paper products, Printing & publications, Transport equipment since $\gamma < 1$. There are diseconomies of scale in the Glass & glass products, Leather & leather products, Non-metalic mineral products, Pottery & China-ware, Textile, Wearing apparel except footwear, Wood & crock products industries.

4.4.1 Interpretations of the Cobb-Douglas estimates

Holding capital input constant we see that 1% increase in the labor input led on an average about 0.717%, 0.452% and 0.452% increase in the output of "Chemical industry" for the model with intrinsically linear, intrinsically nonlinear and intrinsically nonlinear with restricted parameters. We get increased output of the selected industries in the following table for three different models.

Table 4.4 Rate of change of output for a given change in labor input of the industries under study.

Serial number	1% increase in labor input for	% increased in output for		
		Intrinsically linear model	Intrinsically nonlinear model	Intrinsically nonlinear with restricted parameters model
1	Chemical	0.717	0.452	0.452
2	Drugs	-0.174	-0.112	0.000
3	Electrical machinery	0.499	1.192	1.708
4	Food	-0.074	-0.394	0.000
5	Furniture	0.713	0.106	0.106
6	Glass	0.432	0.220	0.220
7	Iron	2.154	2.204	2.205
8	Leather	0.369	0.684	0.269
9	Leather footwear	0.660	0.269	0.685
10	Metal	1.083	0.880	0.880
11	Mineral	0.052	-0.067	0.000
12	Non electrical machinery	0.547	1.604	1.785
13	Other chemical	0.473	-0.097	0.000
14	Paper	0.573	0.461	0.461
15	Plastic	0.090	-0.408	0.000
16	Pottery	0.577	-0.180	0.000
17	Printing	2.016	2.001	1.581
18	Textile	0.236	0.132	0.132
19	Transport	0.863	0.860	0.860
20	Wearing	0.926	-0.449	0.000

Serial number	1% increase in labor input for	% increased in output for		
		Intrinsically linear model	Intrinsically nonlinear model	Intrinsically nonlinear with restricted parameters model
21	Wood	0.317	0.275	0.275

We may comment here that Iron & steel basic industries, Printing & publication and Fabricated metal products follows the forms of labor intensive production function when we consider the intrinsically linear model. Also we observed that Electrical machinery industries, Iron & steel basic industries, Non electrical machinery industries and Printing & publication follows the forms of labor intensive production function for intrinsically nonlinear and intrinsically nonlinear with restricted parameters model.

From the above table, we observed that, if we do not put any restriction on the parameters, we get negative labor elasticity of substitution for some industries. But we know that labor elasticity of substitution is positive. Thus, we put the restriction on the parameters to get the results.

Again holding labor input constant we see that 1% increase in the capital input led on an average about 0.4889%, 0.6143% and 0.6143% increase in the output of "Chemical industry" for the model with intrinsically linear, intrinsically nonlinear and intrinsically nonlinear with restricted parameters. We get increased output of the selected industries in the following table for three different models.

Table 4.5 Rate of change of output for a given change in capital input of the industries under study.

Serial number	1% increase in capital input for	% increased in output for		
		Intrinsically linear model	Intrinsically nonlinear model	Intrinsically nonlinear with restricted parameters model
1	Chemical	0.4889	0.6143	0.6143
2	Drugs	0.4400	0.3182	1.0624
3	Electrical machinery	0.3257	-0.1346	0.0000
4	Food	1.2297	1.4020	1.3340
5	Furniture	0.6982	1.3960	1.3953

Serial number	1% increase in capital input for	% increased in output for		
		Intrinsically linear model	Intrinsically nonlinear model	Intrinsically nonlinear with restricted parameters model
6	Glass	0.3867	0.4204	0.4204
7	Iron	0.0072	0.1510	0.1506
8	Leather	0.5749	0.2860	1.2237
9	Leather footwear	0.3134	1.2240	0.2866
10	Metal	0.4393	0.6020	0.6016
11	Mineral	0.8101	0.8090	0.7556
12	Non electrical machinery	0.1218	-0.4160	0.0000
13	Other chemical	0.7172	1.0730	1.0445
14	Paper	0.3102	0.8020	0.8016
15	Plastic	0.6644	0.4910	0.3280
16	Pottery	0.5415	0.8990	0.7315
17	Printing	-0.0248	-0.3260	0.0000
18	Textile	0.4780	0.3780	0.3775
19	Transport	0.2496	0.3160	0.3163
20	Wearing	0.2358	1.0430	0.6787
21	Wood	0.7281	0.3150	0.3154

We may comment here that only Food manufacturing industries follows the forms of capital intensive production function when we consider the intrinsically linear model. Also we observed that Food, Furniture & fixtures (wooden), Leather footwear, Other chemical products, Wearing apparel except footwear industries follows the forms of capital intensive production function for intrinsically nonlinear model. But Drugs &, Food, Furniture, Leather & leather products, Other chemical products manufacturing industries follows the forms of capital intensive production function for intrinsically nonlinear with restricted parameters model.

From the above table, we observed that, if we do not put any restriction on the parameters, we get negative capital elasticity of substitution for some industries. But we know that capital elasticity of substitution is positive. Thus, we put the restriction on the parameters to get the results.

Rate of change of output for a given change in labor input for the major industries of Bangladesh can be displayed as:

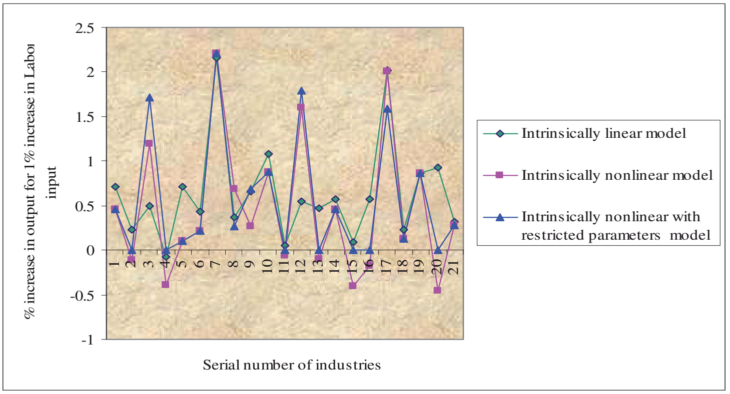

***Figure 4.1** Graph of the Rate of change of output for a given change in labor input for the major industries of Bangladesh.*

From the above figure, we observed that, Iron & steel basic industry provides highest percent increase in output for 1% increase in labor input for three models.

Rate of change of output for a given change in capital input for the major industries of Bangladesh can be displayed as:

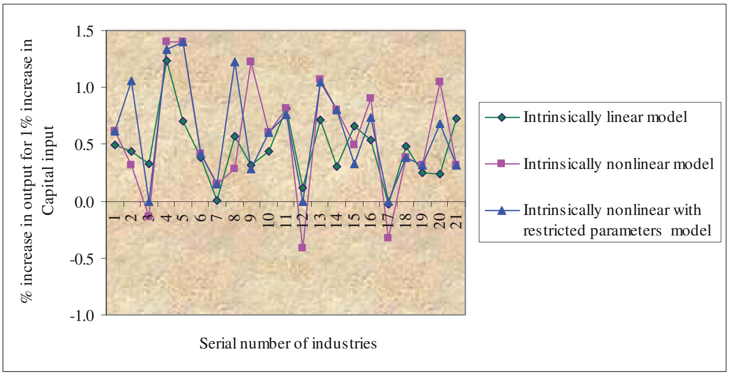

***Figure 4.2** Graph of the Rate of change of output for a given change in capital input for the major industries of Bangladesh.*

From the above figure, we observed that, Food manufacturing industry provides highest percent increase in output for 1% increase in capital input for three models.

4.4.1.1 Returns to scale

Returns to scale refer to the response of output to a proportionate change inputs. It can be increasing or decreasing or it may even remain constant. These three typical cases are described bellow:

- ***Increasing returns to scale:*** If doubling the input of a system results in more than double output, then the system is said to have an increasing returns to scale. In case of Cobb-Douglas production function, we have increasing returns to scale when $\alpha + \beta > 1$.
- ***Decreasing returns to scale:*** If doubling the input of a system results in less than double output, then the system is said to have a decreasing returns to scale. In case of Cobb-Douglas production function, we have increasing returns to scale when $\alpha + \beta < 1$.
- ***Constant returns to scale:*** If doubling the input of a system results in exactly double output, then the system is said to have a decreasing returns to scale. In case of Cobb-Douglas production function, we have increasing returns to scale when $\alpha + \beta = 1$.

Table 4.6 Industries having increasing returns to scale of the three models under study.

Intrinsically linear model	Intrinsically nonlinear model	Intrinsically nonlinear with restricted parameters model
Chemical	Chemical	Chemical
Furniture	Furniture	Furniture
Iron	Iron	Iron
Metal	Metal	Metal
Printing	Printing	Printing
Transport	Transport	Transport
Food	Food	Food
Other chemical		Other chemical
Pottery		
Wearing		

Wood		
	Leather footwear	Leather footwear
	Electrical machinery	Electrical machinery
	Non electrical machinery	Non electrical machinery
	Paper	Paper
		Drug

Table 4.7 Industries having decreasing returns to scale of the three model under study.

Intrinsically linear model	Intrinsically nonlinear model	Intrinsically nonlinear with restricted parameters model
Drugs	Drugs	
Glass	Glass	Glass
Plastic	Plastic	Plastic
Textile	Textile	Textile
Leather	Leather	Leather
Mineral	Mineral	Mineral
Electrical machinery		
Leather footwear		
	Other Chemical	
Non electrical machinery		
	Pottery	Pottery
Paper		
	Wearing	Wearing
	Wood	Wood

Table 4.8 Increasing returns to scale for the industries under study.

Serial number	Name of industry	Increasing returns to scale ($\alpha+\beta>1$)		
		Intrinsically linear model	Intrinsically nonlinear model	Intrinsically nonlinear with restricted model
1	Chemical	1.2055	1.0667	1.0667
2	Drugs			1.0624
3	Electrical machinery		1.0574	1.7080
4	Food	1.1560	1.0080	1.3340
5	Furniture	1.4114	1.5020	1.5017
7	Iron	2.1614	2.3550	2.3552
8	Leather footwear			1.4929

9	Leather		1.4930	
10	Metal	1.5223	1.4820	1.4817
12	Non-electrical machinery		1.1880	1.7854
13	Other chemical	1.1901		1.0445
14	Paper		1.2630	1.2629
16	Pottery	1.1189		
17	Printing	1.9912	1.6750	1.5809
19	Transport	1.1126	1.1760	1.1761
20	Wearing	1.1618		
21	Wood	1.0451		

Table 4.9 Decreasing returns to scale for the industries under study.

Serial number	Name of industry	Decreasing returns to scale ($\alpha+\beta<1$)		
		Intrinsically linear model	Intrinsically nonlinear model	Intrinsically nonlinear and restricted model
2	Drug	0.6688	0.2063	
3	Electrical machinery	0.8251		
6	Glass	0.8188	0.6402	0.6402
8	Leather footwear	0.9435	0.9700	
9	Leather	0.9731		0.9711
11	Mineral	0.8625	0.7420	0.7556
12	Non-electrical machinery	0.6693		
13	Other chemical		0.9760	
14	Paper	0.8834		
15	Plastic	0.7539	0.0830	0.3280
16	Pottery		0.7190	0.7315
18	Textile	0.7138	0.5100	0.5092
19	Wearing		0.5940	0.6787
20	Wood		0.5900	0.5901

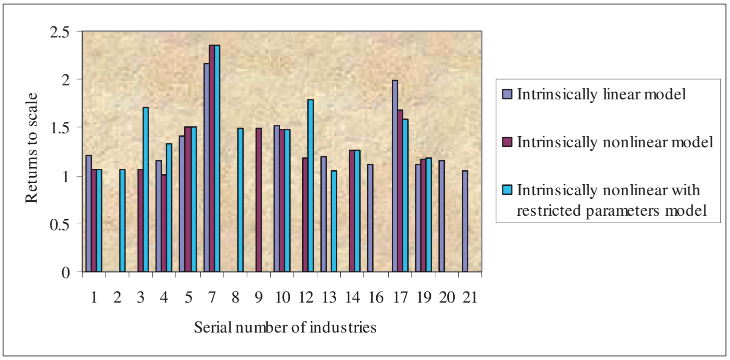

Figure 4.3 *Bar diagram for increasing returns to scale for the industries under study.*

From figure 4.3 we observed that Iron & steel basic industries has highest increasing returns to scale.

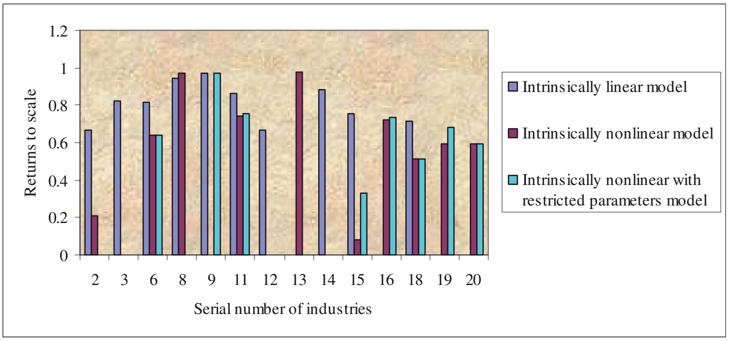

Figure 4.4 *Bar diagram for decreasing returns to scale for the industries under study.*

From the figure 4.4, we may observe that, Leather footwear, Leather & leather products and Other chemical products industries though not exactly but almost appear to follow the concept of constant returns scale.

From the above discussion we may conclude that three models give different outputs. So that, in order to forecast about the production of manufacturing industries, we need to identify the appropriate model. For this purpose, we select better model by model selection criteria.

4.5 Model Selection Criteria

To find the appropriate production function we use model selection criterion. The model that minimizes the criterion is the best model. The general form of model selection criteria is represented as,

$$Crit(m) = 2\ln(\max imized\ likelihood) + f(n,m)$$

Where, m denotes a model, n is the number of sample size and $f(n,m)$ is a function of n and number of independent parameters in the model m. The first term on the right hand side is a measure of fidality of the model to the data (or goodness of fit) and the second term is a “penalty function” which penalizes higher dimensional model.

In practice, one should use these criteria to identify a small group of best models. By increasing the number of the variables in a model the residual sum of squares $\sum \hat{u}_t^2$ will decrease and R^2 will increase, but at the cost of a loss in degrees of freedom.

In general, simpler models are recorded for two technical reasons. First, the inclusion of too many variables makes the relative precisions of individual coefficients worse. Second, the resulting loss of degrees of freedom would reduce the power of tests performed on the coefficients.

The recent years several criteria for choosing among models have proposed. These entire take the form of residual sum of squares (ESS) multiplied by a penalty factor that depend on the complexity of the model. Some of these criteria are discuss below.

4.5.1 Finite Prediction Error (FPE)

Akaike (1970) developed Finite Prediction Error procedure, which is known as FPE. The statistic of this procedure can be represented as,

$$FPE = \left(\frac{ESS}{T}\right)\frac{T+K}{T-K}$$

where, T is the number of observations and K is the number of estimated parameter (See Ramanathan, 1989, pp-165-167).

4.5.2 Akaike Information Criteria (AIC)

Akaike (1974) also developed another procedure which is known as Akaike Information Criteria. The form of this statistic is given below,

$$AIC = \left(\frac{ESS}{T}\right) e^{\left(\frac{2K}{T}\right)}$$

The value of AIC deceases when some variables are dropped (See Ramanathan, 1989, pp-165-167).

4.5.3 Hunnan and Quinn (HQ) criterion

Hunnan and Quinn (1979) developed a procedure which is known as HQ criteria. The statistic of this procedure can be represented as

$$HQ = \left(\frac{ESS}{T}\right) (\ln T)^{(2k/T)}$$

The value of HQ will decrease provided there are at least **16** observations (See Ramanathan, 1989, pp-165-167).

4.5.4 SCHWARZ criterion

Craven and Wahba (1978) developed a procedure which is known as $SCHWARZ$ (BIC) criteria. The form of this procedure is represented as

$$SCHWARZ = \left(\frac{ESS}{T}\right) T^{K/T}$$

The value of *SCHWARZ* will also decrease provided there are at least **8** observations (See Ramanathan, 1989, pp-165-167).

4.5.5 SHIBATA criterion

Craven and Wahba (1981) developed a procedure which is known as *SHIBATA* criteria.

The form of this procedure is represented as

$$SHIBATA = \left(\frac{ESS}{T}\right)\frac{T+2K}{T}.$$

When some variables dropped $SHIBATA$ will increase (See Ramanathan, 1989, pp-165-167).

4.5.6 Generalized Cross Validation (GCV)

Generalized Cross Validation (GCV) is another procedure which is developed by Craven and Wahba (1979). The form of the statistic is given below

$$GCV = \left(\frac{ESS}{T}\right)\left[1-\left(\frac{K}{T}\right)\right]^{-2}$$

If one or more variables are dropped then GCV will decrease (See Ramanathan, 1989, pp-165-167).

4.5.7 Rice criterion

The model selection criteria Rice developed by Craven and Wahba (1984). The form of this criterion can be represented as

$$RICE = \left(\frac{ESS}{T}\right)\left[1-\left(\frac{2K}{T}\right)\right]^{-1}.$$

(See Ramanathan, 1989, pp-165-167).

4.5.8 SGMASQ criterion

The form of this criterion can be represented as

$$SGMASQ = \left(\frac{ESS}{T}\right)\left[1-\left(\frac{K}{T}\right)\right]^{-1}.$$

If $SGMASQ$ decreases (that is $\bar{R}^2$ increases) when one or more variable dropped, then GCV and $RICE$ will also decreases (See Ramanathan, 1989, pp-165-167).

Among the above criteria, BIC and HQ are consistent in this sense that if the set of candidate models contains the true model, then these two criteria select the true model with probability 1 asymptotically. Shibata (1980) shows that AIC is asymptotically efficient, in the sense that it selects the model that is closest to the unknown true model asymptotically.

From Table 4.10 given in Appendix A we observe that, the Cobb-Douglas production function with additive error (2.8) performs better for the selected manufacturing industries based on the data under study period. Thus **the strictly nonlinear models (which are nonlinear with additive error terms) seem to be better than intrinsically linear model (which are nonlinear with multiplicative error terms)**.

4.6 Hypothesis Testing

To investigate the model that is the model is well fitted or not, we have consider the following the null hypothesis,

$H_0 : \theta = 0$, i.e., the model is not fitted well,

against the alternative hypothesis,

$H_0 : \theta \neq 0$, i.e., the model is fitted well,

where θ is the vector of parameters, i.e., $\theta = (A \quad \alpha \quad \beta)'$ for the model (2.8).

Under the null hypothesis, the test statistic is,

$$F = \frac{R^2/(k-1)}{(1-R^2)/(n-k)}$$

where, k is the number of parameter and n is the number of observations.

We reject H_0, if $F = \frac{R^2/(k-1)}{(1-R^2)/(n-k)} > F_{0.05,(k-1),(n-k)}$, which implies that model is fitted well.

The analytical results of the hypothesis testing are presented in the following table:

Table 4.11 The values of test statistic of intrinsically nonlinear model for selected manufacturing industries.

Name of Industry	R^2	F
Chemical	0.94929	121.6798
Drugs	0.85125	37.19748
Electrical machinery	0.87074	43.78624

Name of Industry	R^2	F
Food	0.95603	141.3281
Furniture	0.91378	68.88854
Glass	0.8499	36.80446
Iron	0.8743	45.21042
Leather footwear	0.99588	1571.17
Leather	0.96029	157.1867
Metal	0.79501	25.20886
Mineral	0.78839	24.21688
Non-electrical machinery	0.47428	5.863996
Other chemical	0.96451	176.6502
Paper	0.61092	10.20608
Plastic	0.5107	6.784284
Pottery	0.87927	47.33915
Printing	0.86192	40.57416
Textile	0.49885	6.470169
Transport	0.89209	53.73538
Wearing	0.92919	85.29494
Wood	0.78928	24.34662

From Table 4.9 we observe that R^2 is highly significant for all the manufacturing industries, we can say that the intrinsically nonlinear model (2.8) is fitted well according to the null hypothesis $H_0 : \theta = 0$.

In order to forecast the production of manufacturing industries, we use the following production function:

Table 4.12 Estimated intrinsically nonlinear Cobb-Douglas production functions for the manufacturing industries under study.

Serial number	Name of Industry	Estimated intrinsically nonlinear Cobb-Douglas production function
1	Chemical	$P = 9.64825K^{0.61429}L^{0.45239}$
2	Furniture	$P = 0.04564K^{1.39531}L^{0.10634}$
3	Glass	$P = 580.1568K^{0.42038}L^{0.21982}$
4	Iron	$P = 0.02452K^{0.15062}L^{2.20455}$
5	Leather footwear	$P = 0.00851K^{1.22373}L^{0.26919}$
6	Leather	$P = 780.16K^{0.28655}L^{0.68451}$
7	Metal	$P = 0.43754K^{0.60164}L^{0.88004}$
8	Paper	$P = 1.06585K^{0.80162}L^{0.46125}$
9	Textile	$P = 16743.19K^{0.37752}L^{0.13172}$
10	Transport	$P = 65.66222K^{0.31633}L^{0.85981}$
11	Wood	$P = 2321.932K^{0.31543}L^{0.27466}$

4.7 Concluding Remarks

From Table 4.10 given in Appendix A, we observe that, the Cobb-Douglas production function with additive error (2.8) performs better for the selected manufacturing industries based on the data under study period. Thus **the strictly nonlinear models (which are nonlinear with additive error terms) seem to be better than intrinsically linear model (which are nonlinear with multiplicative error terms)**. Also from Table 4.11, we observe that intrinsically nonlinear model fits well. For forecasting the output of the manufacturing industries we use the estimated production function given in Table 4.12. For the rest of the analysis, we use the strictly nonlinear model.

Chapter 5

Autocorrelation

5.1 Introduction

The purpose of this chapter is to survey the literature on some existing tests and other related issues about autocorrelation. When we are dealing with time series data, a number of special problems arise that often result in the violation of some of the important assumptions of a regression model. Autocorrelation is one of the most serious problems that arise due to such violation.

In the presence of autocorrelation, OLS estimates and forecasts based on them are still unbiased and consistent, but they are not BLUE and hence inefficient. Since, the estimated variances of the parameters become biased and inconsistent, so the test of hypothesis will not be valid and give us misleading conclusions in the presence of autocorrelation. For this reason, it is particularly important to describe the existing test and hence we remove the autocorrelation from the data and finally we re-estimate the parameters.

The organization of the present chapter is as follows. Section 5.2 discusses about the nature of autocorrelation. In section 5.3 we introduce the general forms of autocorrelation. Section 5.4 discusses the sources of autocorrelation. A discussion about the problems of autocorrelation is made in section 5.5 and section 5.6 represents some existing test for autocorrelation. In section 5.7 we discuss the results obtained form the data for different industries under study. Section 5.8 represents some transformations that are used to remove autocorrelation. Finally, section 5.9 contains some results of removing autocorrelation.

5.2 Nature of Autocorrelation

Autocorrelation is a special case of correlation; this is the situation, when the successive residuals tend to be highly correlated. Its existence implies that the total effect of a random error is not instantaneous, but is also felt in future periods, and, after some reflection. It is clear that this is a reasonable assumption for many economic relationships. So, it is a common phenomenon in most economic variables.

The method of least squares has several desirable properties, provided the error terms satisfy a number of assumptions. Among them one important assumption is that the successive values of the random disturbance term are independent that is, $E[u_i, u_j] = 0$, for $i \neq j$.

If this assumption is not satisfied and there exists such dependence that is $E[u_i, u_j] \neq 0$, for $i \neq j$, we say that there exists autocorrelation in the disturbance terms. For example, if we are predicting the growth of stock dividends, an overestimate in one year is likely to lead to overestimates in succeeding years, or, if we are dealing with quarterly time series data involving the regression of output on labor and capital inputs and if say there is a labor strike affecting output in one quarter. Hence, the disruption caused by a strike this quarter may very well affect output next quarter.

Autocorrelated values of the disturbance term may be observed for many reasons. Sometimes autocorrelation occurs due to omitted explanatory variables. It is because, if an autocorrelated variable is excluded from the set of explanatory variables, obviously its influence will be reflected in the random variable u, whose values will be autocorrelated. On the other hand, if we adopt a mathematical form which differs from the true form of the relationship, the u's may show autocorrelation. Autocorrelation may also occur due to interpolations in the statistical observation and for mis-specification of the true random disturbance term u.

The problem of autocorrelation is usually more common in time series data. In time series data, the observations are ordered in chronological order. Therefore, there are likely to be inter-correlations among successive observation especially if the time interval between successive observations is short, such as a day, a week or a month rather than a year.

There exists fourth-order autocorrelation in seasonal data. Although the incidence of autocorrelation is predominantly associated with time series data, it can occur in cross sectional data. The autocorrelation that occurs in cross sectional data is called spatial autocorrelation that is correlation in space rather than over time, (See Gujarati, 2003, pp.- 405).

5.3 General Form of Autocorrelation

There are many possible forms of autocorrelation, and each one lead to a different structure for the error covariance matrix, which are describe by different time series models. For forecasting purposes a model that describes the behaviors of a variable (or a set of variables) in terms of past values called time series model. Depending on the structure of the error u, time series models are different. To analysis the time series models, we can commonly group them to each of the following categories;

i. Standard time series models
ii. Financial time series models.

Standard time series models are those models where only the autocorrelation feature is present. An autoregressive (AR) process, A moving average (MA) process, An autoregressive moving average (ARMA) process and An autoregressive integrated moving average (ARIMA) process are the example of Standard time series models. When heteroscedasticity occur with autocorrelation in time series contexts then the models are known as Financial time series models. To forecast prices of financial assets, such as stock prices and exchange rates. These asset prices are characterized by the phenomenon known as volatility clustering, that is periods in which they exhibit wide swings for an extended time followed by a period of comparative tranquility. The ARCH (autoregressive conditional heteroscedasticity) and GARCH (generalized autoregressive conditional heteroscedasticity) models can capture such volatility clustering. We discuss about structures of these models in the following sections.

5.3.1 Autoregressive (AR) process

The most popular form of autocorrelation and one that has proved to be useful in many application is known as the first-order autoregressive disturbance or AR(1) process, is of the form,

$$u_t = \rho\, u_{t-1} + \varepsilon_t \tag{5.1}$$

where, $\varepsilon_t \sim N(0,1)$.

If we consider the model as,

$$u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \varepsilon_t \tag{5.2}$$

then we say that u_t follows second-order autoregressive AR(2) process.

In general,

$$u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \cdots + \rho_p u_{t-p} + \varepsilon_t \tag{5.3}$$

where, follows u_t an autoregressive process of order p or AR(p) process.

5.3.2 Moving average (MA) process

In some cases, only the particular error terms are correlated while all others have a zero correlation. This can be modeled by a moving average (MA) error process. The first order moving average or MA(1) process is,

$$u_t = \varepsilon_t + \theta \varepsilon_{t-1} \tag{5.4}$$

The second-order moving average MA(2) process is of the form,

$$u_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} \tag{5.5}$$

The general form of moving average or MA(q) process is,

$$u_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} \tag{5.6}$$

5.3.3 Autoregressive moving average (ARMA) process

The autoregressive moving average or ARMA(1,1) process is the case where,

$$u_t = \rho\, u_{t-1} + \varepsilon_t + \theta \varepsilon_{t-1} \tag{5.7}$$

More generally,

$$u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \cdots + \rho_p u_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} \tag{5.8}$$

is an ARMA(p,q) process, where p and q represents autoregressive and moving average parameters respectively.

5.3.4 Autoregressive integrated moving average (ARIMA) process

A very popular process in econometric time series is the autoregressive integrated moving average (ARIMA) process. Time series models are based on the assumption that it is stationary. But many of the econometric time series are nonstationary that is integrated. If a time series is integrated of order one i.e., I(1), its first differences are

$\mathrm{I}(0)$, i.e., stationary. Similarly, if a time series is $\mathrm{I}(2)$, its second difference is $\mathrm{I}(0)$. In general, if a time series is $\mathrm{I}(d)$, then after differencing it d times we get an $\mathrm{I}(0)$ series. Therefore, if we take difference a time series d times and then apply the $\mathrm{ARMA}(p,q)$ model to it, then the time series model is $\mathrm{ARIMA}(p,d,q)$ where, p is number of autoregressive terms, d the number of times the series has to difference and q the number of moving average terms. The ARIMA(1, 1, 1) process can be written as,

$$\Delta u_t = \rho\ \Delta u_{t-1} + \varepsilon_t + \theta\varepsilon_{t-1} \tag{5.9}$$

where, $\Delta u_t = u_t - u_{t-1}$ and $\Delta u_{t-1} = u_{t-1} - u_{t-2}$ are the first differences of u_t.

The ARIMA(2, 2, 2) process is of the form,

$$\Delta^2 u_t = \rho_1\ \Delta^2 u_{t-1} + \rho_2 \Delta^2 u_{t-2} + \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} \tag{5.10}$$

where, $\Delta^2 u_t = u_t - u_{t-1} - u_{t-2}$, $\Delta^2 u_{t-1} = u_{t-1} - u_{t-2} - u_{t-3}$ and $\Delta^2 u_{t-2} = u_{t-2} - u_{t-3} - u_{t-4}$ are second differences of u_t.

Similarly, $\mathrm{ARIMA}(p,d,q)$ is,

$$\Delta^d u_t = \rho_1\ \Delta^d u_{t-1} + \cdots + \rho_p \Delta^d u_{t-p} + \varepsilon_t + \theta_1\varepsilon_{t-1} + \cdots + \theta_q\varepsilon_{t-q} \tag{5.11}$$

where, Δ^d indicates the d–th difference of u_t.

5.3.5 Autoregressive Conditional Heterocedasticity (ARCH) process

The problem of autocorrelation is a feature of time series data and heteroscedasticity is a feature of cross-sectional data. When heteroscedasticity occur in time series contexts, then the current disturbance variance is depend on the previous disturbance information's. Hence the conditional disturbance variance is,

$$\mathrm{var}(u_t) = \sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \cdots + \alpha_p u_{t-p}^2 \tag{5.12}$$

thus arise an autoregressive conditional heteroscedasticity (ARCH) process, developed by Engle (1982).

The first-order autoregressive conditional heteroscedasticity or ARCH(1) process is,

$$u_t = \varepsilon_t (\alpha_0 + \alpha_1 u_{t-1}^2)^{\frac{1}{2}} \tag{5.13}$$

where, $\varepsilon_t \sim N(0,1)$, and the disturbance term is distributed as, $u_t \sim N\left[0, (\alpha_0 + \alpha_1 u_{t-1}^2)\right]$.

And the second-order autoregressive conditional heteroscedasticity process or $\text{ARCH}(2)$ is,

$$u_t = \varepsilon_t(\alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2)^{\frac{1}{2}} \tag{5.14}$$

Thus the autoregressive conditional heteroscedasticity process of order p or $\text{ARCH}(p)$ is,

$$u_t = \varepsilon_t(\alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \cdots + \alpha_p u_{t-p}^2)^{\frac{1}{2}} \tag{5.15}$$

5.3.6 Generalized Autoregressive Conditional Heterocedasticity (GARCH) process

A generalization of the ARCH model where the conditional variance of u_t is dependent not only on the past squared disturbance but also on past conditional variances is called the Generalize Autoregressive Conditional Heteroscedasticity (GARCH) process, introduced by Bollerslev (1986).

The first-order Generalized Autoregressive Conditional Heteroscedasticity process or $\text{GARCH}(1,1)$ is,

$$u_t = \varepsilon_t(\alpha_0 + \alpha_1 u_{t-1}^2 + \gamma_1 \sigma_{t-1}^2)^{\frac{1}{2}} \tag{5.16}$$

where, $u_t \sim N\left[0, (\alpha_0 + \alpha_1 u_{t-1}^2 + \gamma_1 \sigma_{t-1}^2)\right]$.

More generally,

$$u_t = \varepsilon_t(\alpha_0 + \alpha_1 u_{t-1}^2 + \cdots + \alpha_p u_{t-p}^2 + \gamma_1 \sigma_{t-1}^2 + \cdots + \gamma_q \sigma_{t-q}^2)^{\frac{1}{2}} \tag{5.17}$$

is the $\text{GARCH}(p,q)$ process. It expresses the conditional variance as a linear function of p lagged squared disturbances and q lagged conditional variances.

5.3.7 Autoregressive multiplicative seasonal process

A variable may be more closely related to its value in the same quarter (month, week, etc.) of the previous year than to its value in the immediately preceding quarter. Then we can introduce an autoregressive multiplicative seasonal $\text{AR}(1)\times\text{SAR}(1)$ process as,

$$
\begin{aligned}
z_t &= \phi z_{t-4} + u_t, \\
&= \phi z_{t-4} + \alpha u_{t-1} + \varepsilon_t, \\
&= \phi z_{t-4} + \alpha(z_{t-1} - \phi z_{t-5}) + \varepsilon_t \\
&= \alpha z_{t-1} + \phi z_{t-4} - \alpha\phi z_{t-5} + \varepsilon_t .
\end{aligned} \qquad (5.18)
$$

This is seen to be a special case of a general $\text{AR}(5)$, (See Johnston, 1997, pp.- 235).

5.3.8 Moving average multiplicative seasonal process

A moving average multiplicative seasonal $\text{MA}(1)\times\text{SMA}(1)$ process is define as,

$$
\begin{aligned}
z_t &= \beta z_{t-4} + \varepsilon_t, \\
&= \beta z_{t-4} + \theta\varepsilon_{t-1} + \varepsilon_t, \\
&= \beta z_{t-4} + \theta(z_{t-1} - \beta z_{t-5}) + \varepsilon_t \\
&= \theta z_{t-1} + \beta z_{t-4} - \theta\beta z_{t-5} + \varepsilon_t .
\end{aligned} \qquad (5.19)
$$

This is seen to be a special case of a general $\text{MA}(5)$, (See Johnston, 1997, pp.- 235).

Besides the above models, there are PAR (Periodic autoregressive) models, SPAR (Semi-periodic autoregressive) models etc. The models that are discussed above have a unique important place in the field of Economics and Business Statistics, since the series relating to prices, consumption and production of various commodities; money in circulation, bank deposits and bank clearings, sales and profits in a departmental store, agriculture and industrial production, national income and foreign exchange reserves, price and dividend of shares in a stock exchange market, etc.

5.4 Sources of Autocorrelation

Autocorrelation may arise from a good number of causes. Some of them are discussed as follows:

5.4.1 Omitted explanatory variables

If this occurs, since it is known that most economic variables are autocorrelated, then the error will be autocorrelated. Including the omitted variable into the equation should get rid of this problem.

5.4.2 Misspecification of the mathematical form of the model

For different economic problem we use different mathematical model as the problem required. But if we fail to adopt the correct relationship the random variable u may be autocorrelated. For instance, if we specify a linear form when the true form of the model is non-linear then the errors may reflect some dependence.

5.4.3 Interpolation in the statistical observations

This cause arises usually when we work with secondary or published, especially time series data. Because most of the published time series data involves some interpolation and smoothing process which averages out the true disturbances over successive periods. Consequently, the successive value of u is interrelated and exhibit autocorrelation pattern.

5.4.4 Misspecification of the true random error

It may well be expected in many cases for the successive values of u may be related due to purely random factors, such as wars, drought, change in taste, etc, which have an effect over successive periods. In such a case $E(u_t u_{t-j}) \neq 0 \quad \forall j \neq 0$ and the true pattern of the ε_t value will really be miss-specified. This may be called 'true autocorrelation' since the root cause of the auto-regressiveness lies in the nature of the error term.

5.5 Problems of Autocorrelation

Autocorrelation is the most serious problem, which arises in time series analysis. When the random disturbance term exhibits autocorrelation in the regression model, the use of ordinary least squares procedures has a number of important consequences. We summarize some of these problems in the following:

i) If we estimate the model in the presence of autocorrelation, the estimates of the parameters are no longer BLUE and will be inefficient. Hence, forecasts based on these estimates will be still unbiased but they will be also inefficient with larger variances.

ii) The variance of the error term may be seriously underestimated, that is the residual sum of squares, i.e., $\sum_i \hat{\varepsilon}_i^2$ may result in the estimate variances being much less than the true variance, since $s^2 = \frac{\sum_i \hat{\varepsilon}_i^2}{n-k}$.

iii) As a consequence of the problem with the residual sum of squares, the value of multiple correlation coefficients (R^2) becomes overestimate indicating better fit than actually exists and the estimated t–statistics in such a case will tend to appear more significant than they actually are. Therefore, the tests of hypothesis will not be valid in the presence of autocorrelation.

iv) When there exists positive autocorrelation in the random disturbance terms and the independent variable grows over time, the estimated standard errors will be smaller than the true ones and hence will seriously underestimate the variance, standard error, thereby inflating t values. Therefore, the usual t and F tests are not reliable with autocorrelation.

5.6 Tests for Autocorrelation

Autocorrelation is a serious problem in any econometric research and there are a good number of tests to detect the presence of autocorrelation in the data namely graphical method, the Runs test, the χ^2 test of independence of residuals, the Vonn Neumann ratio test, Durbin-Watson d test and so on. These test procedures are briefly described below:

5.6.1 Graphical method

In this method the regression residuals are plotted on a graph sheet either against their own lagged values or against time. This process gives a rough idea about the existence of autocorrelation in the disturbance term.

5.6.2 The Runs test

If the residuals are plot against time, then it could be seen that there are several residuals that are negative, then there is a series of positive residuals and finally there are several residuals that are again negative. If we assign a '$+$'sign for a positive residual and a '$-$' sign for a negative residual, we get several runs which is defined as an uninterrupted sequence of one symbols or attribute, such as $+$ or $-$, now let us assume

$N =$ Total number of observations $= (N_1 + N_2)$

$N_1 =$ Number of positive residuals or '$+$' symbol

$N_2 =$ Number of negative residuals or '$-$' symbol

$n =$ Number of runs

Then under the null hypothesis that the successive outcomes (the residuals) are independent of each other and assuming that $N_1 > 10$ and $N_2 > 10$, the number of run is asymptotically normally distributed with

Mean: $E(n) = \frac{2N_1N_2}{N} + 1$

Variance: $V(n) = \sigma_n^2 = \frac{2N_1N_2(2N_1N_2 - N)}{N^2(N-1)}$

If the hypothesis of randomness is sustainable, the number of runs obtained in a problem, is expected to lie between $\left[E(n) \pm 1.96\sigma_n\right]$ with 95 percent confidence. Therefore, the decision rule for this test be obtained as follows-

The null hypothesis of randomness is to be accepted with 95 percent confidence if

$$\left[E(n) - 1.96\sigma_n \leq n \leq E(n) + 1.96\sigma_n\right]$$

otherwise the null hypothesis is rejected.

5.6.3 Durbin-Watson d test

The Durbin-Watson statistic is a test statistic used to detect the presence of autocorrelation in the residuals from a regression analysis. This test was developed by James Durbin and Geoffrey Watson in 1951. It uses what is usually referred to as the Durbin-Watson d statistic and is based on the sum of the squared differences in successive values of the estimated disturbance terms.

If u_t is the residual associated with the observation at time t, then the test statistic is

$$d = \frac{\sum_{t=2}^{n}\left(\hat{u}_t - \hat{u}_{t-1}\right)^2}{\sum_{t=1}^{n}\hat{u}_t} \qquad (5.20)$$

5.6.3.1 Assumptions involved in d statistic

The Durbin-Watson d statistic involves the following assumptions:

- The regression model includes an intercept term. If such term is not present, as in the case of regression through the origin, it is essential to return the regression including the intercept term to obtain the Residual Sum of Squares (RSS).
- The explanatory variables are non-stochastic or fixed in repeated sampling.
- The disturbances u_t are generated by first order autoregressive scheme:

  $u_t = \rho u_{t-1} + \varepsilon_t$.
- The regression model does not include lagged values of the dependent variables as one of the explanatory variables.
- There is no missing observation in the data.
- The error term u_t is assumed to be normally distributed.

5.6.3.2 Decision taken in Durbin-Watson test

The probability distribution of the d statistic given in (5.20) is difficult to derive because, as Durbin and Watson have show, it depends in a complicated way on the X values present in a given sample. This should be understandable because d is computed from $\hat{u}_t$, which are of course, dependent on the given X's. Therefore unlike the t, F or χ^2 test, there is no unique critical value, which will lead to the rejection or acceptance of the null hypothesis that there is no first order serial correlation in the disturbances u_t. However, Durbin and Watson were successful in deriving a lower bound d_L and an upper bound d_U such that if the computed d these critical values, a decision can be made regarding the presence of positive or negative serial correlation. Moreover, these limilts depends only on the number of observations N and the number of explanatory variables and do not depend on the values taken by these explanatory variables.

The actual test procedure can be explained which shows that the limits of d are 0 and 4. This can be established as follows;

$$d = \frac{2\left(\sum_{t=1}^{n}\hat{u}_t^2 - \sum_{t=2}^{n}\hat{u}_t\hat{u}_{t-1}\right)}{\sum_{t=2}^{n}\hat{u}_t^2}$$

$$\approx 2(1-\hat{\rho})$$

where $\hat{\rho} = \frac{\sum_{t=2}^{n}\hat{u}_t\hat{u}_{t-1}}{\sum_{t=1}^{n}u_t^2}$, is the sample first order coefficient of correlation, an estimate of ρ.

Now from the above relation

$\hat{\rho} = 1$ suggests $d \approx 0$

$\hat{\rho} = 0$ suggests $d \approx 2$

$\hat{\rho} = -1$ suggests $d \approx 4$

From these functional expressions, we have two important conclusions. These are:

i) Values of d lies between 0 and 4, and

ii) If there is no autocorrelation i.,e., $\hat{\rho} = 0$ then $d = 2$.

Therefore, the calculated value of d turns out to be sufficiently close to 2, it can be assumed that there is no first order autocorrelation, either positive or negative. If $\hat{\rho} = +1$, indicating perfect positive correlation in the residuals, then $d \approx 0$. Therefore, the closer d is to 0, the greater is the evidence of positive serial correlation. If $\hat{\rho} = -1$, indicating perfect negative correlation in the residuals, then $d \approx 4$. Hence, the closer d is to 4, the greater is the evidence of negative serial correlation in the residuals.

Since the exact value of d can never be known, there exists a range of a values within which we cannot take any decision whether there exists positive or negative autocorrelation. Specifically, for the two-tailed Durbin-Watson test, we have well defined five regions for the value of d as shown bellow in a tabular form:

Table 5.1: **Durbin-Watson d test: Decision rules**

Null hypothesis	Decision	If
No positive autocorrelaiton	Reject	$0 < d < d_L$
No positive autocorrelaiton	No decision	$d_L < d < d_U$
No negative autocorrelaiton	Reject	$4 - d_L < d < 4$
No negative autocorrelaiton	No decision	$4 - d_U < d < 4 - d_L$
No autocorrelaiton positive or negative	Do not reject	$d_U < d < 4 - d_U$

5.6.3.3 Drawback of Durbin-Watson d test

We have mentioned that Durbin-Watson d test is widely used in empirical analysis. But this test has a great drawback. If it falls in the inconclusive region, one cannot conclude regarding the presence or absence of autocorrelation. To solve this problem, several authors have proposed modifications of the Durbin-Watson d test. This modification states that, in many situations, it has been found that the upper limit d_U is approximately the true signigicance limit and therefore in case the estimated d value lies in the inconclusive region, one can use the following modified d test procedure; given the level of significance α -

- $H_0 : \rho = 0$ vs $H_1 : \rho > 0$

Decision: If the estimated $d < d_U$, H_0 is to be rejected in favor of H_1, at level α, i.,e., there is statistically significant positive autocorrelation.

- $H_0 : \rho = 0$ vs $H_1 : \rho < 0$

Decision: If the estimated $4 - d < d_U$, H_0 is to be rejected in favor of H_1, at level α, i.,e., there is statistically significant evedence of negative autocorrelation.

- $H_0 : \rho = 0$ vs $H_1 : \rho \neq 0$

Decision: If the estimated $d < d_U$ or $4 - d < d_U$, H_0 is to be rejected in favor of H_1, at level 2α, i.,e., there is statistically significant evedence of autocorrelation, positive or negative.

This modification is just a rough and ready procedure, which should not be recommended in general, since the exact significance level of the test will almost certainly differ from the nominal 5% or 1% significance level being used. However, since the consequences of acceptance H_0 when autocorrelation is present are almost certainly serious than the consequences of incorrectly assuming it to be absent, one might have a preference for rejecting the null hypothesis in case of doubt (J. Johnston, 1984).

5.6.4 Durbin's h-test

If the model contains lagged regressand as a regressor, the DW d-test is not appropriate. For such models, Durbin (1970) utilizes the theory of LM test to develop the h-statistic, which is applicable in the presence of lagged dependent variable. This test is known as Durbin's h-test. Let us consider the model as,

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \cdots + \beta_k x_{kt} + \gamma y_{t-1} + u_t \qquad (5.21)$$

where, y_{t-1} is the lagged values of y_t, γ is the regression coefficient of one period lagged value of y_t. The test statistic of h-test is,

$$h = \left(1 - \tfrac{1}{2} d\right) \left[\frac{n}{(1 - n\hat{v}(\gamma))}\right]^{1/2} \qquad (5.22)$$

where, d is the DW statistic as defined in (5.20), n is the sample size and $\hat{v}(\gamma)$ is the estimate of variance of γ from least squares analysis. For large sample size, Durbin has shown that under the null hypothesis, the h-statistic (5.22) follows the standard normal distribution.

To apply the test, we proceed as follows: First, we estimate the model (5.21) by OLS and then $\hat{v}(\gamma)$ is estimated. Next we determine the value of d using the equation (5.20). Using the values of these estimates we compute the h-statistic using (5.22). Now if the sample size is reasonably large and if the computed $|h|$ exceeds 1.96, we can conclude that there is evidence of first-order autocorrelation in the model.

5.6.5 Detection of higher order autocorrelation: The Breusch-Godfrey (BG) test

To avoid some of the pitfalls of the Durbin-Watson d test of autocorrelation, statisticians Breusch and Godfrey have developed a test of autocorrelation that is general in the sense that it follows for

i) non-stochastic regressors, such as the lagged values of the regressand
ii) higher order autoregressive schemes, such as $AR(1)$, $AR(2)$, etc. and
iii) simple or higher order moving average of white noise error terms.

The Breusch-Godfrey (BG) test is also known as the Lagrange Multiplier test. Engle (1982) has shown that, for large sample, the sample size (n) is multiplied by the unadjusted R^2, which is obtained from the auxiliary regression has the chi-square distribution with degrees of freedom equal to the number of restrictions in the null hypothesis. Thus in this case, $LM = (n-p)R^2$, follows chi-square distribution with $(n-p)$ degrees of freedom, (See Ramanathan, 1995).

To illustrate this test let us write the model for t time periods as,

$$y_t = \beta_1 + \beta_2 x_t + u_t \tag{5.23}$$

with

$$u_t = \rho \, u_{t-1} + \varepsilon_t \tag{5.24}$$

where, $\varepsilon_t \sim NID(0, \sigma_\varepsilon^2)$

The steps for the LM test are described as follows:

Step 1 $H_0 : \rho_1 = \rho_2 = \cdots = \rho_p = 0$

H_a : at least one $\rho_i \neq 0$, $i = 1, 2, \cdots, p$.

Step 2 Estimate the model (2.91) by OLS and then obtain the estimated residuals from this model as, $\hat{u}_t = y_t - \hat{\beta}_1 - \hat{\beta}_2 x_t$.

Step 3 Regress $\hat{u}_t$ against all the independent variables (X) in equation (5.23) and $\hat{u}_{t-1}, \hat{u}_{t-2}, \cdots, \hat{u}_{t-p}$. We refer this regression as the auxiliary regression. The effective number of observations used in this auxiliary regression is $(n-p)$.

Step 4 Now, we compute the value of $(n-p)R^2$, where, R^2 is obtained from the auxiliary regression run in Step 2. If the calculated value of $(n-p)R^2$ exceeds the critical value of chi-square distribution with $(n-p)$ degrees of freedom then we reject the null hypothesis.

5.7 Test Results

In the present study, we have considered Durbin-Watson d test procedure to detect the presence of autocorrelation. Most of the cases d statistic fails to detect autocorrelation. The limits of d are obtained at 5% level of significance. Here for $k=2$, $n=16$, $d_L=0.982$ and $d_U=1.539$, where, k is the number of explanatory variables excluding the constant term and n is the total number of observations.

In case of inconclusive situation of d statistic, Run test is used to detect the presence or absence of autocorrelation where d statistic failed to detect. But in the present study we are not able to adopt the Run test because of the small number of observations. So to solve this problem, several authors have proposed modifications of the Durbin-Watson d test but they are rather involved and are beyond the scope of this text or study.

By using model selection criteria we observe that the Cobb-Douglas model with additive error term is better. Thus **the strictly nonlinear models** (which are non linear with additive error terms) **seem to be better than intrinsically linear model** (which are non linear with multiplicative error terms).

The computer program SPSS and SAS performs an exact d test (it gives the p value, the exact probability, of the computed d value) and those with access to this program may want to use that test in case the usual d statistic lies in the indecision zone. In many situations, however, it has been found that the upper limit d_U is approximately the true significance limit and therefore, in case the estimated d values lies in the indecision zone, one can use the modified d test procedure (D. N. Gujarati, 1995).

After applying the test we found, there exists positive autocorrelation of some manufacturing industries under the study.

Table 5.2 Result for testing autocorrelation for the Cobb-Douglas production function with additive error terms in different industries by applying Durbin-Watson test.

Name of industry	Durbin Watson (d)	$(4-d)$	Comment based on Durbin-Watson d test	Comment based on modified d test
Chemical	2.017	1.983	No autocorrelation	
Drug	1.187	2.813	No decision about no positive autocorrelation	Positive autocorrelation
Electrical machinery	1.702	2.298	No autocorrelation	
Food	1.988	2.012	No autocorrelation	
Furniture	1.235	2.765	No decision about positive autocorrelation	Positive autocorrelation
Glass	1.652	2.348	No autocorrelation	
Iron	1.021	2.979	No decision about positive autocorrelation	Positive autocorrelation
Leather footwear	0.591	3.409	Positive autocorrelation	
Leather	0.922	3.078	Positive autocorrelation	
Metal	1.401	2.599	No decision about positive autocorrelation	Positive autocorrelation
Mineral	1.643	2.357	No autocorrelation	
Non-electrical machinery	2.254	1.746	No autocorrelation	
Other chemical	0.681	3.319	Positive autocorrelation	
Paper	0.311	3.689	Positive autocorrelation	
Plastic	0.866	3.134	Positive autocorrelation	
Pottery	2.113	1.887	No autocorrelation	
Printing	1.704	2.296	No autocorrelation	

Name of industry	Durbin Watson (d)	$(4-d)$	Comment based on Durbin-Watson d test	Comment based on modified d test
Textile	1.21	2.79	No decision about positive autocorrelation	Positive autocorrelation
Transport	2.171	1.829	No autocorrelation	
Wearing	1.796	2.204	No autocorrelation	
Wood	0.902	3.098	Positive autocorrelation	

5.8 Treatment for Autocorrelation

5.8.1 Reformulation of model

If autocorrelation exists in the model, we have to take remedial action to correct it. Since the nature and cause of the autocorrelation are generally unknown, so there is no estimation procedure that can guarantee the elimination of autocorrelation. The solution here is reformulating the model to include the quadratic term and transforming the original model so that there is no autocorrelation.

5.8.1.1 Specify a more general dynamic structure

Let,

$$y_t = \alpha + \beta X_t + u_t \tag{5.25}$$

Where, y_t and X_t are respectively the values of dependent and explanatory variables at time t, α, β are the parameters of the model and u_t is the random disturbance term.

Let us consider the following model that relates the dependent variable to its own lagged value, an explanatory variable and its lag,

$$y_t = \beta_1 y_{t-1} + \beta_2 y_{t-2} + \beta_3 y_{t-3} + \varepsilon_t, |\beta_1| < 1 \tag{5.26}$$

Where, ε_t assumed to have mean zero and constant variance and be serially independent.

When variables are expressed as deviations from the mean the model (5.25) becomes,

$$y_t = \beta X_t + u_t \tag{5.27}$$

Solving for u_t in terms of the others and substituting for it in equation,

$$u_t = \rho u_{t-1} + \varepsilon_t, -1 < \rho < 1, \tag{5.28}$$

We get, $y_t - \beta X_t = \rho\left(y_{t-1} - \beta X_{t-1}\right) + \varepsilon_t$, which can be rearranged as follows,

$$y_t = \rho y_{t-1} + \beta X_t - \beta\rho X_{t-1} + \varepsilon_t \tag{5.29}$$

Comparing equations (5.26) and (5.27) we observe, $\beta_1 = \rho$, $\beta_2 = \beta$, $\beta_3 = -\rho\beta$ and the parameters satisfy the nonlinear restriction,

$$\beta_3 + \beta_1\beta_2 = 0 \tag{5.30}$$

If this restriction is satisfied by equation (5.26) then the model reduces to the static model as in equation (5.27) with autoregressive error structure in equation(5.28). After formulation the model (5.26) we test the non-linear restriction (5.30) on it and if it is accepted, the model is simplified along the lines of equations (5.27) and(5.28).

5.8.1.2 Model formulation of first differences

In empirical econometric work, a common way to get around the problem of spurious regression is to formulate models in terms of first difference, which is the difference between the value at time t and that at time $t-1$. Here we estimate, $\Delta y_t = \beta\Delta X_t + \varepsilon_t$. where, $\Delta y_t = y_t - y_{t-1}$ and $\Delta X_t = X_t - X_{t-1}$. The solution of using first differences might not always be appropriate and to seed whether it is appropriate or not, the first difference model can be rewritten as,

$$y_t = y_{t-1} + \beta X_t - \beta X_{t-1} + \varepsilon_t \tag{5.31}$$

Hence comparing this equation with (5.26), we see that this model is a special case with β_1 and $\beta_2 + \beta_3 = 0$. Now, we test these two restrictions first and if both are accepted, then a first difference specification is used. The first difference transformation is appropriate if the coefficient of autocorrelation is very high.

5.8.2 Estimation procedures

When the modified functional forms do not eliminate autocorrelation, several estimation procedures are available that will produce more efficient estimates than those obtained by the OLS procedure. These methods need to be applied only for time series data.

5.8.2.1 Generalized least squares (GLS) estimation

When ρ is a known parameter, the GLS estimate is best linear unbiased and it is given by,

$$\hat{\beta} = \left(X'\psi^{-1}X\right)^{-1} X'\psi^{-1}y \tag{5.32}$$

Also, we can calculate $\hat{\beta}$ using the following steps:

Step 1: Find a matrix P such that $P'P = \psi^{-1}$

Step 2: Calculate the transformed observation $y^* = Py$ and $X^* = PX$

Step 3: Apply least squares to the transformed model,

$$y^* = X^*\beta + e^* \tag{5.33}$$

Where, $e^* = Pe$, to obtain the generalized least square estimator,

$$\hat{\beta} = \left(X^{*\prime} X^*\right)^{-1} X^{*\prime} y^* \tag{5.34}$$

.

This estimator is identical to the one in (5.32). To find the transformed matrix P, we first need to specify the inverse of ψ using direct multiplication, it can be shown that,

$$\psi = \begin{bmatrix} 1 & -\rho & 0 & \cdots & 0 & 0 \\ -\rho & 1+\rho^2 & -\rho & \cdots & 0 & 0 \\ 0 & -\rho & 1+\rho^2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1+\rho^2 & -\rho \\ 0 & 0 & 0 & \cdots & -\rho & 1 \end{bmatrix} \tag{5.35}$$

Then it is possible to show that the appropriate transformation matrix is,

$$P = \begin{bmatrix} \sqrt{1-\rho^2} & 0 & 0 & \cdots & 0 & 0 \\ -\rho & 1 & 0 & \cdots & 0 & 0 \\ 0 & -\rho & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & -\rho & 1 \end{bmatrix} \tag{5.36}$$

Appling this matrix to step (ii) we get the transformed matrix y^* and X^*.

5.8.2.2 Cochrane-Orcutt iterative procedure

The Cochrane-Orcutt iterative procedure requires the transformation of the regression model,

$$y_t = \beta_1 + \beta_2 X_{t1} + ... + \beta k X_{tk} + u_t \tag{5.37}$$

to a form in which the OLS procedure is applicable. Now, rewriting equation (5.37) for the period $t-1$ we get,

$$y_{t-1} = \beta_1 + \beta_2 X_{(t-1)1} + ... + \beta_k X_{(t-k)k} + u_{t-1} \tag{5.38}$$

Multiplying (5.38) term by ρ and substituting from (5.37), we get,

$$y_t - \rho y_{t-1} = \beta_1(1-\rho) + \beta_2\left[X_{t2} - X_{(t-1)2}\right] + ... + \beta_k\left[X_{tk} - X_{(t-k)k}\right] + \varepsilon_t \tag{5.39}$$

where, $u_t = \rho u_{t-1} + \varepsilon_t$.

The equation (5.39) can be rewritten as follows,

$$y_t^* = \beta_1^* + \beta_2 X_{t2}^* + \beta_3 X_{t3}^* + ... + \beta_k X_{tk}^* + \varepsilon_t \qquad (5.40)$$

Where, $y_t^* = y_t - \rho y_{t-1}$, $\beta_1^* = \beta_1(1-\rho)$ and $X_{ti}^* = X_{ti} - \rho X_{(t-1)i}$ for $t = 2,3,...,T$ and $i = 2,3,...,k$.

The transformation that generates the variables y^* and the X^*'s is known as generalized differencing. Hence the error term in equation (5.40) satisfied all the properties needed for applying the least squares procedure. If ρ is known we can apply OLS method to (5.40) and the estimates that are obtained becomes BLUE.

5.8.2.3 Hildreth-Lu search procedure

A frequently used alternative to Cochrane-Orcutt procedure is the Hildreth-Lu search procedure in which we first choose a value of ρ (say ρ_1), then using this value we transform the variables and estimate the equation (5.40) by OLS.

From these estimates we derive $\hat{\varepsilon}_t$ from equation (5.40) and the error sum squares associated with it. Let it be $ESS_\varepsilon(\rho_1)$. Next we choose a different value of (ρ_2) and then estimate $ESS_\varepsilon(\rho_2)$ in the same way as for $ESS_\varepsilon(\rho_1)$.

By varying the value of ρ from -1 to $+1$ in some systematic way (say, at steps of length 0.05 or 0.01), we get a series of values of $ESS_\varepsilon(\rho)$. Then we choose that ρ for which ESS_ε is minimum. This is the final value of ρ that globally minimizes the error sum of squares of the transformed model. Then equation (5.40) is estimated with this final ρ as the optimum solution.

The above procedures can be applied to remove autocorrelation if we are able to detect autocorrelation properly, which are uses the estimated autocorrelation coefficient predefined by any other test such as Durbin-Watson test.

5.8.2.4 Theil-Nagar procedure

Theil and Nagar have suggested that in small samples ρ can be estimated as

$$\hat{\rho} = \frac{n^2\left(1-\frac{d}{2}\right)+k^2}{n^2-k^2} \tag{5.41}$$

where, $n =$ total number of observations, $d =$ Durbin-Watson d, and $k =$ number of coefficients (including the intercept) to be estimated.

5.9 Results

From Table 5.2 we observed that, the autocorrelation is present in Drug, Furniture, Iron, Leather footwear, Leather, Metal, Other chemical, Paper, Plastic, Textile and Wood industry for Cobb-Douglas model with additive error terms. In order to remove this autocorrelation at first we estimate the value of ρ by Theil-Nagar method since it is based on small sample size. Then we fit again the model for transformed data and we get the following estimates by SAS program.

Table 5.3 Result for testing autocorrelation for the Cobb-Douglas production function with additive error in different industries by applying Durbin-Watson test (transformed data).

Name of industry	$\hat{\rho}$	Durbin-Watson (d)	Comment
Drug	0.457749	2.043	No autocorrelation
Furniture	0.432874	2.260	No autocorrelation
Iron	0.543773	2.483	No autocorrelation
Leather footwear	0.766607	2.473	No autocorrelation
Leather	0.595077	2.496	No autocorrelation
Metal	0.34685	2.067	No autocorrelation
Other chemical	0.719968	1.896	No autocorrelation
Paper	0.911709	1.646	No autocorrelation
Plastic	0.624097	2.280	No autocorrelation
Textile	0.44583	1.701	No autocorrelation
Wood	0.605441	2.066	No autocorrelation

Form Table 5.3 we observed that the problem of autocorrelation successfully removed by taking suitable steps.

After removing the autocorrelation problem we again estimate the parameters of the production function and are given in the following table.

Table 5.4 Results for intrinsically nonlinear Cobb-Douglas production function (transformed data).

Name of industry	Intercept	Capital elasticity (α)	Labor elasticity (β)
Drug	79.5673	0.7709	0.0000
Furniture	0.3762	1.1344	0.1325
Iron	0.0014	0.1676	2.6591
Leather footwear	0.0052	2.4482	0.0146
Leather	528.3416	0.3364	0.6449
Metal	12.5420	0.5046	0.6669
Other chemical	0.3071	1.1225	0.1525
Paper	0.7952	0.6766	0.8622
Plastic	18961.1100	0.0744	0.0609
Textile	148795.4400	0.1786	0.0899
Wood	28943.7300	0.1247	0.2013

5.10 Concluding remarks

We observed that, there exists a positive autocorrelation of some manufacturing industries. So, we successfully remove the problem of autocorrelation and re-estimate the parameters by using SAS program.

Chapter 6

Multicollinearity

6.1 Introduction

The most important stage of an econometric research is assessing the model and the method of model estimation by econometric criteria. The acceptability of any set of parameter estimates depends on whether they possess all the econometric criteria. If these criteria are not satisfied, a model through theoretically good enough might perform imprecisely. Eventually the standard errors of the parameter estimates may be very high or the marginal effects of the explanatory variables may be entangled. Therefore these problems should be encountered and eliminated as far as possible.

Various problems arise in empirical econometric analysis. Among these problems the multicollinearity is one of the fatal problems.

This chapter is organized into the following sections. Section 6.2 represents the literature about multicollinearity. In section 6.3, we discuss about the reasons of multicollinearity and section 6.4 describe of types of multicollinearity. Section 6.5 represents the consequence of multicollinearity. In section 6.6, we discuss different test procedure of detecting multicollinearity. Section 6.7 represents the results of the study of multicollinearity detection. Section 6.8 contains some method of removing multicollinearity. Lastly in section 6.9 consists of some concluding remarks.

6.2 Multicollinearity

Multicollinearity means interrelationship among the explanatory variables and when this is strong, it is difficult to separate the effects of explanatory variables on the explained variable. The term 'Multicollinearity' is due to Ranger Frisch (1934) refers to a situation where the variables dealt with subject to two or more relations. Multicollinearity is an inherent phenomenon.

One of the main assumptions of the classical linear regression model is that there is no multicollinearity among the explanatory variables included in the regression model. If the main assumption is violated we say that there is multicollinearity in the data.
So, there can be two extreme cases:

- If the explanatory variables are perfectly linearly correlated (correlation coefficient $=1$), then the parameters become indeterminate.
- If the explanatory variables are perfectly interrelated at all (correlation coefficient $=0$), the explanatory variables are orthogonal and there is no problem.

In linear regression analysis, the term "multicollinearity" refers to the existence of more than one exact linear relationship. But this distinction is rarely maintained in practice, and multicollinearity refers to both the cases (Gujarati 1988).

In linear regression and econometrics, neither of the two extreme cases (orthogonality or perfect linear relationships among the columns) is often met in practice, because it is very rare for any two variables to be no or exactly interrelated. Usually the lack of orthogonality is not serious enough to affect the analysis. Multicollinearity is thus a persistent problem in regression as well as in econometrics. It is not a condition that either exists or does not exists in regression (economic) magnitudes. There is no conclusive evidence concerning the degree of multicollinearity which, if present, will affect seriously the coefficient estimates.

A distinction is often made between exact multicollinearity and near multicollinearity defines as: Exact multicollinearity exists when the rank of the data matrix is less than the number of columns i.,e., at least two columns of the data matrix are exactly linearly dependent. Near multicollinearity exists when the columns of the data matrix are approximately or nearly dependent i.,e., the relationship is not exact but strong. Near multicollinearty is the prevalent case in so such econometric work, especially with time series data, is one of high but not exact multicollinearity (Johnston 1988).

Professor Ranger Frisch, awarded the first Nobel memorial prize in Economics, was the first research to seriously study the multicollinearity problem in 1934 and was responsible for the introduction of the term ‘multicollinearity’ in his book “*Statistical Confluence Analysis by Means of Complete Regression System*”. Before 1960 there were no much works or publications about multicollinearity. During the time period 1960-1990, a lot of research has been done on multicollinearity and many methods, informal and formal, have been developed for detecting multicollinearity and its remedies in single equation linear regression model.

Longley (1967) used a famous set of data to demonstrate the numerical consequences of highly intercorrelated predictor. Beaton, Rubin and Barone (1976), Dent and Canvander (1977) and Judge et al. (1980) continue this discussion especially in numerical and theoretical consequences of multicollinearity. In the case of polynomials, the collinearity of the predictors is, in a sense, self-introduced and is often resolved by simply centering the data (see Marquardt and Snee (1975) and Bradley and Srivastava (1979)).

The consequences of linear dependencies in the predictors were emphasized in two papers by Hoerl and Kennard (1970a, b) where they proposed the concept of ‘ridge regression’ as an alternative to OLS in such situations. The ‘ridge trace’ provided graphic evidence of the effect of multicollinearity in the coefficient estimates. The quest for the appropriate ‘biasing parameter’, the concepts and characterizations of ridge regression and discussions of its effectiveness as well as applications of the technique have been the subject of roughly 25 papers in Technometrics in the past twelve years (Hocking 1983). In addition, many related techniques have been proposed and related to ridge regression in papers by Marquardt (1970), Mayer and Wilke (1973) and Hocking, Speed and Lynn (1976).Walker and Birch (1988) show that when ridge regression is used to mitigate the effects of multicollinearity, the influence of some observations can be drastically modified.

Gaylon and Merrill (1968) recommended taking additional data in a carefully designed experiment. Mason, Gunst and Webster (1975) discussed the sources of multicollinearity and distinguished between multicollinearities that are inherent in the system and those that may be unique to the sample. Willan and Watts (1978) suggests that we should be restrained from predicting too far from the original data and recommended an ellipsoidal region of “effective predictability” (see O. Beilley 1975, 1976).

6.3 Reasons of Multicollinearity

Problems of multicollinearity may arise mainly for two reasons:

- There is tendency of economic variables to move together over time. Economic magnitudes are influenced by the same factors. Therefore, one such influencing factor become operative, all the variables tend to change in the one direction. For example, income, saving, investment, consumption, prices and employment tend to rise in the periods of boom and decreases in the periods of depression. In time series data, therefore, growth and trend factors are the main causes of multicollinearity.
- Multicollinearity arises due to the use of lagged values of certain explanatory variables in the regression model. For example, to estimate the consumption function, past income may also be included as a separate variable along with the present incomes. Hence the problem of multicollinearity is generally observed in distributed lag models.

6.4 Types of Multicollinearity

Multicollinearity can be of the following two types

i) Perfect Multicollinearity

ii) Nearly Perfect Multicollinearity

6.4.1 Perfect multicollinearity

If the explanatory variables are perfectly linearly correlated, that is if the correlation coefficient for this variables is equal to unity, then this type of multicollinearity is called “Perfect or exact multicollinearity”.

For the regression model involving k-explanatory variables, $X_1, X_2, ..., X_k$ an exact multicollinearity is said to exist if the following condition is satisfied-

$$\lambda_1 X_1 + \lambda_2 X_2 + ... + \lambda_k X_k = 0 \qquad (6.1)$$

Where, $\lambda_1, \lambda_2, ..., \lambda_k$ are constants, such that not all of them are simultaneously zero.

Assuming $\lambda_2 \neq 0$, equation (6.1) can be written as-

$$X_2 = -\frac{\lambda_1}{\lambda_2} X_1 - \frac{\lambda_3}{\lambda_2} X_2 - \ldots - \frac{\lambda_k}{\lambda_2} X_k \qquad (6.2)$$

This shows how X_2 is exactly linearly related to other variables or how it can be derived from a linear combination of other explanatory variables. In this situation, the correlation coefficient between X_2 and other X variables is bound to be unity.

6.4.2 Near multicollinearity

If the explanatory variables are nearly or highly correlated then their exists near multicollinearity and this type of multicollinearity is called "Nearly Perfect Multicollinearity".

For the regression model involving k explanatory variables, $X_1, X_2, \ldots, X_k$, a near linear relationship is said to exist if

$$\lambda_1 X_1 + \lambda_2 X_2 + \ldots + \lambda_k X_k + V_i = 0 \qquad (6.3)$$

Where $\lambda_1, \lambda_2, \ldots, \lambda_k$ are constants, such that not all of them are zero simultaneously and V_i is a stochastic error term. Assuming $\lambda_2 \neq 0$, equation (6.3) can be written as

$$X_2 = -\frac{\lambda_1}{\lambda_2} X_1 - \frac{\lambda_3}{\lambda_2} X_2 - \ldots - \frac{\lambda_k}{\lambda_2} X_k - \frac{V_i}{\lambda_2} \qquad (6.4)$$

This shows that X_2 is not an exact linear combination of other X's because it requires error term V_i to be determined. In this situation, the correlation coefficient between X_2 variable and other X variable is not unity.

6.5 Consequences of Multicollinearity

In case of perfect or exact multicollinearity assumptions of OLS procedures breaks down, the regression coefficient of the explanatory variables are indeterminate and their standard errors are infinitely large, but this is a rare situation. The most common feature of multicollinearity, which have the following consequences:

- The variances and covariances of OLS estimates become larger and hence the confidence interval constructed from OLS estimates is wider. Consequently, prediction based on OLS estimates becomes less efficient.
- The effects of regressions on dependent variable are entangled which makes it impossible to separate the marginal effects of the regressors.
- As the standard errors of parameter estimates becomes larger which are used as denominators for computing test statistics, the effect (s) of some explanatory variables might be shown to be insignificant.
- As covariance estimates are high, one may get a few significant t-ratios in a model with which the coefficient of multiple determination $\left(R^2\right)$ is very high. This means that the effect of one regressor is reflected upon the other regressor (s) and happens due to entanglement of marginal effect on the regressors. Even for this reason, the signs and magnitudes of the coefficients may change.

6.6 Detecting multicollinearity

All of the Econometricians give the tests of multicollinearity high attention. There are several methods for testing multicollinearity. The most commonly known ones are Fisher's confluence analysis, Farrar-Glauber series of tests, thumbs rules, Method of Variance Inflation factor etc. But on the belief of some authors Condition Index (CI) is the best available multicollinearity diagnostic (Gujarati, 1995). Following is a discussion on some test rules regarding multicollinearity.

6.6.1 The High R^2

The high R^2 is a classical symptom of multicollinearity. If R^2 is high, say, in excess of 0.8, in most cases the F-test will reject the null hypothesis that partial slope coefficients are significantly equal to zero, but the individual t-test will that very few, if not all, partial slope coefficients are statistically different from zero.

6.6.2 High pairwise correlation among regressions

All pairwise or zero order correlation coefficients can be obtained from the correlation matrix $Z'Z$. The (i, j)th element of $Z'Z$, r_{ij} $\left(i \neq j = 1, 2, ..., k\right)$ gives the simple (total) correlation coefficient between X_i and X_j. A suggested rule of thumb is that if the pairwise correlation coefficient between two explanatory variables is greater than 0.80 then multicollinearity is a serious problem (Judge et al. 1982). High pairwise correlations are a sufficient but not a necessary condition for the existence of multicollinearity because it may exist even though the pairwise correlation coefficient is less than 0.50 when number of explanatory variables are more than two (Gujarati, 1988). Thus, of course, if there are only two explanatory variables in the model, the pairwise correlations will suffice. But this method will not provide an infallible guide to the presence of multicollinearity in case of more than two explanatory variables because it is possible that three or more variables are multicollinear while no pair of the variables taken alone are highly correlated. Again, a highly correlation indicates multicollinearity, but absence of high correlation can not be viewed as evidence of no problem (Belsley et al.1980).

In this method there is no suitable calibration point, even in case of two regressors, by which we can determine which correlation is high that can be considered as multicollinearity indicator. By this method we can not identify which variables are involved in multicollinear relations, how many linear dependency exist among the explanatory variables and which coefficient have been affected by multicollinearity in case of more than two explanatory variables model.

6.6.3 Bunch-Map analysis based on Frisch's 'Confluence Analysis'

Frisch (1934) showed that highly inaccurate results may be obtained without their inaccuracy being shown up by the standard errors, and he therefore recommended a systematic approach to the problem through what is known as Confluence analysis or Bunch-map analysis.

Yet none of the standard errors, the simple correlation coefficients (r_{ij}) and the overall R^2 (coefficient of determination) criteria by itself is a satisfactory indicator of multicollinearity, because

- Large standard errors may arise for various reasons and not only because of the presence of linear relationships among the explanatory variables.

- A high r_{ij} is sufficient but not a necessary condition for the existence of multicollinearity.
- The overall R^2 may be high relative to r_{ij}'s and yet the estimates may be highly imprecise and insignificant (with wrong signs/or large standard error).

However, in order to gain as much knowledge as possible as to the seriousness of multicollinearity, a combination of all above criteria may help the detection multicollinearity, a method which is in its essence a revised version of Frisch's 'Confluence analysis' (or 'Bunch-Map Analysis'). This procedure is to regress the dependent variables separately. Thus all the elementary regressions are obtained to examine their results on the basis of a priori and statistical criteria.

An elementary regression is to be chosen which appears to give the most plausible results, on both a prior and statistically criteria. Then introduced the remaining explanatory variables gradually into the chosen regression to examine their effects on the individual coefficient, on their standards errors and on the overall R^2. A new variable is classified as useful, superfluous or detrimental as follows:

- If the new variables improve R^2 without rendering the individual coefficients unacceptable on a priori considerations, the variable is considered useful and is retained as an explanatory variable.
- If the new variable does not improve R^2 and not affect to any considerable extent the values of the individual coefficients, it is considered as superfluous and is rejected.
- If the new variable affects considerably the signs or the values of the coefficients, it is considered as detrimental and this is an indication of serious multicollinearity.

6.6.4 Determinant of Cross-product matrix

If, some of the columns of the data matrix X is perfectly multicollinear, the X is not of full column rank p and hence the rank of the cross product matrix $X'X$ is less than p. Then the determinant of $X'X$ becomes zero. Multicollinearity in a sample as a departure of the observed X 's from orthogonality. The stronger the departure from orthogonality that is the closer the value of the determinant to zero, the stronger the degree of multicollinearity and the vice versa (Farrar and Glauber1967). Johnston (1984) illustrates, by treating some numerical examples, that the determinant declines in value with increasing collinearity, tending to zero as collinearity becomes exact. When the data is not exact multicollinear but near multicollinear then the value of $|X'X|$ will be small (Belsley et al.1980) and $|X'X|$ will be near to zero (very small) if the data is near perfect multicollinear (Hocking 1983; Stewart 1987) where $|.|$ denotes determinant.

But, still, one does not know what is small. Smallness is a vague concept. Generally, $|X'X|=0$ indicates exact multicollinearity and the value of $|X'X|$ near to zero indicates near multicollinearity (strong but not exact).

The main difficulties in working with determinant are that:

- There is no calibration scale for assessing what is serious and what is very serious.
- It is excessively sensitive to scaling. Stewart (1987) says that even X is scaled so that its column has length unity, the determinant of the cross-product matrix deceases factorially with the increase of the number of variables.

The determinant of $Z'Z$ falls in the interval $[0,1]$. The position of $|Z'Z|$ in the interval $[0,1]$ provides an objective measure of multicollinearity and the closer $|Z'Z|$ is to zero, the more severe the multicollinear problem. Though $|X'X|$ or $|Z'Z|$ gives us a rough idea about multicollinearity, but it definitely ignores the questions as what is the pattern of multicollinearity, which variables are multicollinear, how many linear dependency exist among the columns and which regression coefficients are affected by multicollinearity and to what extent, etc.

6.6.5 Eigen values and Condition Index

On the belief of some authors, Condition Index (C.I) is the best available multicollinearity diagnostic (Gujarati 1995). SAS and SPSS output uses eigen values and condition index to diagnose multicollinearity. This method is used to detect the degree of multicollinearity in the present study. CI is used calculated by using eigen values. The condition number k can be derived from these eigen values, which is defined as:

$$k = \frac{(\text{Maximum eigen value})}{(\text{Minimum eigen value})}$$

And the Condition Index (CI) is defined as:

$$CI = \left[\frac{\text{Maximum eigen value}}{\text{Minimum eigen value}}\right]^{\frac{1}{2}} = \sqrt{k}$$

If k is between 100 and 1000 there is moderate to strong multicollinearity; if it exceeds 1000 there is multicollinearity. Alternatively, if $\sqrt{k}$ is lies 10 to 30, there is moderate to strong multicollinearity and if the value of CI exceeds 30 then there is severe multicollinearity.

Condition Index is calculated by using eigen values. For the matrix $X'X$, eigen values are obtained by solving the matrix equation $|X'X - \lambda I| = 0$ for λ; where the elements of λ (say, λ_i) are eigen values of the matrix $X'X$ and I is the unit matrix. The formula of CI is:

$$CI = \sqrt{\frac{\lambda_i}{\min(\lambda_i)}}$$

According to thumb rule, if $10 \le CI \le 30$ there is moderate strong multicollinearity and if $CI > 30$ the problem of multicollinearity is severe. The larger the value of CI, the stronger the multicollinearity is.

6.6.6 Using auxiliary regressions

This is another procedure that is sometimes suggested as a way of detecting the presence and nature of multicollinearity is to regress each of the explanatory variables on the other $(k-1)$ regressions (Judge et al. 1982). Since multicollinearity arises because one or more of regressors are exact or approximate linear combinations of the remaining regressors, one way of finding out which X variable is related to other X variables is to regress each X_i for all $i = 1, 2, ..., k$ on the other $(k-1)$ regressors of X and compute the corresponding R^2, which we designate as R_i^2; each one of these regressions is called an auxiliary regression, auxiliary to the main regression of y on the X's. A high R_i^2 indicates a near exact linear dependence among the columns of X. Comparison of the R_i^2's shows which coefficients are likely to be most seriously affected by multicollinearity. But R_i^2's do not suffer from extensive multicollinearity, if it involves only a few variables. Unfortunately, if there are several complex linear associations, this curve fitting exercise may not prove to be of much value as it will be difficult to identify the separate interrelationships (Judge et al.1982). There is no calibration point which is able to differentiate high R_i^2. This method can not identify the nature of linear dependence if the auxiliary regressions again suffer from multicollinearity.

6.6.7 The Farrar-Glauber method

D. E. Farrar and R. R. Glauber (1967) method suggested three test statistic for detecting multicollinearity on the basis of the assumptions that

i) The $n \times k$ data matrix X is a sample of size n from a k-variate normal population.
ii) Multicollinearity in a sample as a departure of the observed X's from orthogonally and
iii) The null hypothesis is that the columns of X's are orthogonal.

The three test statistic are –

- A ***chi-square*** test for detecting the existence and severity of multicollinearity in a model including several explanatory variables.
- An ***F***-test for locating the variable(s) responsible for multicollinearity.
- A ***t***-test for finding out the pattern of multicollinearity.

6.6.8 Method of Variance Inflation Factor (VIF)

Farrar and Glauber (1967) were the first to suggest looking at the values of the $(i, j)^{\text{th}}$ element of the inverse correlation matrix to diagnose multicollinearity. Later Marquardt (1970) suggested the term "variance Inflation factor" as a measure of multicollinearity.

We partition the data matrix X as $X = [x_i, X_i]$ where x_i denotes the i$^{\text{th}}$ explanatory variable and X_i denotes the sub matrix of the $(k-1)$ remaining explanatory variables of X, then we get,

$$X'X = \begin{bmatrix} x_i'x_i & x_i'X_i \\ X_i'x_i & X_i'X_i \end{bmatrix};\ i = 1, 2, \ldots, k$$

The leading term in $(X'X)^{-1}$ is

$$[x_i' x_i - x_i' X_i \left(X_i' X_i \right)^{-1} X_i' x_i]^{-1} = \left(x_i' M x_i \right)^{-1}$$

Where $M_i = I - X_i \left(X_i' X_i \right)^{-1} X_i'$

Thus the sampling variance of the OLS estimate $\hat{\beta}_i$ is

$$Var\left(\hat{\beta}_i\right) = \frac{\sigma^2}{x_i' M_i x_i} \qquad (6.5)$$

Where $x_i' M_i x_i$ is the residual sum of squares from regression of the i^{th} explanatory variable on the other $(k-1)$ explanatory variables and the residual sum of squares decreases with increasing multicollinearity between the i^{th} explanatory variable and the remaining explanatory variables, and thus the sampling variance of $\hat{\beta}_i$ increases. Not all coefficients will be affected similarly by multicollinearity. We have,

$$R_i^2 = 1 - \frac{x_i' M_i x_i}{x_i' x_i}$$

Where $x_i' x_i$ is the total sum of squared derivations for i^{th} explanatory variable (corrected)

or, $$x_i' M_i x_i = x_i' x_i (1 - R_i^2).$$

Therefore, $$Var\left(\hat{\beta}_i\right) = \frac{\sigma^2}{x_i'x_i} * \frac{1}{1-R_i^2} \qquad (6.6)$$

In the orthogonal case $R_i^2 = 0$, letting $\hat{\beta}_{io}$ denote the estimate of $\hat{\beta}_i$, then

$$Var\left(\hat{\beta}_{io}\right) = \frac{\sigma^2}{x_i'x_i}$$

Thus if R_i^2 increases, (6.6) shows that $Var\left(\hat{\beta}_i\right)$ increases when $x_i'x_i$ is fixed. When R_i^2 nearer to one $Var\left(\hat{\beta}_i\right)$ tends to infinity. Thus if $x_i'x_i$ is held to fixed, the magnification of the sampling variance with increasing multicollinearity is given by

$$\frac{Var\left(\hat{\beta}_i\right)}{Var\left(\hat{\beta}_{io}\right)} = \frac{1}{1-R_i^2} \qquad (6.7)$$

Johnston (1984) shows by some illustrative conclusions that the relationship is highly non-linear and the magnification factor increases dramatically as R_i^2 exceeds 0.9. This magnification factor is known as 'Variance Inflation Factor' (due to Marquardt 1970) and is written as

$$VIF_i = \frac{1}{1-R_i^2} \qquad (6.8)$$

Clearly a high variance inflation factor indicates high R_i^2 (near unity), and hence points to the presence of multicollinearity. This measure is therefore, of some use as an overall indication of multicollinearity. This factor also indicates which variance of the estimate is inflated by the existence of multicollinearity. It is difficult to say, since the numerical value of VIF lies from one to infinity and there is no cutoff point to differentiate, which value of VIF is large. Its weaknesses like those off correlation matrix, auxiliary regression etc., lie in its inability to distinguish among several co-existing near dependencies and can not identify which variables are involved. Marquardt (1970) suggests a rule of thumb that VIF_i is greater than 5 indicates harmful multicollinearity (Chatterjee and Price 1977; Vinod and Ullah 1981). This measure is therefore, some use as an overall indication of multicollinearity.

6.7 Test Results

In the present study, we have used CI method for investigating the presence of multicollinearity in the data and results are presented below.

Table 6.1 Results for testing multicollinearity.

Name of Industry	Condition Index (CI)	Comment
Chemical	108.7907	Severe multicollinearity
Drug	71.5807	Severe multicollinearity
Electrical machinery	115.5392	Severe multicollinearity
Food	175.1538	Severe multicollinearity
Furniture	98.0506	Severe multicollinearity
Glass	60.7156	Severe multicollinearity
Iron	92.2044	Severe multicollinearity
Leather footwear	173.1511	Severe multicollinearity
Leather	136.6657	Severe multicollinearity
Metal	71.9928	Severe multicollinearity
Mineral	67.3298	Severe multicollinearity
Non electrical machinery	87.2222	Severe multicollinearity
Other chemical	201.8131	Severe multicollinearity
Paper	271.991	Severe multicollinearity
Plastic	96.3403	Severe multicollinearity
Pottery	129.5811	Severe multicollinearity
Printing	468.4329	Severe multicollinearity
Textile	75.6827	Severe multicollinearity
Transport	77.2806	Severe multicollinearity
Wearing	320.6777	Severe multicollinearity
Wood	104.4159	Severe multicollinearity

From the above table we observed that there is severe multicollinearity presents in the data for each industry.

6.8 Solution of the Problem of Multicollinearity

There is no doubt that the problem of multicollinearity should be encountered. But the solutions that may adopted to remove this problem may vary depending on the severity of multicollinearity, on the availability of sources of data, on the importance of factors which are multicollinear, on the purpose for which the function is being estimated and some other consideration.

Among the different remedial measures, the most commonly used ones are briefly discussed bellow

6.8.1 Increases the sample size

Multicollinearity can be removed or reduced to an acceptable level by increasing the sample size by gathering more observations. High covariance among estimated parameters resulting from multicollinearity, in an equation can be reduced by increasing the sample size, because these covariances are inversely related to the sample size.

6.8.2 Introduction to additional equation in the model

Multicollinearity can be removed or reduced to an acceptable level by introducing additional equations in the model to express meaningfully the relationship between multicollinear X 's. The addition of new equation transforms the single equation or original model to simultaneous equation model. The reduced form method can then be applied to avoid multicollinearity.

6.8.3 Dropping explanatory variable (s)

In case of several multicollinearity, one of the simplest ways is to drop one or more of the multicollinear variables. But in dropping a variable from the model, we may face the problem of specification bias, which arises from incorrect specification of the model used in the analysis.

Hence the remedy may create a worse situation than the problem itself in some situation because while multicollinearity prevent precise estimation of the parameters of the model, omitting a variable may seriously mislead as to true value of the parameters.

6.8.4 Pooling cross section and Time series data

This method is a special case of the method of Restricted Least Square method because in this method, at first, one of the parameters of the original function is estimated and then this estimated parameter is used as a restriction on the original function to estimate the remaining parameter (s).

6.8.5 Transformation of variables

One of the most important reasons for high multicollinearity between the variables in time series data is that the variables tend to move in the same direction. One of the ways to minimize this dependency is to transform the form of the variables. For that the following process is followed-

If the relation

$$Y_t = \beta_o + \beta_1 X_{1t} + \beta_2 X_{2t} + \beta_3 X_{3t} + ... + \beta_k X_{kt} + U_t \qquad (6.9)$$

Holds at time t, it must also at time (t-1) because the origin of time is arbitrary any way. Therefore we have

$$Y_{t-1} = \beta_o + \beta_1 X_{1(t-1)} + \beta_2 X_{2(t-1)} + ... + \beta_k X_{k(t-1)} + U_{t-1} \qquad (6.10)$$

If we subtract (6.10) from (6.9), we obtain

$$\Delta Y_t = \beta_1 \Delta X_{1t} + \beta_2 \Delta X_{2t} + ... + \beta_k \Delta X_{kt} + \Delta U_t \qquad (6.11)$$

where, $\Delta Y_t = Y_t - Y_{t-1}$, $\Delta X_{1t} = (X_{1t} - X_{1(t-1)})$ and so on.

Equation (6.11) is known as the first difference form, because we run the regression not on the original variables but on the difference of the successive values of the variables.

The first difference regression model often reduces the severity of multicollinearity because although the levels of X_i and X_j $(i \neq j)$ may be highly correlated, there is no reason to believe that their difference will also be correlated.

In this method, there is a loss of one observation due to the differencing procedure and therefore, the degrees freedom is reduced by one.

6.9 Concluding Remarks

In the present study the remedial measures like increases the sample size and Introduction to Additional Equation in the Model can not be applied to reduce the problem of multicollinearity because the deficiency of data. Pooling Cross Section and Time series data is beyond the scope of the study. Therefore, we have used the method of Transformation of variables to reduce the problem of multicollinearity. But the problem of multicollinearity is so serve in this study that it could not be taken care of properly by it. In this study, we consider only two explanatory variables, so we cannot use Dropping variable method.

Chapter 7

Summary and Conclusions

The growth of a country can be measured by Gross Domestic Product (GDP). GDP is substantially affected by the industrial output. Industrial gross output is a function of a capital and labor input mainly. If the effect of labor and capital input to output is at a satisfactory level in an industry or in a group of industries, then industrial investment will increases. As a result, the number of industries will increase, which will directly affect GDP and also will decrease the unemployment rate. This is why industrial input-output relationship is so important for any industry as well as for the overall industrial sector of a country.

A firm's output decision depends critically on the quantities of inputs it uses to produce the desired level of output. The production function analysis helps a firm to select the optimal combination of inputs by which it can produce the desired level of output with minimum cost and maximum profitability. In the present study, the productivity behavior in some selected manufacturing industries in Bangladesh, we use the concept of production function.

In this thesis, we investigate production behavior of some manufacturing industries by linear and nonlinear Cobb-Douglas production function. As dependent variable, we consider the Gross output of an industry and we choose two independent variables, which are Total fixed assets and Total person engaged.

In chapter four, we estimate the parameters of intrinsically linear and intrinsically nonlinear Cobb-Douglas production function. Here, we also compute the values of different model selection criteria and we compare the intrinsically linear and intrinsically nonlinear Cobb-Douglas production function. We observed that, the Cobb-Douglas-type production function (2.8) is the appropriate model than (2.7). Thus **the strictly nonlinear models (which are nonlinear with additive error terms) seem to be better than intrinsically linear model (which are nonlinear**

with multiplicative error terms). To use this selected model for forecasting purpose, we fit the model by estimating the parameters. Thus we use **numerical nonlinear estimation methods** to fit the intrinsically nonlinear model. For the intrinsically nonlinear model, we observe that the value of R^2 are very high. So it can be say that the Cobb-Douglas production function which is intrinsically nonlinear fully fits well to the yearly data.

In Chapter five, we review the literature about some existing tests for detecting autocorrelation. This chapter also review some other related issues about autocorrelation such as nature of autocorrelation, general forms of autocorrelation, problems and remedial measures of autocorrelation. We use Durbin-Watson d and Durbin's h test to detect the problem of autocorrelation. From these test, we observe that there exists a positive autocorrelation of some manufacturing industries. To solve this problem, we estimate the value of ρ by using Theil-Nagar procedure since our sample size is small. Then we again estimate the intrinsically nonlinear model for transformed data given in Table 5.4.

In Chapter six, we discuss the problem about multicollinearity. We use the "Condition Index" method to detect the strength of multicollinearity and observe that there exists sever multicollinearity among the explanatory variables. In the present study the remedial measures like increases the sample size and Introduction to Additional Equation in the Model can not be apply to reduce the problem of multicollinearity because the deficiency of data. Pooling Cross Section and Time series data is beyond the scope of the study. Therefore, we have use the method of Transformation of variables to reduce the problem of multicollinearity. But the problem of multicollinearity is so serve in this study that it can not be take care of properly by it. In this study, we consider only two explanatory variables, so we cannot use Dropping variable method. Thus, we are not able to remove the problem of multicollinearity.

From Table 4.3, we observe that the major manufacturing industries including Chemical, Furniture & fixtures (wooden), Iron & steel, Fabricated metal products, Printing & publications, Transport equipment, Food, Other chemical products, Leather footwear, Electrical machinery, Non electrical machinery, Paper & paper products and Drugs & pharmaceuticals provide good economy according to their contribution. We are to improve the efficiency of the industries like Glass & glass products, Leather & leather products, Non-metalic mineral products, Pottery & China-ware, Textile, Wearing apparel except footwear, Wood & crock products mainly to ensure a better economy since there is diseconomy of scale for these industries.

Also from Table 4.4 and Table 4.5, we may say that by enhancing more capital we can get more production especially in the Furniture & fixtures (wooden), Food, Drugs & pharmaceuticals and Other chemical products industries. The proverb “The more the labor input, the more is gross output” is appropriate for the industries like Electrical machinery, Iron & steel, Non-electrical machinery and Printing & publications.

Appendix A

Table 4.10 Values of different model selection criteria of three models under study.

Name of the industry	FPE			AIC		
	Additive error	Multiplicative error	Additive error with restricted parameter	Additive error	Multiplicative error	Additive error with restricted parameter
Chemical	12532.692	16079.234	12536.271	12476.551	16007.206	12480.114
Drug	1282500	1371944.9	1652605.41	1276755	1365799.2	1645202.5
Electrical machinery	74337.5	99879.428	193519.719	74004.501	99432.013	192652.84
Food	322451.92	463383.7	680624.77	321007.48	461307.95	677575.87
Furniture	3907.7885	13530.495	3907.38558	3890.2833	13469.885	4635.5796
Glass	217.22115	238.74623	217.267113	216.2481	237.67675	216.29385
Iron	141129.81	179930.52	141169.815	140497.61	179124.51	140537.44
Leather footwear	8400.1923	171926.6	8399.80676	8362.5632	171156.44	8362.1793
Leather	8288.75	8347.3093	8288.96014	8251.6201	8309.9171	8251.8293
Metal	33642.788	37811.141	33641.101	33492.084	37641.764	33490.404
Mineral	58178.365	70628.853	61592.9987	57917.752	70312.467	61317.089
Non electrical machinery	900.49038	3291.6714	2062.08486	896.45659	3276.9262	2052.8476
Other chemical	15035.577	36825.908	15108.0721	14968.224	36660.944	15040.395
Paper	51208.654	62323.417	51205.1224	50979.262	62044.235	50975.746
Plastic	1206.6827	2381.1099	1247.95866	1201.2773	2370.4436	1242.3684
Pottery	789.23077	2891.8991	814.097687	785.69536	2878.9446	810.45089
Printing	25229.808	33368.609	25540.4571	25116.789	33219.132	25426.047
Textile	3564326.9	3832727.9	3564604.69	3548360.3	3815559	3548636.8
Transport	62334.615	70507.198	62339.6664	62055.384	70191.357	62060.412
Wearing	2830817.3	8198048.2	2971141.55	2818136.5	8161324.6	2957832.1

Wood	661.71154	1147.2951	661.681388	658.74736	1142.1558	658.71735

Table 4.10 (Continued)

Name of the industry	HQ			SCHWARZ		
	Additive error	Multiplicative error	Additive error with restricted parameter	Additive error	Multiplicative error	Additive error with restricted parameter
Chemical	12569.45	16126.4	12573.03627	14421.4	18502.4	14425.492
Drug	1286261	1375968	1657452.024	1475773	1578697	1901653.6
Electrical machinery	74555.51	100172	194087.2562	85540.2	114931	222683.21
Food	323397.6	464743	682620.8455	371046	533216	783195.17
Furniture	3919.249	13570.2	3918.844813	4496.69	15569.5	4496.23
Glass	217.8582	239.446	217.9042947	249.956	274.725	250.00934
Iron	141543.7	180458	141583.8253	162398	207046	162444.16
Leather footwear	8424.828	172431	8424.440957	9666.1	197836	9665.6606
Leather	8313.059	8371.79	8313.269258	9537.87	9605.25	9538.1094
Metal	33741.45	37922	33739.76059	38712.8	43509.3	38710.827
Mineral	58348.99	70836	61773.63316	66945.9	81272.6	70875.086
Non electrical machinery	903.1313	3301.32	2068.132363	1036.19	3787.73	2372.8418
Other chemical	15079.67	36933.9	15152.37968	17301.4	42375.6	17384.864
Paper	51358.83	62506.2	51355.29223	58925.8	71715.6	58921.753
Plastic	1210.222	2388.09	1251.618558	1388.53	2739.94	1436.0265
Pottery	791.5454	2900.38	816.4852003	908.168	3327.71	936.7825
Printing	25303.8	33466.5	25615.35986	29031.9	38397.3	29389.413
Textile	3574780	3843968	3575058.642	4101472	4410321	4101791.9
Transport	62517.42	70714	62522.49062	71728.5	81132.7	71734.276
Wearing	2839119	8222091	2979855.044	3257422	9433497	3418893.6
Wood	663.6521	1150.66	663.6219084	761.432	1320.19	761.39701

Table 4.10 (Continued)

Name of the industry	SHIBATA			GCV		
	Additive error	Multiplicative error	Additive error with restricted parameter	Additive error	Multiplicative error	Additive error with restricted parameter
Chemical	11790.625	15127.174	11793.9918	12989.349	16665.117	12993.058
Drug	1206562.5	1290711.3	1554753.77	1329230.8	1421934.8	1712821.8
Electrical machinery	69935.938	93965.515	182061.314	77046.154	103518.76	200571.04
Food	303359.38	435946.51	640324.619	334201.18	480268.13	705424.86
Furniture	3676.4063	12729.348	3676.02723	4050.1775	14023.509	4049.76
Glass	204.35938	224.60994	204.402613	225.13609	247.44548	225.18373
Iron	132773.44	169276.74	132811.076	146272.19	186486.7	146313.65
Leather footwear	7902.8125	161746.73	7902.44978	8706.2722	178191.13	8705.8726
Leather	7797.9688	7853.0608	7798.16645	8590.7692	8651.4623	8590.987
Metal	31650.781	35572.324	31649.1937	34868.639	39188.875	34866.89
Mineral	54733.594	66446.882	57946.0449	60298.225	73202.374	63837.278
Non electrical machinery	847.17188	3096.7698	1939.98773	933.30178	3411.6108	2137.2216
Other chemical	14145.313	34645.427	14213.5152	15583.432	38167.743	15658.569
Paper	48176.563	58633.214	48173.2402	53074.556	64594.31	53070.896
Plastic	1135.2344	2240.1232	1174.06637	1250.6509	2467.871	1293.4308
Pottery	742.5	2720.6682	765.894535	817.98817	2997.2719	843.76117
Printing	23735.938	31392.836	24028.1932	26149.112	34584.469	26471.081
Textile	3353281.3	3605790.1	3353542.57	3694201.2	3972382	3694489.1
Transport	58643.75	66332.43	58648.502	64605.917	73076.286	64611.152
Wearing	2663203.1	7712637.5	2795218.69	2933964.5	8496762.5	3079401.8

Wood	622.53125	1079.3632	622.502885	685.82249	1189.0994	685.79124

Table 4.10 (Continued)

Name of the industry	RICE			SGMASQ		
	Additive error	Multiplicative error	Additive error with restricted parameter	Additive error	Multiplicative error	Additive error with restricted parameter
Chemical	13720	17602.5	13723.91774	10553.8	13540.4	10556.86
Drug	1404000	1501919	1809168.029	1080000	1155322	1391667.7
Electrical machinery	81380	109342	211853.1659	62600	84109	162963.97
Food	353000	507283	745105.0114	271538	390218	573157.7
Furniture	4278	14812.3	4277.558954	3290.77	11394.1	3290.43
Glass	237.8	261.364	237.8503133	182.923	201.049	182.96178
Iron	154500	196977	154543.7976	118846	151520	118879.84
Leather footwear	9196	188214	9195.577922	7073.85	144780	7073.5215
Leather	9074	9138.11	9074.230046	6980	7029.31	6980.177
Metal	36830	41393.2	36828.15266	28330.8	31841	28329.348
Mineral	63690	77320	67428.12492	48992.3	59476.9	51867.788
Non electrical machinery	985.8	3603.51	2257.440273	758.308	2771.93	1736.4925
Other chemical	16460	40314.7	16539.36312	12661.5	31011.3	12722.587
Paper	56060	68227.7	56056.13403	43123.1	52482.9	43120.103
Plastic	1321	2606.69	1366.186319	1016.15	2005.15	1050.9126
Pottery	864	3165.87	891.2227316	664.615	2435.28	685.55595
Printing	27620	36529.8	27960.07934	21246.2	28099.9	21507.753
Textile	3902000	4195828	3902304.08	3001538	3227560	3001772.4
Transport	68240	77186.8	68245.52957	52492.3	59374.5	52496.561
Wearing	3099000	8974705	3252618.114	2383846	6903620	2502013.9
Wood	724.4	1255.99	724.3669937	557.231	966.143	557.20538

Appendix B

The intrinsically nonlinear Cobb-Douglas type production function can be written as

$$y_t = \beta_1 L_t^{\beta_2} K_t^{\beta_3} + e_t$$

In order to estimate the parameters we minimize the following error sum squares

$$S(\beta) = \sum_{t=1}^{n} \left(y_t - \beta_1 L_t^{\beta_2} K_t^{\beta_3} \right)^2$$

In case of nonlinear estimation we use the score vector and Hessian matrix. The elements of score vector are given below:

$$\frac{\partial S(\beta)}{\partial \beta_1} = -2 * \sum_{t=1}^{n} \left[\left(y_t - \beta_1 L_t^{\beta_2} K_t^{\beta_3} \right) * \left(L_t^{\beta_2} K_t^{\beta_3} \right) \right]$$

$$\frac{\partial S(\beta)}{\partial \beta_2} = -2 * \sum_{t=1}^{n} \left[\left(y_t - \beta_1 L_t^{\beta_2} K_t^{\beta_3} \right) * \left(\ln(L_t) \right) * \left(\beta_1 L_t^{\beta_2} K_t^{\beta_3} \right) \right]$$

$$\frac{\partial S(\beta)}{\partial \beta_3} = -2 * \sum_{t=1}^{n} \left[\left(y_t - \beta_1 L_t^{\beta_2} K_t^{\beta_3} \right) * \left(\ln(K_t) \right) * \left(\beta_1 L_t^{\beta_2} K_t^{\beta_3} \right) \right]$$

Also the elements of Hessian matrix are given below:

$$\frac{\partial^2 S(\beta)}{\partial \beta_1^2} = 2 * \sum_{t=1}^{n} \left(L_t^{\beta_2} K_t^{\beta_3} \right)^2$$

$$\frac{\partial^2 S(\beta)}{\partial \beta_1 \partial \beta_2} = 2 * \sum_{t=1}^{n} \left[\left((\ln(L_t)) * \left(\beta_1 L_t^{\beta_2} K_t^{\beta_3} \right) * \left(L_t^{\beta_2} K_t^{\beta_3} \right) \right) - \left(\left(y_t - \beta_1 L_t^{\beta_2} K_t^{\beta_3} \right) * (\ln(L_t)) * \left(L_t^{\beta_2} K_t^{\beta_3} \right) \right) \right]$$

$$\frac{\partial^2 S(\beta)}{\partial \beta_1 \partial \beta_3} = 2 * \sum_{t=1}^{n} \left[\left((\ln(K_t)) * \left(\beta_1 L_t^{\beta_2} K_t^{\beta_3} \right) * \left(L_t^{\beta_2} K_t^{\beta_3} \right) \right) - \left(\left(y_t - \beta_1 L_t^{\beta_2} K_t^{\beta_3} \right) * (\ln(K_t)) * \left(L_t^{\beta_2} K_t^{\beta_3} \right) \right) \right]$$

$$\frac{\partial^2 S(\beta)}{\partial \beta_2^2} = 2 * \sum_{t=1}^{n} \left[\left((\ln(L_t)) * \left(\beta_1 L_t^{\beta_2} K_t^{\beta_3} \right) \right)^2 - \left(\left(y_t - \beta_1 L_t^{\beta_2} K_t^{\beta_3} \right) * (\ln(L_t))^2 * \left(\beta_1 L_t^{\beta_2} K_t^{\beta_3} \right) \right) \right]$$

$$\frac{\partial^2 S(\beta)}{\partial \beta_2 \partial \beta_3} = 2 * \sum_{t=1}^{n} \left[\begin{array}{l} \left((\ln(K_t)) * \left(\beta_1 L_t^{\beta_2} K_t^{\beta_3} \right) * (\ln(L_t)) * \left(\beta_1 L_t^{\beta_2} K_t^{\beta_3} \right) \right) \\ - \left(\left(y_t - \beta_1 L_t^{\beta_2} K_t^{\beta_3} \right) * (\ln(K_t)) * (\ln(L_t)) * \left(\beta_1 L_t^{\beta_2} K_t^{\beta_3} \right) \right) \end{array} \right]$$

$$\frac{\partial^2 S(\beta)}{\partial \beta_3^2} = 2 * \sum_{t=1}^{n} \left[\left((\ln(K_t)) * \left(\beta_1 L_t^{\beta_2} K_t^{\beta_3} \right) \right)^2 - \left(\left(y_t - \beta_1 L_t^{\beta_2} K_t^{\beta_3} \right) * (\ln(K_t))^2 * \left(\beta_1 L_t^{\beta_2} K_t^{\beta_3} \right) \right) \right]$$

Gauss Code 1

@ The following program is used to calculate the estimated @ value of the parameters of the nonlinear Cobb-Douglas @ type production Function

@ MAIN PROGRAM

```
new;
n=16;k=3;
load x[n,k]=drug.txt;
x=x;
y=x[.,1];
k=x[.,2];
l=x[.,3];

bb=8|0.5|.5;
vv=0;
bbb=0;
start=bB;

@ Call of Gauss optimization subroutines/built-in
@ functions at the final stage

library co;
coset;
#include co.ext;
_co_GradProc=&gp;
_co_HessProc=&Hsp;
_co_algorithm=2;
_co_options=("none");
p=b[1,1].*(k^b[2,1]).*(l^b[3,1]);
{bb,f,g,ret}=co(&fct,start);

bb;f;g;
```

@ **SUBROUTINES**

*@ The following subroutine-**fct** estimate optimal parameters @ of the model*

```
proc fct(Bb);
local ff;
ff=(y-(Bb[1]*((k^Bb[2]).*(l^Bb[3]))));
retp(ff'ff);
endp;

@ The following subroutine-GP declares gradients of each
```

```
@ log likelihood function

proc gp(bb);
local b1,b2,b3,f1,d1,db1,d2,db2,d3,db3, sb;
B1=BB[1];
B2=BB[2];
B3=BB[3];
sb=zeros(3,1);
f1=b1*((l^b2).*(k^b3));
d1=(y-f1).*((l^b2).*(k^b3));
db1=-2*sumc(d1);
d2=(y-f1).*((ln(l)).*b1*(l^b2).*(k^b3));
db2=-2*sumc(d2);
d3=(y-f1).*((ln(k)).*b1*(l^b2).*(k^b3));
db3=-2*sumc(d3);
sb[1,1]=db1;
sb[3,1]=db3;
sb[2,1]=db2;

retp(sb);
endp;

@ The following subroutine-HSP declares Hessian of each
@ log likelihood function

proc hsp(bb);
local
lnl,b1,b2,b3,f1,hh,f2,db11,db12,db13,db22,db23,db33;
b1=bb[1];
b2=bb[2];
b3=bb[3];
hh=eye(3);
f1=b1*((l^b2).*(k^b3));
f2=(l^b2).*(k^b3);
db11=2*sumc(f2^2);
lnl=ln(l);

db12=2*sumc(((lnl.*f1).*f2)-(((y-f1).*(lnl.*f2))));
db13=2*sumc(((ln(k).*f1).*f2)-((y-f1).*(ln(k).*f2)));
db22=2*sumc(((ln(l).*f1)^2)-((y-f1).*(ln(l)^2).*f1));
db23=2*sumc(((ln(k).*f1).*((ln(l).*f1)))-
((y-f1).*ln(l).*ln(k).*f1));
db33=2*sumc(((ln(k).*f1)^2)-((y-f1).*((ln(k)^2).*f1)));

hh=db11~db12~db13|db12~db22~db23|db13~db23~db33;
vv=hh;
retp(hh);
endp;
```

Gauss Code 2

```
@ The following program is used to calculate the
estimated value of the parameters of the restricted and
strictly and nonlinear Cobb-Douglas type Production
Function
@                    Main Program

new;
rndseed 1;
format /rd 16,5;
n=16; k=3;
load x[n,k]=wood.txt;
p=x[.,1];
k=x[.,2];
l=x[.,3];
beta=1|.5|.5;
s=1~.5~.5;
start4=s';
lk=0~9999999|0~9999999|0~9999999;

@ Call of Gauss optimization subroutines/built-in
@ functions at the final stage

coset;
library co;
#include co.ext;
_co_bounds=lk;
e=rndn(16,1);
pp=beta[1]*(k^beta[2]).*(l^beta[3])+e;
{kkd4, f4,g4,ret4}=co(&fct4,start4);
u=p-kkd4[1]*(k^kkd4[2]).*(l^kkd4[3]);
ess=u'u;
tss=(p-meanc(p))'(p-meanc(p));
rsq=1-(ess/tss);
bb=kkd4;
bb;
rsq;
ess;

@                    SUBROUTINES
@ The following subroutine-fct4 estimate optimal
parameters @ of the model
proc fct4(bb4);
local z;
z=(p-bb4[1]*(k^bb4[2]).*(l^bb4[3]))'(p-
bb4[1]*(k^bb4[2]).*(l^bb4[3]));
retp(z); endp;
```

Bibliography

Allen R. G. D. (1968), *Macro-Economic Theory: a mathematical treatment*, Macmillan & Co Ltd., New York.

Arrow, K., H. Chenery, B. Minhas and R. Solow (1961), Capital-Labor Substitution and Economic Efficiency: *Review of Economics and Statistics*, 43, 225-250.

Breusch, T. S. (1978), Testing for autocorrelation in dynamic linear models, *Australian Economic Papers*, 17, 334-55.

Chiang A. C. (1984), *Fundamental Methods of Mathematical Economics*, 3rd ed., McGraw-Hill, Inc., New York.

Christensen, L. R., D. W. Jorgenson and L. J. Lau (1973), Transcendental Logarithmic Production Functions, *Review of Economics and Statistics*, 55, 28-45.

Christensen R. (2001), *Advance Linear Modeling*, 2nd ed., Springer, New York.

Cox D. R. and P. J. Solomon (1986), Analysis of variability with large numbers of small samples, *Biometrika*, 73, 543-554.

Cox D. R. and P. J. Solomon (1988), On testing for serial correlation in large numbers of small sample, *Biometrika*, 75, 145-148.

Dent, W. T., and Cassing, S. (1978), On Durbin's and Sim's residuals in autocorrelation tests, *Econometrica*, 46, 1489-1492.

Dhrymes P. J. and B. M. Mitchell (1969): "Estimation of joint production functions", *Econometrica*, 37, 732-736.

Donald Erdman and Mark Little, *Nonlinear Regression Analysis and Nonlinear Simulation Models*, SAS Institute Inc., Cary, NC.

Dorothy J. Hodges (1969): "A note on estimation of Cobb-Douglas and CES production function models", *Econometrica*, 37, 721-725.

Draper, N. R., and Smith, H. (1966), *Applied regression analysis*, John Willey and Sons Inc., New York.

Duncan K. Foley, Thomas R. Michl (2001), The production function and productivity, *The Journal of Economic Perspectives*, 15, 257-258.

Durbin, J., and Watson, G. S. (1950), Testing for serial correlation in least squares regression, *Biometrica*, 37, 409-428.

Durbin, J., and Watson, G. S. (1970), Testing for serial correlation in least squares regression when some of the regressors are lagged dependent variables, *Econometrica*, 38, 410-421.

Durbin, J., and Watson, G. S. (1951), Testing for serial correlation in least squares regression, *Biometrica*, 37, 159-178.

Durbin, J. (1950), Testing for serial correlation in systems of simultaneous regression equations, *Biometrica*, 44, 370-377.

Gaofeng Han, Kaliappa kalirajan, Nirvikar Singh (2002), Productivity and economic growth in East Asia: innovation, efficiency and accumulation, *Japan and the World Economy*, 14, 401-424.

Goldfield S. M. and Quandt R. E. (1976), *Nonlinear Methods of econometrics*, North-Holland publication Company, Amsterdam, New York.

Graybill, F. A. (1961), *An Introduction to Linear Statistical Models,* Vol. 1, McGraw-Hill, Inc., New York.

Greene, W. H. (1997), *Econometric Analysis*, 3rd ed., Prentice Hall, Upper Saddle River, New Jersey.

Griffiths, W. E., Hill, R. C., and Judge, G. G. (1993), *Learning and practicing Econometrics*, New York.

Gujarati, D. N. (2003), *Basic Econometrics*, 4th ed., McGraw Hill, New York.

Hebden Julia (1983), *Application of Econometrics*, Heritage Publishing, U. K.

Henderson, J. M. and Quandt, R. E. (1971), *Micro economic Theory: a mathematical approach*, McGraw Hill, New York.

Hong, H., O. Scaillet (2006): "A fast subsampling method for nonlinear dynamic models", *Jounal of Econometrics*, 133, 557-578.

Intriligator Michael D. (1980), *Econometric Models, Techniques and Applications*, Prentice-Hall, Inc. Englewood Cliffs, N. J., USA.

Imhof, J. P. (1961), Computing the distribution of quadratic forms in normal variables, *Biometrika*, 48, 419-426.

Johnson, R. A., and Wichern, D. W. (1998), *Applied multivariate statistical analysis*, 5th ed., Prentice-Hall International, New York.

Johnston, J., and Dinardo, J., (1997), *Econometric methods*, 4th ed., McGraw Hill, New York.

Judge, G. G., and Yancy, T. A. (1986), *Improved Methods of Inference in Economatrices,* North-Holland, Amsterdam.

Judge, G. G., Griffiths, W. E., Hill, R, C., Lutkephol, H., and Lee, T. C. (1985), *The theory and practice of Econometrics,* 2nd ed., John Wiley and Sons, New York.

Kenneth J. White (1992), The Durbin-Watson test for autocorrelation in nonlinear models, *The review of Economics and Statistics*, 74, 370-373.

Kmenta J. (1964), Some properties of alternative estimates of the Cobb-Douglas production function, *Econometrica*, 32, 183-188.

Koutsoyiannis, A. (1996), *Theory of Econometrics*, 2nd ed., Macmillan Press Ltd., Hampshire.

Maddala, G. S. (1992), *Introduction to Econometrics*, 2nd ed., Maxwell Macmillan, New York.

Maddala, G. S. (1977), *Econometrics*, International student ed., McGraw-Hill, New York.

Madnani, G. M. K. (1992), *Introduction to Econometrics-Principles and applications*, 4th ed., Oxford and IBH publishing Co. Pvt. Ltd., New Delhi.

Michael Bruno, Edwin Burmeister, Eytan Sheshinski (1968), The badly behaved production function: comment, *The Quarterly Journal of Economics*, 82, 524-525.

Michel R. Veall (1990), Testing for a global maximum in an Econometric context, *Econometrica*, 58, 1459-1465.

Mood, A. M., Graybill, F. A., and Boes, D. C. (1999), *Introduction to the theory of statistics,* 3rd ed., McGraw-Hill, New York.

Nerlove, M. (1965), *The Estimation and Identification of Cobb-Douglas Production Functions*, Amsterdam.

Nerlove, M., and Wallis, K. F (1966), Use of the Durbin-Watson test in inappropriate situations, *Econometrica*, 34, 235-238.

Neter, J., Wasserman, W., and Kutner, M. H. (1990), *Applied linear statistical models,* 3rd ed, IRWIN, Homewood, Boston.

Park, S. B. (1975), On the small-sample power of Durbin's h-test, *Journal of American Statistical Association*, 70, 60-63.

Potluri Rao (1969), A note on Econometrics of joint production, *Econometrica*, 37, 737-738.

Ramanathan, R. (1995), *Introductory Econometrics with applications*, 3rd ed., Harcourt College Publishers, New York.

Reports on Bangladesh Census of Manufacturing Industries (1984), Bangladesh Bureau of Statistics, Planning division, Ministry of Planning, Dhaka, Bangladesh.

Reports on Bangladesh Census of Manufacturing Industries (1987), Bangladesh Bureau of Statistics, Planning division, Ministry of Planning, Dhaka, Bangladesh.

Reports on Bangladesh Census of Manufacturing Industries (1992), Bangladesh Bureau of Statistics, Planning division, Ministry of Planning, Dhaka, Bangladesh.

Reports on Bangladesh Census of Manufacturing Industries (1997), Bangladesh Bureau of Statistics, Planning division, Ministry of Planning, Dhaka, Bangladesh.

Reports on Bangladesh Census of Manufacturing Industries (2000), Bangladesh Bureau of Statistics, Planning division, Ministry of Planning, Dhaka, Bangladesh.

Reports on Bangladesh Census of Manufacturing Industries (2002), Bangladesh Bureau of Statistics, Planning division, Ministry of Planning, Dhaka, Bangladesh.

Reports on Bangladesh Census of Manufacturing Industries (2004), Bangladesh Bureau of Statistics, Planning division, Ministry of Planning, Dhaka, Bangladesh.

Robert M. Coen and Bert G. Hickman (2006), An econometric model of potential output, productivity growth and resource utilization, *Journal of Macroecomics*.

Ryuzo Sato, Ronald F. Hoffman (1968), Production function with Variable Elasticity of Factor Substitution: Some Analysis and Testing, *The Review of Economics and Statistics*, 50, 453-460.

Seber G. A. F. and Wild C. J. (1989), *Nonlinear Regression*, Wiley series in probability and mathematical statistics, John Wiley & Sons, New York.

Thomas S. Shively, Craig F. Ansley and Robert Kohn (1990), Fast evaluation of the distribution of the Durbin-Watson and other invariant test statistics in time series regression, *Journal of the American Statistical Association*, 85, 676-685.

Singh, S. P. and Parashar, A. K. (1999), *Econometric and Mathematical Economics*, 7th ed., Chand and company Ltd., New Delhi.

Statistical Year Book (1984), 5th ed., Bangladesh Bureau of Statistics, Statistics division, Ministry of Planning, Dhaka, Bangladesh.

Statistical Year Book (1986), 7th ed., Bangladesh Bureau of Statistics, Statistics division, Ministry of Planning, Dhaka, Bangladesh.

Statistical Year Book (1995), 16th ed., Bangladesh Bureau of Statistics, Statistics division, Ministry of Planning, Dhaka, Bangladesh.

Statistical Year Book (1997), 18th ed., Bangladesh Bureau of Statistics, Statistics division, Ministry of Planning, Dhaka, Bangladesh.

Statistical Year Book (2000), 21th ed., Bangladesh Bureau of Statistics, Statistics division, Ministry of Planning, Dhaka, Bangladesh.

Statistical Year Book (2001), 22th ed., Bangladesh Bureau of Statistics, Statistics division, Ministry of Planning, Dhaka, Bangladesh.

Statistical Year Book (2003), 24th ed., Bangladesh Bureau of Statistics, Statistics division, Ministry of Planning, Dhaka, Bangladesh.

Theil, H. (1971), *Principles of Econometrics,* New York: John Wiley.

Vinod H. D. (1969), Econometrics of joint production-A reply, *Econometrica*, 37, 739-740.

http://www.statsoft.com/textbook/stnonlin.html

http://www.bus.lsu.edu/economics/faculty/cpapageorgiou/personal/papers/jeg6.pdf

http://econ.la.psu.edu/~hbierens/EasyRegTours/NLLS.PDF

http://mpra.ub.uni-muenchen.de/932/01/MPRA_paper_932.pdf

http://dde.sdo.teilar.gr/mcda64/en/..%5Cpapers%5Cmcda64-9.pdf

http://www.econ.iastate.edu/tesfatsi/ProductionCobbDouglasFct.EmpValid.Fraser.pdf

http://ocw.mit.edu/NR/rdonlyres/Political-Science/17-874Spring2004/77594DDE-4F1A-44FA-9277-7AAD32FA15F8/0/stats_4_note_04.pdf http://robinson.gsu.edu/resources/files/ec_for/schotastic_JE.pdf

http://msu.edu/~ec/faculty/wooldridge/current%20research/panel8r2.pdf

http://www.economics.harvard.edu/faculty/doraszelski/papers/ip_september25.pdf

http://client.norc.org/jole/SOLEweb/huang.pdf

http://www3.interscience.wiley.com:8100/legacy/college/besanko/0471457698/chaps/ch06.pdf

http://www.sztaki.hu/~viharos/homepage/Publications/2002_GEPESZET_2002_OPTIMIZATION/Paper_Termelesi%20rendszerek%20optimalizalasa%20szimulacios%20es%20mesterseges%20intelligencia%20eszkozok%20alkalmazasaval.pdf

http://www2.eur.nl/WebDOC/doc/econometrie/feweco20030514165056.pdf

http://www.searchanddiscovery.net/documents/2006/06001sinha/images/sinha.pdf

http://www.fordham.edu/iped/ecga5510/serial.pdf

http://lcb1.uoregon.edu/sergiok/DSC330HSP04/week9/TASgenerated.pdf

http://www.albany.edu/econ/Research/2004/convergence.pdf

http://www.ide.go.jp/English/Publish/Ideas/pdf/machine_04_1_3.pdf

http://banglapedia.search.com.bd/HT/I_0062.htm

http://www.google.com.bd/search?q=manufacturing+industries+in+bangladesh&hl=en&ie=UTF-8&start=10&sa=N

http://papers.ssrn.com/sol3/papers.cfm?abstract_id=923295

https://idl-bnc.idrc.ca/dspace/handle/123456789/6620

http://ideas.repec.org/p/wbk/wbrwps/3988.html

http://caleb.wabash.edu/econometrics/EconometricsBook/chap20.htm

http://www.fmrib.ox.ac.uk/analysis/techrep/tr01mw1/tr01mw1/node7.html