2014

Impact of Education on Rice Production in Bangladesh: A Case Study of Three Villages of Shibganj Upazila of Bogra

Islam, Md. Sariful

University of Rajshahi

IMPACT OF EDUCATION ON RICE PRODUCTION IN BANGLADESH: A CASE STUDY OF THREE VILLAGES OF SHIBGANJ UPAZILA OF BOGRA

M.PHIL DISSERTATION

Submitted by

Md. Sariful Islam

M.Phil Fellow

Roll No. 07

Session: 2011-2012

INSTITUTE OF EDUCATION AND RESEARCH
UNIVERSITY OF RAJSHAHI
RAJSHAHI-6205
BANGLADESH

August 2014

Impact of Education on Rice Production in Bangladesh: A Case Study of Three Villages of Shibganj Upazila of Bogra

A dissertation submitted in partial fulfillment of the requirements for the Degree of Master of Philosophy

Researcher **Md. Sariful Islam**

M.Phil Fellow
Roll No. 07
Session: 2011-2012

Supervisor **Dr. M.Moazzem Hossain Khan**

Professor
Department of Economics
University of Rajshahi, Rajshahi
Bangladesh.

Co-Supervisor **Mr. Kazi Julfikar Ali**

Assistant Professor
Department of Economics
University of Rajshahi, Rajshahi
Bangladesh.

INSTITUTE OF EDUCATION AND RESEARCH
UNIVERSITY OF RAJSHAHI
RAJSHAHI-6205
BANGLADESH

DEDICATED TO

My Departed Parents

Late Alhazz Hafizar Rahman

And

Late Shamsunnahar (Sajeda)

Declaration

I do hereby declare that the thesis entitled ***"Impact of Education on Rice Production in Bangladesh: A Case Study of Three Villages of Shibganj Upazila of Bogra"*** is submitted to the Institute of Education and Research (IER), University of Rajshahi, Rajshahi, Bangladesh, for the Degree of Master of Philosophy (M. Phil) in Economics is an original and independent research work. No part of this thesis, in any form, has been submitted to any other university or institution for any other degree.

(Md. Sariful Islam)

M.Phil Fellow

Session: 2011-2012

Institute of Education and Research (IER)

University of Rajshahi

Bangladesh.

CERTIFICATION OF DISSERTATION

This is to certify that the thesis entitled **"Impact of Education on Rice Production in Bangladesh: A Case Study of Three Villages of Shibganj Upazila of Bogra"** is submitted by Md. Sariful Islam, a Fellow of the Institute of Education and Research (IER), University of Rajshahi, Rajshahi, Bangladesh, in partial fulfillment for the requirements of the Degree of Master of Philosophy (M. Phil) in Economics. It is also certified that the research work embodied in this thesis is original and carried out by him under our supervision and used in this thesis is genuine and original. No part of the work has been submitted for any others degree.

(Dr.M.Moazzem Hossain Khan)

Professor and Supervisor

Department of Economics

University of Rajshahi, Rajshahi

Bangladesh.

(Mr.Kazi Julfikar Ali)

Assistant Professor and Co-Supervisor

Department of Economics

University of Rajshahi, Rajshahi

Bangladesh.

Acknowledgement

The author expresses all the admiration to the almighty Allah, the most merciful and the most beneficial that has enabled the author to complete his humble work.

This is an opportunity as well as obligation for me to express my deep gratitude to all persons and institutions that helped me directly and indirectly in preparing this thesis.

First of all, I must express my deepest gratitude to my honorable supervisor Dr. M. Moazzem Hossain Khan, Professor, Department of Economics, and co-supervisor Mr. Kazi Julfikar Ali, Assistant Professor, Department of Economics, University of Rajshahi, Rajshahi, Bangladesh, for suggesting the problem, giving valuable advice, instruction, constant encouragement, constructive criticisms and scholastic guidance from the very beginning for the completion of this research work and for their cordial efforts in the preparation of the final manuscript.

I also express with great pleasures to all teachers, especially Professor Dr. A.N.K.Noman, Professor Dr. Md. Elias Hossain, Associate Professor Dr. Abdur Rashid Sarker, Associate Professor Mr. Md. Atiqul Islam, Assistant Professor Mr. Md. Selim Reza Department of Economics, University of Rajshahi, for their valuable advice, inspiration, and direct or indirect help at the time of my research work.

I would like to acknowledge my indebtedness to the Institute of Education and Research (IER), University of Rajshahi, Rajshahi, Bangladesh, for offering me all kinds of academic support and facilities for this research work. I would like to acknowledge the great help of Professor Dr. Golam Kabir, Director of the IER. I also express my gratitude to all the teachers of the Institute of Education and Research (IER), University of Rajshahi, Rajshahi, Bangladesh for their cordial help and encouragement. I am also thankful to all the staff of the institute for their full cooperation and support.

I express my deep sense of gratitude to those teachers of the Department of Economics, Government Azizul Haque College, Bogra especially Associate professor Mr.S.M. Mozaffer Hossain and Assistant Professor Mr. Md. Zubaidur Rahman Al Galib who have extended their generous help and active cooperation for preparing the study. I am also grateful to all the people who helped me serving as subjects of the study.

I am grateful to the Institute of Education and Research (IER) authority for providing me with financial assistance through the (IER) M. Phil Fellowship Program, to the Directorate of Secondary and Higher Education (DSHE) for processing my application for deputation for this study, and to the Ministries of Education and Establishment for granting me deputation, without which this study would have been impossible.

I would like to offer cordial thanks to the fellows of IER, University of Rajshahi for their support and cooperation conducting this research work.

I also extend my cordial thanks to officers and other staff of central library of Rajshahi University, for their cooperation and assistance during this study.

I also extend my cordial thanks to Aftab Ali Fakir, Serajul Islam, Saidur Rahman, Zillur Rahman, Saiful Islam, Matiul Islam and Bayezid Ahmed for their cooperation and assistance during this study.

I also extend my cordial thanks to officers and other staff of Upazila Agricultural Office, Shibganj, Bogra, for their cooperation and assistance during this study.

Finally, the gratitude is extended to my departed parents Late Alhazz Hafizar Rahman and Late Shamsunnahar (Sajeda) who have always encouraged me and extended their moral support. I express my sincere thanks and gratefulness to my spouse, Nusrat Jafrin Akhtar (Mala) for her kind inspiration and encouragement. It would have been an impossible task for me to complete this research work without her constant inspiration, help and cooperation. I also express my thanks to my sons Adri and Arnab. I convey my heart-felt thanks to my younger brother and all the members of my family who have helped me with their kind company and encouragement to perform such a research work.

The Author

Rajshahi, August, 2014.

ABSTRACT

This study examines the impact of education on rice production in Bangladesh. The study has been employed on farm level cross sectional data from the villages of Chapachil, Paschim Saidpur and Asrafpur of Shibganj *Upazila* of Bogra. The research approach is both qualitative and quantitative in nature. Farm level data used in this study are collected by employing random sampling technique. Structured questionnaire is used to collect data from 358 rice cultivators in the study area. The *chi-sq*uare test, analysis of variance (ANOVA) and the econometric techniques ordinary least squares (OLS) and ridge regression methods are used to assess the impact of education on rice production. The results of the study show that education has a statistically significant and positive effect on rice production. The study also shows that input cost, labour cost, cultivable land and extension service have also statistically significant and positive effect on rice production. The policy suggestion of the study is that the government should emphasize on education through literacy campaigning, training and adult continuing education programs so that rice production is increased. In addition, the government should also take initiative so that farmers can easily adopt modern agricultural inputs.

Contents

Page

List of Tables

List of figures

List of abbreviations and acronyms

ANOVA	Analysis of Variance
BADC	Bangladesh Agricultural Development Corporation
BBS	Bangladesh Bureau of Statistics
BRRI	Bangladesh Rice Research Institute
DAE	Department of Agricultural Extension
DAP	Drought Animal Power
DTW	Deep Tube Well
DW	Durbin-Watson
FY	Financial Year
GDP	Gross Domestic Product
GLS	Generalized Least Squares
GOB	Government of Bangladesh
HYV	High Yielding Variety
IPM	Integrated Pest Management
IRRI	International Rice Research Institute
Kg	Kilogram
LLP	Low Lift Pump
LM	Lagrange Multiplier
MES	Mean Square Error
MOA	Ministry of Agriculture
MP	Muriate of Potash
MV	Modern Variety
OLS	Ordinary Least Squares
SAAOs	Sub-Assistant Agricultural Officers
SSE	Some Square Error
SSP	Single Super Phosphate
STW	Shallow Tube Well
TOL	Tolerance
TSP	Triple Super Phosphate
VIF	Variance Inflation Factor
WLS	Weighted Least Squares
WB	World Bank

Chapter One

Introduction

1.1 Background of the study

Farmers' education is an important factor of rice production in Bangladesh. It is because education possesses productive value in enabling the farmers to produce a larger quality of output from the same amount of inputs and thus enhances the productivity of farmers in various ways. Even it raises consciousness among the farmers in adopting modern technology for rice cultivation. Besides, educated farmers can utilize modern agricultural methods efficiently better than uneducated counterpart. Rice is the staple food of more than 95 per cent people of our country with the increasing population. The demand for rice is also increasing day by day in our country. On the other hand, cultivable land is decreasing with the demand for habitation of the increased population as well as due to various national calamities. Therefore, the production of rice needs to be increased in order to fulfill the demand for food of the growing population. The people living in rural areas are mainly engaged in rice production. So, in order to increase rice production as well as to improve farmers' productive efficiency, education should be provided to the farmers who are directly engaged in rice production. The rate of education of our country lowers than that of in comparison to other leading rice producing countries. It is well recognized that farmers' ability as the producer can be improved through education and training within the context of changing physical and economic environment. Education may affect agricultural productivity as well as rice productivity in a number of different ways. Firstly, education is supposed to let farmers become better "managers" by enhancing their decision-making skills (Asadullah and Rahman, 2009). Secondly, education can improve farmers' access to information and therefore should allow them to potentially pay and receive better prices for the inputs used and the outputs sold (Jamison and Lau, 1982). Thirdly, better-educated farmers are able to adopt modern technologies faster as they have opportunity to be benefitted being first-mover in taking the risk of using new technology (Feder, Just, and Zilberman, 1985; Hossain et al., 1990; Lin, 1991; Asfaw and Admassie, 2004; Weir and Knight, 2004). Finally, better-educated farmers generally prefer riskier production technology (typically promising higher returns) to less risky ones since they are able to evaluate adequately the implied opportunities of technology (Asadullah and Rahman, 2009).

In order to ensure food security of our country rice production needs to be increased. To increase rice production it is urgent to enhance the productive efficiency of farmers for which proper training as well as education is urgently needed. Besides, if farmers are educated and given proper training, their use of extension service will be more effective and the educated farmers can utilize and take various agricultural training more significantly as well as accurately than the uneducated ones. Development in agriculture as well as rice production mostly comes as a result of technological change, and technological change, in turn, depends much on farmers' education. Educated farmers either formal or non-formal are believed to be more productive, particularly in use of modernizing agricultural methods. Education enhances rice productivity and consequently contributes to rural development through increased food production. Therefore, this study examines the effect of farmers' education on rice productivity.

1.2 Statement of the problem

There is a significant relationship between education and rice production. Rice is the staple food of 153.6 million people (GOB, 2013) of our country. Educated farmers can adopt new technology as well as modern inputs rapidly. It is well known that our country needs to import rice almost every year as it faces a deficit of rice. In 2011-12 FY, the total import of rice through public and private sectors was 5.23 lakh metric tons (GOB, 2012). This deficit can be overcome by enhancing the productivity of rice. Rice productivity can be enhanced both by technological improvement and by increasing of farmers' efficiency. Adoptions of new technologies are difficult especially for a developing country with low-skilled workers, low GDP and huge unemployment, where most of the farmers are illiterate and live on subsistence farming. As a result, their income level is very low comparing to developed countries. So, it is difficult for them to gear up income without education. Although agriculture is the main stream of Bangladesh economy, education for scientific method of agriculture is still felt necessity in this country. It means that lack of productive education is too acute in her agriculture. It is notable that education increases income through productivity. Most of the people live in rural areas and maintain their livelihood from the cultivation of rice.

Rice cultivation also provides a safety net for the poor. A number of studies have examined the impact of education on rice production using time series data (Minh-Phuong Ngo 2006; Alam et al.2009; Reimers &Klasen2012). However the use of time series data at aggregate level for education to analyse its impact on rice production could be erroneous. Aggregate level data cannot capture the variations in farm level micro-data. This study, therefore, examines the effect of individual farmer level of education on rice production using cross-section data. Given the importance of rice production in the country, this study aims to estimate the impact of education on rice production.

1.3 Research question

This study endeavours to find the answer of the following research question.

1. How is the effect of the different levels of education on rice production?

1.4 Objectives of this study

The principal objective of this study is to analyse the impact of education on rice production taken Bogra district of Bangladesh as a case study. The specific objectives are:

1. to provide an account of socio-economic conditions of rice farmers.
2. to examine the effect of primary, secondary, higher secondary and tertiary levels of education on rice production; and
3. to guide some policy suggestions based on the findings of the study.

1.5 Rationale of the study

Rice is the main and most dominant food crop in our country. It provides 47.5 per cent of rural employment (GOB, 2013). More than 95per cent of population consumes rice and it alone provides 76 per cent of calorie and 66percent of total protein requirement of daily food intake (Bhuiyan*et al.,* 2002). About 77 per cent area of arable land is used for rice production of our country (IRRI, 2012). Given the importance of rice in Bangladesh, this study focuses the impact of education on rice production. Agricultural sector alone is employing 47.3 per cent of the total labour force of the country (GOB, 2010). That is why; rice is not only the mainstay of rural populace, but also the main occupation of the nation as a whole.

1.6 Limitations and scope of the study

Every research has some limitations. The study focuses the impact of education on rice production. Only *aman* rice is taken as sample. Due to time and budget constraints, the study cannot include *boro* rice. Cultivators do not keep written documents about rice cultivation. For this reason, when the researcher interviewed them, they depended upon their memory to recall, or called their family members to supply perfect information about rice cultivation. This study is only conducted within three villages of Shibganj *Upazila* of Bogra. It is another limitation of this research. In spite of having these limitations, it is expected that the study is conducted fruitfully by using scientific methodology of research. Aman is one of the main crops in the study area and by considering its importance; this study is conducted to show the impact of education only on *aman*.

1.7 Organization of the thesis

The present thesis has been divided into seven Chapters. Chapter one deals with introduction, statement of the problems, objectives of the study, rationale of the study, limitations of the study and organization of the thesis. Chapter two deals an overview of rice production in Bangladesh. Chapter three provides the review of related literature and research gap. Chapter four deals with methodology of the study , different concepts, selection of study areas, source of data, selection of sample size, sampling methods, methods of data collection, period of data collection, nature of data, techniques of data processing and analysis, model used for the study and validity and reliability of collected data have been explained in this chapter. Socio-economic condition of the respondents is represented in chapter five. Results and discussions of the study are explained in chapter six. Chapter seven contains the summary of the main findings, policy suggestions and conclusion of this study.

Chapter Two

An Overview of Rice Production in Bangladesh

2.1 Introduction

Since the independence of Bangladesh, the country has been striving hard for rapid development of its economy. In all respects, the economic development is based on agriculture. It is often argued that the future development of the country depends particularly on the development and proper management of the agricultural sector. More specifically the food problem is the most crucial aspect of economic development from political point of view. It bears upon the rate and structure of economic growth, rate of inflation, poverty and nutrition, and the balance of trade and government's fiscal position. The contribution of agricultural sector in Gross Domestic Product (GDP) stood at 19.29 per cent in FY2011-12 (GOB, 2012). In agricultural sector, the crop-sub-sector dominates with 14.32 per cent in GDP of which rice alone contributes 53 per cent. In Bangladesh, although 43.6 per cent of the total labor force is directly engaged in agriculture and 78 per cent of total crop is devoted to rice production, the country has still shortage of food grains. It needs to mention here that the total food grains imported by public and private sector in financial year 2011-12 are 22.91 lakh metric tons (GOB, 2012). Bangladesh government receives 0.55 lakh metric tons of food grains as food aid in financial year 2011-12 (GOB, 2012).

Bangladesh provides nearly 47.5 (GOB, 2012) per cent of rural employment, about two-third of total calorie supply and about one-half of the total protein intakes of an average person in the country (BRRI, 2013). Rice sector contributes one-half of the agricultural GDP and one-sixth of the national income in Bangladesh (BRRI, 2013). Rice is not only used in course form but also it is by products are utilized in many sectors such as straw is used as fodder, packing material and manufacturing cardboard. Its flour is used in confectionary and bakery products. Bangladesh stands as fourth rice growing country in the world (Hussain and Iqbal, 2011). Almost all of the 13 million farm families of the country grow rice. Rice grows on about 10.5 million hectares, which has remained almost stable over the past three decades. About 75per cent of the total cropped areas and over 80 per cent of the total irrigated areas are used planting rice (BRRI, 2013).Thus, rice plays a vital role in the livelihood of the people of Bangladesh.

Total rice production in Bangladesh was about 10.75 million tons in the year 1971 when the country's population was only about 70.88 millions (BRRI, 2013). However, the country is now producing about 35.00 million tons to feed her 160 million people (BRRI, 2013). This indicates that the growth of rice production is much faster than the growth of population. This increased rice production has been possible largely due to the adoption of modern rice varieties on around 66 per cent of the rice land, which contributes to about 73 per cent of the country's total rice production (BRRI, 2013).

However, there is no reason to be complacent. The population of Bangladesh is still growing by two million every year and may increase by another 30 million over the next 20 years (BRRI, 2013). Thus, Bangladesh will require about 27.26 million tons of rice for the year 2020 (BRRI, 2013). During that time, total rice area will also shrink to 10.28 million hectares. Rice yield therefore, needs to be increased from the present 2.74 to 3.74 thousand/hectares (BRRI, 2013).

It is a prime concern to enhance the growth of rice production through increasing land productivity to meet the increasing food demand for the vast population of the country. As the country has serious land constraints, significant differences in rice productivity among the different regions are also barriers to the production growth (Hossain, 1990; Hossain and Ahmed, 1989). Many steps are taken to enhance the growth from the part of the government and non-government agencies since the independence of the country.

2.2 Share of agricultural production in GDP of Bangladesh

Bangladesh agriculture is the tiger point of all economic activity. Harvesting period and bumper production are the two key factors for generating economic activity. Over the last three decades, contribution of agriculture to GDP is declining. Expansion of garments, manufacturing and service sectors and overall low growth of agricultural sector are the vital factors of declining trend of agriculture-GDP ratio in financial year 2012-13, the combined contribution of all sub-sectors of agriculture (Crop, livestock, Forestry and Fisheries) to GDP is 18.70 per cent. Table 2.1 shows the share of agricultural crop sectors in GDP. The crop sector alone projects to contribute 10.25 per cent to Gross Domestic Product (GDP), of the total labour force in Bangladesh, 47.5 per cent is engaged in agriculture (GOB, 2013).

Table 2.1: Share of agricultural sectors in GDP in 2002-03 to 2012-13. (Per cent)

Crops sector	2002-03	2003-04	2004-05	2005-06	2006-07	2007-08	2008-09	2009-10	2010-11	2011-12	2012-13
Crops & vegetables	13.43	13.23	12.51	12.28	12.00	11.64	11.43	11.42	11.32	10.86	10.25

Source: Economic Review 2013.

2.3 Rice seasons in Bangladesh

The country grows a wide variety of crops which are broadly classified according to seasons in which they are grown into two groups (a) Kharif and (b) Rabi crops.

Kharif crops grow in the spring or the summer season and are harvested in late summer or in early winter. Rabi crops grow in winter and are harvested in the spring or early summer. Table 2.2 provides a brief description of crop seasons in Bangladesh.

Table 2.2: Time of sowing, transplanting, and time of harvesting of different rice

Crop	Time of Sowing/transplanting	Time of harvest
1.Aus		
a) Broad cast	Mid March to mid May	July to mid August
b) Transplant	Mid April to mid June	-do-
c)HYV	Mid march to mid April	-do-
2.Aman		
a) Transplant	Mid June to mid August	Mid November to mid January
b) Broad cast	March	-do-
c) HYV	Mid June to mid August	-do-
3.Boro		
a) Local	Mid November to mid January	Mid April to mid June
b) HYV	Mid December to mid February	Mid April to mid June

Source: DAE and MOA

2.4 Input use in rice production of Bangladesh

A descriptive picture of changes in input use in rice production of Bangladesh has been discussed. The description of input use follows the input classification used in this study. Main inputs are classified as land, labour, fertilizer and irrigation. In the post-independence period, a rapid expansion of irrigation, fertilizer and modern variety (MV) seeds caused a breakthrough on rice production in Bangladesh that made steady progress in rice production. The main inputs, which are used in rice production of Bangladesh, are briefly discussed below.

2.4.1 Land

Land is the main source of all agricultural activities. As an agrarian economic base, land occupies a dominant position in Bangladesh rice sector. Since independence, like other inputs, the use of land changes substantially over time. Bangladesh agriculture has witnessed changes in the patterns of land utilization. Table 2.3 presents the historical development of trend of changing land utilization in agriculture of Bangladesh since 1981-82 to 2007-08. Table 2.3 depicts scenario of land utilization for the period 1981-82 to 2006-07. During this period, total cropped area increased due to increasing the sown area more than one and that was the result of changing land use patterns or changing land utilization. Culturable waste and current fallow decrease sharply and the land not available for cultivation increases a little. Furthermore, demographic pressures and increased urbanization have caused cultivated area to decline at a rate of 1 percent per year, while cropping intensity has virtually reached its limit (GOB, 2008).

Table 2.3 Trends of changing land utilization in agriculture of Bangladesh from 1981-82 to 2005-06 ('000' acres)

Year	Forest	Not available for cultivation (a)	Culturable waste (b)	Current fallows (c)	Net cropped area	Area sown more than once	Total cropped area (d)
1981-82	5298	6837	611	1350	21212	11426	32638
1982-83	5296	6876	572	1196	21369	11761	33130
1983-84	5205	7156	810	1124	21442	11571	33013
1984-85	5297	7193	721	1221	21353	11143	32496
1985-86	5237	7220	670	997	21661	11798	33459
1986-87	4910	8141	660	973	21878	11097	34883
1987-88	4703	7685	890	2913	20478	13670	34148
1988-89	4703	7645	888	3285	20148	13739	33887
1989-90	4703	7783	863	2686	20633	14117	34750
1990-91	4693	7958	1442	2379	20198	14482	34680
1991-92	4674	9885	1532	862	19716	14405	34121
1992-93	4674	10137	1512	928	19418	14438	33856
1993-94	4674	10355	1566	984	19090	14225	3331
1994-95	4861	10128	1547	1000	19133	14280	33413
1995-96	5317	9788	1314	969	19281	14110	33391
1996-97	5329	9681	1295	963	19401	14688	34089
1997-98	5572	9268	1241	898	19690	15210	34810
1998-99	5572	9141	1100	1115	19741	14752	34493
1999-00	6490	8435	781	862	20101	15166	35267
2000-01	6490	8427	794	987	19970	15366	35335
2001-02	6365	8676	799	1005	19824	15252	35076
2002-03	6418	8685	764	957	19845	15281	35126
2003-04	6418	8697	736	957	19843	15286	35129
2004-05	6420	8724	663	1157	19703	5142	134845
2005-06	6420	8802	640	1518	19289	14655	33944

Source: Agriculture Statistics, BBS, 2008.

Notes: (a) Agriculture included crops, livestock, fishery and forestry; (b) Culturable waste is the area suitable for cultivation but lying fallow for more than one year; (c) Current fallow is the area already brought under cultivation, but not cultivated during the year; (d) Total cropped area is the sum of the net-cropped area and the sown more than once.

Total cropped area in Bangladesh is graphically shown in Figure 2.1. From this figure, it is observed that a slow increasing tendency was found for the total cropped area in Bangladesh while the curve for the last two years showed a declining tendency. This may be caused by the natural disasters. In 1993─1994 and 1994─1995 food grain production declined due to depressed prices of rice and natural disasters particularly floods, and droughts in the north west region of the country. During these years, more than 2 per cent reduction in area shown was also observed (GOB, 1998).

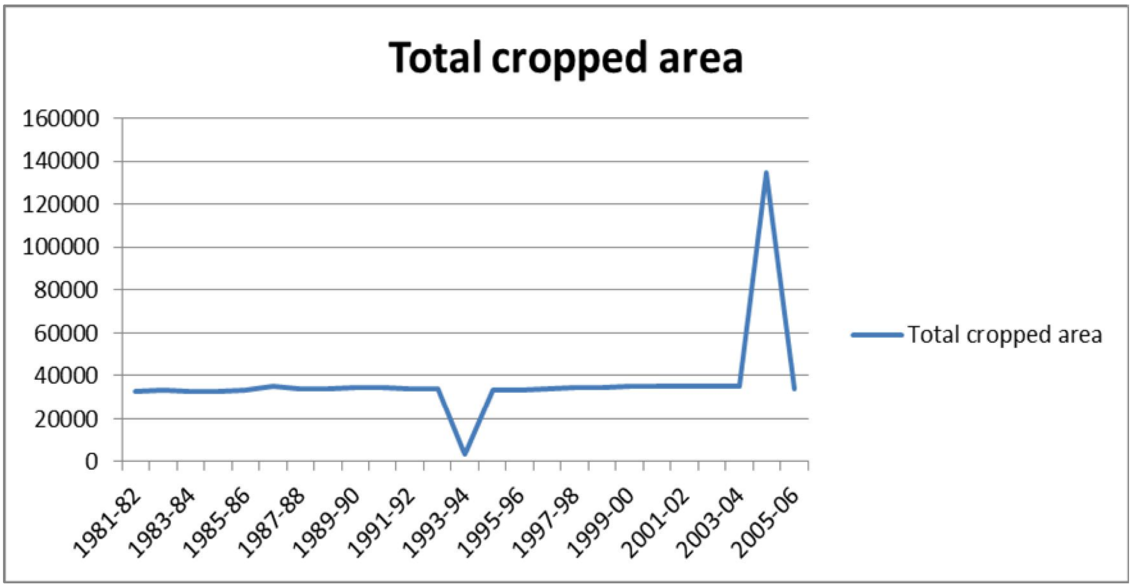


Figure 2.1: Total cropped area in Bangladesh from 1981-82 to 2005-06

Total rice area in Bangladesh is presented in Table 2.4. From Table 2.4, it is observed that the total rice area has gradually increased overtime. During 1971-72 only 9278.7 thousand hectares were rice area but in 2010-11 it gradually increased to 11528.51 thousand hectares.

Table 2.4: Total rice area in Bangladesh (000'ha)

Year	Aus	Aman	Boro	Total
1971-72	3001.60	5410.70	866.40	9278.7
1972-73	2930.00	5713.80	1002.60	9646.4
1973-74	3107.90	5718.70	1222.70	10049.3
1974-75	3179.10	5449.90	1161.20	9790.2
1975-76	3419.90	5759.90	1147.90	10327.7
1976-77	3217.10	5806.40	854.20	9877.7
1977-78	3161.70	5771.20	1093.70	10026.6
1978-79	3234.60	5805.10	1071.80	10111.5
1979-80	3036.30	5972.70	1148.40	10157.4
1980-81	3111.20	6035.80	1160.00	10307
1981-82	3145.60	6010.30	1301.70	10457.6
1982-83	3158.10	5993.00	1432.80	10583.9
1983-84	3138.10	6006.70	1401.20	10546
1984-85	2937.60	5710.20	1574.40	10222.2
1985-86	2844.90	6018.90	1533.20	10397
1986-87	2903.60	6052.40	1651.70	10607.7
1987-88	2788.30	5590.40	1942.60	10321.3
1988-89	2683.46	5100.80	2438.30	10222.56
1989-90	2255.00	5702.50	2453.60	10411.1
1990-91	2107.30	5775.30	2547.90	10430.5
1991-92	1915.90	5692.30	2634.90	10243.1
1992-93	1735.10	5843.70	2598.90	10177.7
1993-94	1649.40	5843.30	2580.80	10073.5
1994-95	1663.75	5594.17	2663.54	9921.46
1995-96	1541.85	5646.40	2753.57	9941.82
1996-97	1592.29	5802.49	2782.59	10177.37
1997-98	1565.88	5808.45	2888.56	10262.89
1998-99	1442.26	5165.50	3526.67	10134.43
1999-00	1351.32	5704.87	3651.89	10708.08
2000-01	1325.23	5709.96	3761.84	10797.03
2001-02	1242.18	5647.22	3771.34	10660.74
2002-03	1243.72	5682.11	3844.84	10770.67
2003-04	1202.58	5677.61	3943.50	10823.69
2004-05	1024.68	5279.92	4063.79	10368.39
2005-06	1034.27	5429.01	4065.81	10529.09
2006-07	905.71	5415.62	4250.10	10571.43
2007-08	918.66	5048.16	4607.85	10574.67
2008-09	1065.56	5497.77	4716.31	11279.64
2009-10	984.22	5662.89	4706.60	11353.71
2010-11	1112.87	5645.64	4770.00	11528.51
2011-2012	1200.00	5630.00	4780.00	11610.00

Source: BBS and MOA, 1972-2012

Total rice area in Bangladesh is graphically displayed in Figure-2.2. From this figure, it can be seen that the total rice area increases gradually while fluctuation is observed for some specific years.

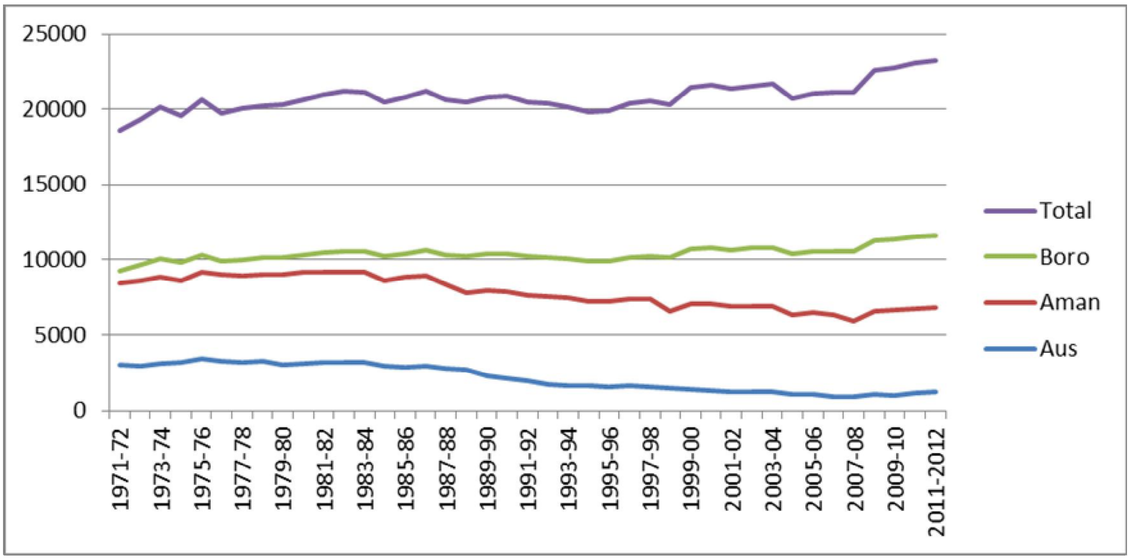


Figure 2.2: Total rice area in Bangladesh

Table 2.5 presents the trends of land utilization, which exhibits the statistics of single-cropped area, double-cropped area and triple-cropped area. It is seen from Table 2.5 that in 1974-75 single-cropped area was 12481 thousand acres and it gradually decreased to 7228 thousand acres in 1994-95. While single-cropped area has been decreased over time, double-cropped area and triple-cropped area has gradually been increased.

Table 2.5: Trends of land utilization in Bangladesh, 1974-75 to 1994-95. (In thousand acres)

Year	Single-cropped	Double-cropped	Triple-cropped
1974-75	12481	6711	1279
1975-76	12250	7269	1449
1976-77	11911	7072	1462
1977-78	11683	7520	1489
1978-79	11364	7826	1610
1979-80	11390	7865	1618
1980-81	11456	8040	1661
1981-82	11464	8070	1678
1982-83	11336	8251	1689
1983-84	11526	8340	1512
1984-85	11626	8199	1472
1985-86	11516	8492	1653
1986-87	10781	9189	1908
1987-88	9168	8949	2362
1988-89	8825	8908	2415
1989-90	8980	9191	2463
1990-91	8140	9634	2424
1991-92	8387	10722	2652
1992-93	7416	9566	2436
1993-94	7229	9497	2364
1994-95	7228	9530	2375

Source: Statistical Yearbook of Bangladesh (Various issues), Dhaka: Bangladesh Bureau of Statistics.

2.4.2 Labour

Agriculture is the largest employer of labour. As a primary sector in the economy, agriculture provides the source of income. In the past, the crop sector generated most of the employment opportunities in the agricultural sector. Overtime, importance of labour in the crop sector has been gradually decreasing. Since smaller farms need fewer hired workers, the continuing fragmentation of land holding and resultant proliferation of smaller farms have tended to decrease the overall demand for agricultural workers.

According to various censuses, it is observed that the participation of labour force in the rice sector declined but in 2011, it increased. Table 2.6 shows the present nature of employment. Table 2.6 shows that in 1974, 77.2 per cent of labour force engaged themselves in the traditional agriculture while the remaining 22.8 per cent was engaged in non-agricultural activities. However, in1991 the occupational dependencies in agriculture sector declined. This happened due to the transfer of agricultural labour force in the traditional agriculture sector. It should be noted that household-based economic activities were excluded in these censuses. The household activities consist of threshing, cleaning and processing of food grains, care of livestock, poultry etc., which are mostly done by the females.

Table: 2.6 Distribution of economically active population by sex and by major occupation groups in Bangladesh, 1961-2011 census years.

Year	Sex	Major Occupation (in percent)	
		Agriculture %	Non-Agriculture/Other %
1961	Both sex	86.0	14.0
	Male	85.0	15.0
	Female	91.8	8.2
1974	Both sex	77.2	22.8
	Male	77.5	22.5
	Female	69.8	30.2
1981	Both sex	61.3	38.7
	Male	63.0	37.0
	Female	28.0	72.0
1991	Both sex	54.6	45.4
	Male	57.5	42.5
	Female	18.0	82.0
2001	Both sex	50.9	49.1
	Male	52.2	47.8
	Female	43.9	56.1
2011	Both sex	76.9	23.1
	Male	78.2	21.8
	Female	56.8	43.2

Source : Bangladesh Population Census, National Series (2001 &2011), BBS.

It is interesting to note that women also play a significant role in this sector. The major portion of rural women largely appears as unrecognized contributors to the agricultural sector and other economic activities. In Bangladesh, as a subsistence family occupation, women used to be involved in post-harvest work by using the dheki (one type of food operated mortar and pestle which is used for husking, grinding etc. of food grains). But after the introduction of mechanized husking and polishing of grains men are now engaged in these operating in modern mills. About the participation of women in the agricultural sector, the World Development Report 1990 reported that emerging financial institutions like Grameen Bank also helped to grow the women workers forum in Bangladesh. This report showed how the women workers used the loans of Grameen Bank, which helped greater participation of women in the agricultural sector.

According to the labour force survey 2010, a number of agricultural labour forces (both sex) decreased to 47.3 percent from 48.1 percent in labour force survey 2005-06 .On the other hand, the number of the labour forces (both sex) in the non-agricultural sector increased to 52.7 percent while it was 51.9 percent in 2005-06 (Labour Force Survey,2010). As compared to the labour force survey 2010, both male and female labour forces in the agricultural sector also decreased according to this survey.

In rural areas, major portion of non-farm household depends on agriculture due to lack of other sources of employment. Households having less than 0.05 acre of cultivated land are treated as non-farm households. Agricultural labour households mean those households whose earnings come from labour services provide to the agriculture sector (BBS, 1989). In Bangladesh, members of agriculture labour households are often found to be engaged in auxiliary occupations such as livestock rising, poultry rising, cottage industry, fishing etc. as sources of subsidiary income.

There are two types of labour used in agriculture sector. There are (i) Hired labour and (ii) Family labour.

(i) Hired labour cost: Hired labour cost is calculated by summing up the amount of man-days (one man-days =8 hours) employed for cultivating per acre land during the crop season. It is also estimated by adding up total labour employed times the prevailing local wage rate.

(ii) Family labour cost: Family labour cost is estimated by adding up total labour spend during the crop season times 80 per cent of local wage rate. But here, it should be noted that in calculating profitability of tenanted land all family labours have been transferred into hired labour in such a way that family labour would reflected the hired labour cost in the production process.

2.4.3 Fertilizer

Fertilizer is of immense importance for sustained increase in crop production. Since independence, the use of chemical fertilizer has steadily increased. Due to large production and expended marketing system, there has been a rapid expansion of fertilizer use in Bangladesh. In 1950, chemical fertilizer was first introduced in Bangladesh but significant use of chemical fertilizer was not found until 1960. At that time, chemical fertilizer was mostly used in tea gardens and government experimental farms. Since then Bangladesh Agricultural Development Corporation (BADC) started to sell chemical fertilizers to the farmers.

The government initially gave subsidy to make fertilizer more popular to the farmers. Gradually this subsidy has been reduced. The Government introduced a new marketing system in 1982, which deregulated retail trade in fertilizer. Due to withdrawal of explicit subsidy from the phosphate and potash fertilizers and handing over of their trade to the private sector, prices of these items increased gradually. But in 1994 and in early 1995, the price of urea was reduced (GOB, 1998).

With the introduction of HYVs of rice in Bangladesh, consumption of fertilizer increased sharply. Fertilizer is mostly applied on the *boro* rice under irrigated conditions. The July-October period is the fertilizer season for *aman* rice while the various *rabi* crops like Potato, wheat, mustard and sugarcane are associated with fertilizer from November to March. Fertilizer is used on the *aus* rice and Jute from April to June. In case of application rate of fertilizer, *boro* rice alone consumes nearly 2.5 times more fertilizer than all crops combined (except potato). Since independence, fertilizer consumption is increasing markedly. The annual average growth rate was above 9 percent between 1969-1970 and 1989-1990 (GOB, 1990).

The various types of chemical fertilizers like, Urea, TSP, SSP, MP etc. are used in Bangladesh. Urea is the most important used fertilizer than triple super phosphate (TSP) and Muriate of Potash (MP). Table 2.7 presents the consumption pattern of various types of fertilizer in Bangladesh. A rapid transition in fertilizer consumption was observed during 1972-73 to 1994-95 periods. In 1972-73, urea consumption was only 277 thousand metric tons while it rapidly increased to about 1746 thousand metric tons in 1994-95. Triple super phosphate (TSP) gradually rose up to 1990-91 and from 1991-92 it continued to fall till the last stage of the period, perhaps due to the introduction of SSP (single super phosphate) fertilizer. Muriate of potash (MP) also significantly increased during this period.

Table 2.7: Distribution of fertilizer consumption in Bangladesh from 1972-1973 to 1994-1995.
(In thousand metric tons)

Year	Urea	TSP	MP
1972-1973	277	89	18
73-74	268	94	18
74-75	176	76	18
75-76	312	111	22
76-77	353	126	22
77-78	480	192	41
78-79	471	178	47
79-80	353	203	46
80-81	560	215	45
81-82	519	208	45
82-83	619	203	50
83-84	708	261	63
84-85	832	346	69
85-86	647	259	52
86-87	841	311	68
87-88	886	328	72
88-89	1145	416	94
89-90	1369	480	119
90-91	1322	515	147
91-92	1531	457	136
92-93	1545	407	122
93-94	1577	234	138
94-95	1775	180	140
95-96	2045	111	155
96-97	2119	72	219
97-98	1872	62	193
98-99	1902	170	210
99-00	2151	259	239
00-01	2121	399	123
01-02	2247	401	233
02-03	2247	375	270
03-04	2324	361	240
04-05	2523	420	260
05-06	2451	436	290
06-07	2515	340	230
07-08	2762	392	262
08-09	2532	156	75
09-10	2409	420	263
10-11	2652	564	482
2011-12	2296	678	613

Source: Yearbook of Agricultural Statistics of Bangladesh, 2011, BBS, Ministry of Agriculture, Economic Review 2011-13 and Khalil, M.I (1991).

2.4.4 Irrigation

Irrigation is the artificial application of water to the land or soil. It is used to assisting in growing of agricultural crops. Farmers largely depend on rainwater for cultivation. But rainwater is gradually decreasing due to climate change. It is possible to increase crop production through proper water management. Irrigation ensures a constant supply of water, which is essential not only to crops growing but also to the quality of the crop. Water supplies two essential elements hydrogen and oxygen to the crop. Irrigation is one of the key factors making the country self-sufficient in food grain production and contributes greatly towards agriculture Gross Domestic Product (GDP). Irrigation particularly ground water irrigation plays an important role for rice production. Because if the irrigation is done properly the farmers can work properly on their fields without facing any problem of water, it is also necessary because during every famine, no field is left without water and even then our country can progress well. Efforts have been continuing to ensure the use of underground and surface water in an integrated and planned manner to increase cropping intensity, diversification and yield while maintaining environmental balance.

There are two types of irrigation system in Bangladesh, (i) Traditional system likes swing baskets, dhones etc. and (ii) Modern irrigation system likes Low Lift Pump (LLP), Shallow tube wells (STWs) etc. Before introduction of modern irrigation system, farmers mostly used traditional methods like swing baskets and dhones for lifting surface water.

Since independence, use of irrigation water has been increasing and it should be regarded as the most modernized component of Bangladesh agriculture. Land utilization can be increased through irrigation in recent years. In the regard, the government resulting in a rapid increase of irrigation undertook vast programs. Effective flood control, drainage, and minor irrigation on supplementary basis play a vital role in the education of low-risk modern technology. Among the modern technology Low lift pumps (LLPs), Shallow tube wells (STWs) and Deep tube wells (DTWs) are dominant.

At present both traditional and modern methods exist in Bangladesh. Among the traditional methods, swing basket is dominant. As primitive method, open dug well is widely used, particularly in areas where water level is relatively higher. Among modern systems, LLP, DTW, and STW are the most important ones. For the small-scale irrigation, hand tube well (HTW) is also used by the poor farmers. Bangladesh Agricultural Development Corporation (BADC) mostly undertakes major irrigation projects such as canal digging and DTWS.

The history of post-independence irrigation development can be divided into four phases. In the first phase, public sector LLPs and DTWs covered most of the irrigated area up to 1979. In the second phase, mainly STWs in the private sector liberally expanded from 1979 to 1989. In the third phase, form 1984 to1987, expansion of irrigation was constrained by the limitations of STW operation in the private sector. Drop in the ground water level in many places where STWs operated and government imposed embargo on diesel engines added to the administrative bottlenecks created by privatization. In the latest phase, starting from 1987, import bar on diesel engines has been removed and there has been rapid expansion in private sector sales of STWs and DTWs. From1988-89 under the new policies, duties and standardization restriction of imported small diesel engines were removed (GOB,1990). These helped to encourage the rapid expansion of private sales of STWs and LLPs. Since 1973-74, sale of STWs increased rapidly as it could be purchased with credit facilities available from Bangladesh Krishi Bank and its Rajshahi Division counterpart Rajshahi Krishi Unnayan Bank.

LLPs are used in areas where surface water is available. Area irrigated by LLPs has increased overtime. To enable participation of the farmers in the canal digging programs the objective is to supply pump gets free of cost to the participating farmers. The main objective of this scheme is to preserve surface water, which could be utilized during winter .At the beginning of modern irrigation, surface water, which is extracted by LLP, is utilized more than other methods. Following the subsidized rental system, the expansion of LLPs started from the mid-1970s.

The spurt of expansion of ground water development began in 1967-68. In recent years, the relative importance of ground water has been growing. Total use of tube wells is higher than LLPs. Irrigation by DTWs and STWs has grown in those areas where adequate surface water is not available. Area irrigated by DTWs is increasing gradually. In recent years, STWs has become the most popular irrigation method due to ease of installation. However, the Briand of northern Bangladesh and Chittagong Hill Tracts is unsuitable for STW irrigation.

Area irrigated by different methods in Bangladesh is presented in Table 2.8. From the Table 2.8 it is observed that area irrigated by modern methods has gradually increased overtime. During 1979-80, only 180.50 thousand hectares were irrigated by DTWs but in 2004-05, it gradually increased to716.68 thousand hectares. In case of STWs, only 55.10 thousand hectares were irrigated in 1979-80 but it rapidly increased to 3027.40 thousand hectares in 2004-05. From Table 2.8, it is also observed that to irrigate the cultivated area; STWs played the most dominant role. Area irrigated by LLPs was about 621.40 thousand hectares in 1979-80 and in 2004-05; it increased to 802.07 thousand hectares.

Table: 2.8 Distribution of irrigated area by different methods in Bangladesh from 1979-80 to 2004-05. (Figures in Thousand hectares)

Year	Methods		
	LLPs	DTWs	STWs
1979-80	621.40	180.50	55.10
1980-81	665.80	259.50	99.20
1981-82	704.20	323.40	202.10
1982-83	746.60	390.10	298.70
1983-84	666.70	415.40	303.50
1984-85	580.50	441.20	300.40
1985-86	608.70	358.50	534.50
1986-87	659.60	387.50	521.20
1987-88	527.20	432.30	868.50
1988-89	657.70	501.90	899.10
1989-90	657.30	428.10	1045.10
1990-91	674.80	615.50	1131.11
1991-92	684.70	656.00	1293.40
1992-93	686.20	645.40	1337.70
1993-94	627.60	693.70	1431.10
1994-95	656.96	667.96	1572.90
1995-96	677.50	676.80	1680.00
1996-97	670.30	475.40	2104.15
1997-98	621.70	464.90	2181.90
1998-99	668.76	661.94	2111.74
1999-00	741.80	664.10	2290.98
2000-01	756.75	693.62	2437.39
2001-02	768.08	677.43	2632.44
2002-03	783.34	697.65	2764.78
2003-04	785.08	719.11	2994.61
2004-05	802.07	716.68	3027.40

Source: Handbook of agriculture Statistics December 2007, Bangladesh Bureau of Statistics (BBS).

Area irrigated by DTWs in Bangladesh is graphically displayed in Figure 2.3a.This figure shows that area irrigated by DTWs increased gradually with some fluctuation. Figure 2.3b indicates the area irrigated by STWs in Bangladesh. This figure shows the insignificant performances of STWs in the initial years. Area irrigated by LLPs in Bangladesh is represented in Figure 2.3c. As compared to DTWs and STWs, slow increasing tendencies of LLPs are found in this figure where quite fluctuations are also observed.

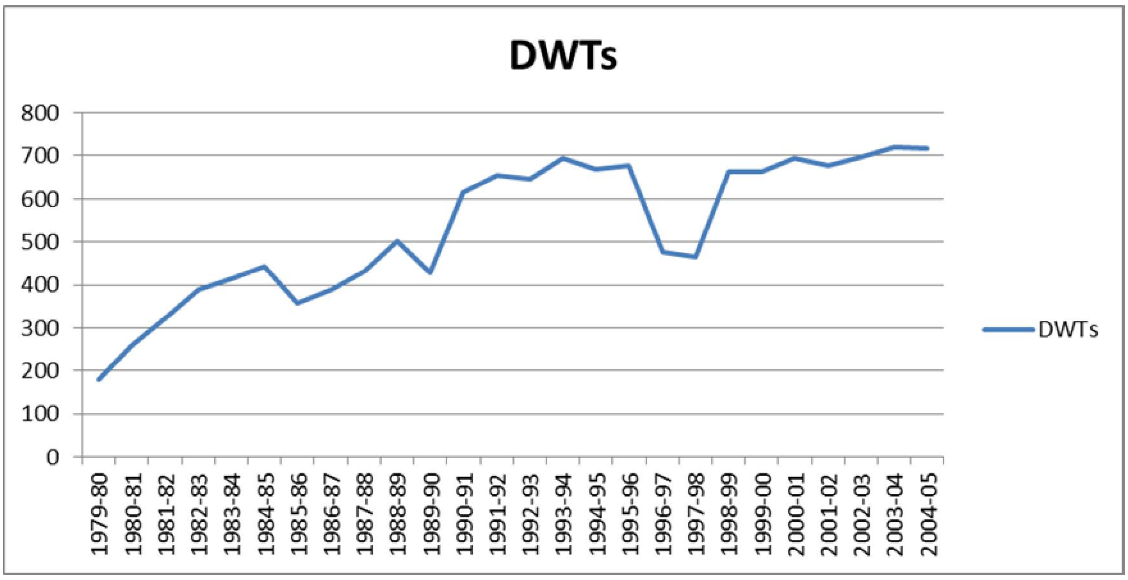


Figure 2.3a: Area irrigated by deep tube wells (DTWs) in Bangladesh from 1979-80 t0 2003-04

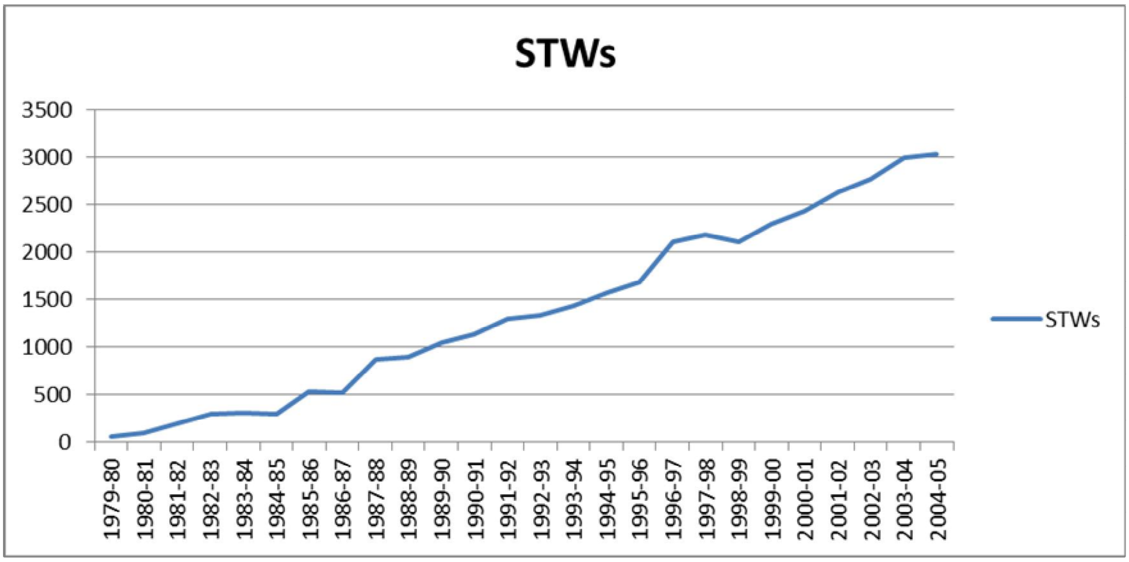


Figure 2.3b: Area irrigated by shallow tube wells (STWs) in Bangladesh from 1979-80 t0 2003-04

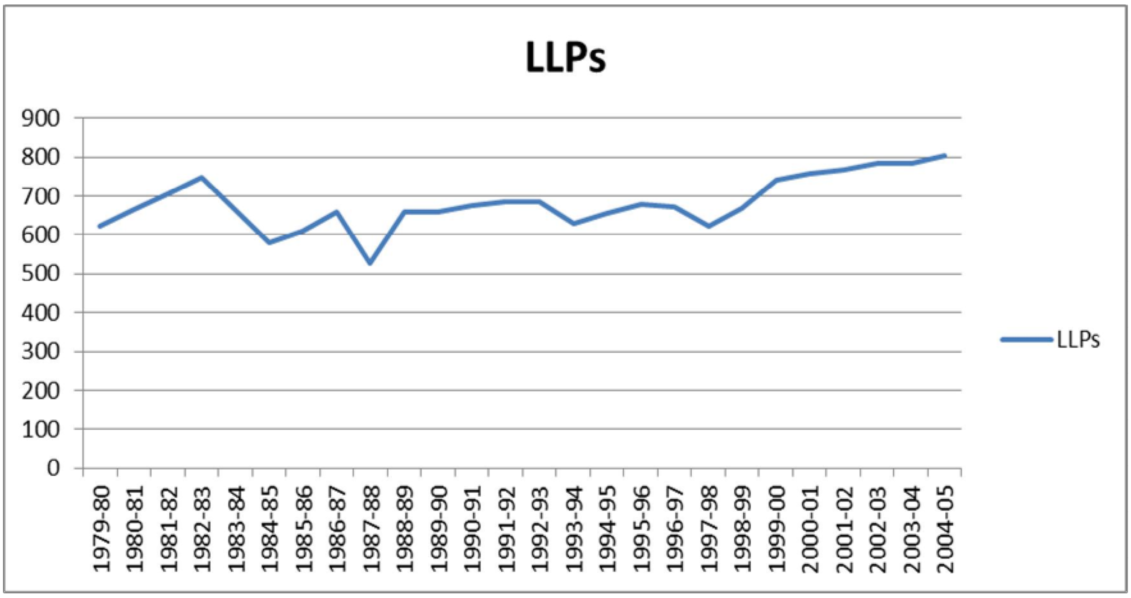


Figure 2.3c: Area Irrigated by Low Lift Pumps (LLPs) in Bangladesh from 1979-80 t0 2003-04

Area irrigated by different crops has also increased over the years. Table 2.9 presents the relevant trends of area irrigated by crops for the 1979-80 to 2004-05 years of the sample period. Table 2.9 shows that area under different crops of which rice is dominant and it has been gradually increasing over the years.

Table 2.9: Irrigated area under different crops from 1979-80 to 2004-05 (Figures in thousand hectares)

Year	Rice	Wheat	Potato	Vegetables	Others*	Total	Total Cropped Area	% of Irrigated Area
1979-80	1228.9	172.50	64.60	40.00	63.10	1569.10		
1980-81	1258.79	184.80	71.40	43.50	80.60	1639.09	13160.00	12.46
1981-82	1339.0	189.60	76.90	46.90	73.40	1725.80	13200.00	13.07
1982-83	1459.5	193.50	73.10	50.10	71.80	1848.00	13000.00	14.22
1983-84	1501.6	214.50	73.80	44.40	85.70	1920.00	13360.00	14.37
1984-85	1582.1	283.30	69.80	49.20	88.30	2072.70	13150.00	15.76
1985-86	1613.4	267.00	67.80	53.30	96.10	2097.60	13540.00	15.49
1986-87	1717.3	253.60	68.40	57.30	102.40	2199.00	13340.00	16.48
1987-88	1956.5	196.39	56.60	33.70	104.30	2347.49	13820.00	16.99
1988-89	2221	260.50	76.60	62.60	116.70	2737.40	13710.00	19.97
1989-90	2405.66	277.40	78.40	70.16	105.70	2937.32	14060.00	20.89
1990-91	2479.67	282.46	81.89	74.87	109.12	3028.01	14030.00	21.58
1991-92	2702.69	260.30	79.90	77.80	109.91	3230.60	13810.00	23.39
1992-93	2706.78	271.20	82.40	82.60	110.92	3253.90	13700.00	23.75
1993-94	2718.1	271.63	88.20	86.20	125.64	3289.77	13480.00	24.40
1994-95	2838.93	283.02	94.79	89.15	123.89	3429.78	13520.00	25.37
1995-96	2940.5	298.40	100.60	82.45	131.95	3553.90	13510.00	26.31
1996-97	3047.65	316.10	104.60	90.20	109.70	3668.25	13800.00	26.58
1997-98	3125.89	345.75	113.36	98.45	86.05	3769.50	14090.00	26.75
1998-99	3209.87	361.94	145.43	115.63	106.93	3939.80	13960.00	28.22
1999-00	3371.01	372.25	163.15	117.40	162.75	4186.56	14270.00	29.34
2000-01	3606.51	383.64	150.14	117.76	161.06	4419.11	14300.00	30.90
2001-02	3754.21	401.85	152.16	121.40	166.72	4596.34	14194.56	32.38
2002-03	3857.28	394.30	159.84	123.42	189.69	4724.53	14174.55	33.33
2003-04	3942.78	382.02	180.49	134.35	295.82	4935.46	14039.31	35.15
2004-05	4001.05	364.21	199.10	152.16	318.07	5034.59	14104.09	35.70

Source: Handbook of Agriculture Statistics December 2007, Bangladesh Bureau of Statistics (BBS).

2.4.5 Production technology

As rice is traditionally a subsistence sector in Bangladesh, it has been operating with Drought Animal Power (DAP) for various farm activities until early seventies. Traditional capital of rice farm consists of ploughs, bullock, and homemade irrigation. However, since the mid-seventies, agriculture of Bangladesh has been undergone some adoption to the new seed-fertilizer-irrigation technology (Hossain, 1995:19). Green revolution technology is the major technological breakthrough in agricultural history of Bangladesh. Over the last three decades (since early 1980s), the national policies had been directed towards transforming rice through rapid technological progress. Development programs have undertaken to diffuse modern varieties of rice and wheat with corresponding support in the provision of modern inputs, such as, chemical

*Others means others cereals, pulses, oil seeds, sugarcane, cotton and others crops.

fertilizer, pesticides, irrigation, equipment’s, institutional credit, product procurement facilities (Rahman and Karim, 1999).

Use of mechanical sources of power has gradually been increasing. Due to excessive sub-divisions and fragmentation of holdings, small farm size and capital constraints, complete mechanization is a time consisting matter in Bangladesh. Despite these constraints, it is realized that major portion of farmers in Bangladesh has mechanized their farming activities partially. They adopt and use different combinations of mechanical devices, such as mechanical irrigation, power tiller and tractor sprayers and thresh in their farming activities.

The rice sector of Bangladesh has made remarkable progress in the 20th century. Since independence in 1971, agriculture sector has been growing over time. There is tremendous technological development in the 20th century. Until 2000, Bangladesh releases more than 50 modern varieties of rice. Improved varieties technologies are complemented with better crop husbandry practices, efficient utilization of fertilizer and irrigation, and pest management methods. Modern varieties and other inputs, production technologies and knowledge are disseminating to farmers' field through extension services and other public and private agencies. Table 2.10 show changes over the period of 1983-84 to 1996-97 in the uses of fertilizer and irrigated area.

Table 2.10: Changes over the period of 1983-84 to 1996-97 in the uses of fertilizer and irrigated area. (hectare)

Variable	1983-84	1996-97	Change over 83-84 to 96
Irrigated area(hac)	1620938	3762514	132.12
Farm holdings use of fertilizer (m.ton)	6176100	9782685	58.40

Source : Agriculture Census, 1996.

As a result, total production level of food grain was increasing from 12.14 million metric tons in 1973-74 to 26.24 million in 1983-84. It leads to production, which is more than double in 2000-01 than that of 1973-74 (Deb et al. 2004).

2.5 Trends of rice production in Bangladesh

Since independence, technological transformation has accelerated with the available of modern techniques, inputs and equipment. Yet the modernization process is far from complete due mostly farmer`s lack of resource for investment, deficiency input delivery, and inadequate infrastructure. Recent performance in the agriculture sector is influenced by two major factors, natural phenomena, e.g. flood, drought, cyclone, etc. and policy changes particularly in case of input distribution and pricing system (GOB, 1990).

In Bangladesh, out of the net-cropped area of 7.60 million hectare, about 55 per cent is double cropped and approximately 15 per cent is triple cropped while about 30 per cent is still single cropped area (GOB, 1988). In 1968, modern variety (MV) rice was first introduced in Bangladesh, which needed irrigation facilities and application of fertilizer. In increasing the rice production, modern variety (MV) *boro* plays significant role in Bangladesh agriculture. Total rice production in Bangladesh is presented in Table 2.11.

Table 2.11 shows that the total rice production has gradually been increased overtime. During 1971-72 only 9774, thousand metric tons were rice production but in 2011-12, it gradually increased to 34750.00 thousand metric tons.

Table 2.11: Rice production in Bangladesh from 1971-72 to 2011-12. (Thousand metric tons)

Year	Aus	Aman	Boro	Total
1971-72	2341.00	5695.00	1738.00	9774.00
1972-73	2243.00	5587.00	2071.00	9901.00
1973-74	2801.00	6699.00	2220.00	11720.00
1974-75	2859.00	6000.00	2250.00	11109.00
1975-76	3229.00	7045.00	2286.00	12560.00
1976-77	3014.00	6905.00	1650.00	11569.00
1977-78	3103.00	7422.00	2239.00	12764.00
1978-79	3287.00	7429.00	1929.00	12645.00
1979-80	2809.00	7303.00	2427.00	12539.00
1980-81	3289.00	7964.00	2630.00	13883.00
1981-82	3270.00	7209.00	3152.00	13631.00
1982-83	3065.00	7516.00	3548.00	14129.00
1983-84	3222.00	7843.00	3350.00	14415.00
1984-85	2783.00	7930.00	3909.00	14622.00
1985-86	2828.00	8542.00	3671.00	15041.00
1986-87	3130.00	8267.00	4010.00	15407.00
1987-88	2993.00	7690.00	4731.00	15414.00
1988-89	2856.00	6857.00	5831.00	15544.00
1989-90	2475.00	9202.00	6033.00	17710.00
1990-91	2261.00	9167.00	6357.00	17785.00
1991-92	2179.00	9269.00	6807.00	18255.00
1992-93	2075.00	9680.00	6586.00	18341.00
1993-94	1850.20	9419.20	6772.20	18041.60
1994-95	1790.70	8504.00	6538.70	16833.40
1995-96	1676.00	8790.00	7220.60	17686.60
1996-97	1870.00	9551.00	7460.00	18881.00
1997-98	1874.60	8849.80	8137.30	18861.70
1998-99	1616.90	7735.80	10551.90	19904.60
1999-00	734.00	10306.00	12027.00	23067.00
2000-01	1916.00	11249.00	11920.50	25085.50
2001-02	1808.00	10726.00	11766.00	24300.00
2002-03	1850.70	11118.40	12222.20	25191.30
2003-04	1831.80	11520.50	12837.10	26189.40
2004-05	1500.00	9819.00	13837.10	25156.10
2005-06	1745.00	10810.00	13975.30	26530.30
2006-07	1512.00	10841.00	14965.00	27318.00
2007-08	1507.00	9662.00	17762.00	28931.00
2008-09	1895.00	11613.00	17809.00	31317.00
2009-10	1709.00	12207.00	18059.00	31975.00
2010-11	2132.82	12791.50	18616.00	33540.32
2011-12	2750.00	13300.00	18700.00	34750.00

Source: BBS and DAE, 2012.

2.6 Conclusion

This chapter contains a description of changing input use of rice production. A significant change in the use of major inputs is witnessed due to adoption of modern technology like fertilizer and irrigation in the rice sector of Bangladesh. As a core factor of production, total cropped area increased moderately during 1979-80 to 2001-02 (BBS, 2007). Labour force in the rice sector has decreased substantially due to the transfer of agriculture labour force to non-agriculture sector. Consumption of chemical fertilizers increased markedly and as the most modernized component, the use of irrigation increased during this period. In increasing the crop production, the MV *boro* plays an important role.

Chapter Three

Review of Literature

3.1 Introduction

In any relevant research work, there should be a clear and logical presentation of the literature review. The literature review mainly helps to find out or to highlight the research gap in the field of research work and provides the foundation for developing a conceptual framework. Besides this, literature review also helps the researcher to focus further research work more accurately and precisely.

3.2 Review of related literature

The relevant research works that are studied in the field of impact of education on rice production nationally and internationally are discussed below:

3.2.1 National

Asadullah and Rahman (2009) discover that the various stages of education have a very significant and positive effect on rice production in Bangladesh. It is found that the primary and secondary stage of education is more relevant in rice production than tertiary level. It is also found that the education is a significant matter in raising production, boosting potential output and improving farmer’s productive efficiency.

Salehin et al. (2009) finds that the educated farmers play a significant and positive effect on rice production in Bangladesh. It is found that educated farmers are likely to be more receptive to the modern ideas and technology; they have much mental strength in any decision over a matter related to the adoption of technologies as well as solving problems in their daily life. It is also found that the farmers, who have higher education likely to have higher adoption of rice production technologies.

Haq (2012) shows that primary education has a significant impact on rice productivity as it has a great positive value. It is found that farmers who have primary education seem to be effective for rising per unit of rice productivity in Bangladesh. It is also found that farmers with primary school degree might spend enough time for farm production.

Nargis and Lee (2013) find that education statistically plays a positive and significant role on rice production in Bangladesh. It is found that the more the farmers are educated, the more they are efficient compared to less-educated counterparts of them, probably because of their better skills, better access to information, and better farm planning.

Rahman and Haque (2013) find that the relationship between education and adoption of modern variety of wheat production practices in Bangladesh is very significant. It is found that education helps the individual to eliminate adoption gaps to increase productivity as well as net return of wheat production.

3.2.2 International

Gross and Tales (1952) examine that the educated farmers can differentiate themselves from uneducated ones in comparison with the acceptance of recommended farm practices. It is found that education helps a farmer greatly to solve any problem regarding rice production. It is also found that education encourages a farmer in taking risk and adoption new technology.

Nelson and Phelps (1966) show that farmers who have a relatively high level of education may have a higher probability of adopting new technologies than those with little education. It has also been examined that education enhances one’s ability to receive, decode, and understand information.

Welch (1970) finds that education directly affects agricultural production in three ways. He shows that "worker effect" improves the quality of labour and helps the farmer to produce more with a given bundle of non-labour inputs and "allocative effect" develops the farmer's ability to process information and to allocate inputs across competing uses. He also shows that "input selection effect" strengthens the selection of purchased inputs in the short run and the scale of operation in the long run. He also finds that the workers who are well educated are simply able to use a given amount of resources more efficiently.

Singh (1974) finds that literacy of the farm decision-maker has positive and significant role in determining the production level of the farm in India. He finds that the level of farm production is significantly higher on farms with literate decision makers than on farms with illiterate decision-makers. He finds that there is a positive relationship between the level of education of farm decision-maker and the level of farm production. He studies that the secondary education brings a qualitative improvement in managerial skills of farm decision-makers and hence marked increase in farm production. He also finds that the size of the farm has a significant influence on the relationship between education and farm production.

Huffman (1974) finds that the farmer's education has significant effect on agricultural production in United State. He observes that allocative efficiency is related to the education of decision makers. He also finds that the allocative effect of education is economically important in a dynamic environment, perhaps more important than the traditional worker effect of education.

Schultz (1975) finds that it is education which is expected to promote the quality of labour and mainly the impact depends on the environment. For example, the impact of education is higher in a rapidly changing technological or economic environment. Schultz also finds that the impact of education would be the strongest of all things in an upgrading and in a modernizing environment.

Wu (1977) finds that the way by which education contributes to production is quite different for small-scale, labour-intensive farming from for a large-scale, capital-intensive farming in Taiwan. He also finds that the minimum level of education is necessary for farmers to learn effective use of modern agricultural inputs.

Lockheed et al. (1980) describe that the gains of education under technological as well as modernizing conditions is considerably higher than that of a traditional environment.It is also found that schooling has a positive effect on rice production and the effectiveness of education is enhanced in modernizing environment.

Pudasaini (1983) examines that farmer's education raises agricultural production and productivity in Nepal in three ways. Firstly, it improves farmers' skills enabling them to achieve higher output for given inputs. Secondly, it enhances their ability to obtain, understand, and utilize new inputs and practices. Thirdly, it improves farmers' overall managerial ability. He also finds that education enhances agricultural production by improving farmers' decision-making ability and alleviating their technical efficiency.

Cotlear (1986) argues that education plays a vital role on agricultural production, but mostly this role depends upon its technical and economic situation. He finds that the completing primary education in urban areas and rural schools has different effect in nature. He also finds that the various levels of education are related with higher possibility of adoption but only in the initial stages of the distribution process.

Duraisamy (1989) studies that education has a positive and significant effect on rice production in India. He examines that education expands the probability of adoption of modernization of new techniques in rice production. He shows that the higher level of education is required to better understand, make out new information and utilise in an effective way. He also finds that the level of using high-yielding rice varieties in India is positively related to level of education.

Lin (1991) studies that farmer`s level of education has positive and statistically significant effects on the household's probability and intensity of adopting of hybrid seed in China. He also examines that farmer's education level is an important factor in the adoption of hybrid rice and this reason increasing state investments in rural education in order to facilitate technological change in agriculture.

Appleton & Balihuta (1996) examine that different stage of schooling of farmers are statistically significant and have a positive effect on agricultural production in Uganda. They find that the production of crops increases if the education level of farmers increases. They also find that with the help of education farmers can improve their practices and techniques. They also show that the use of controlling capital and purchasing inputs are affected by education.

Yang (1997) finds that the education has positive and significant effect upon the head as well as the earning members of the households in agricultural production in China. He finds that production is deeply associated and influenced by average schooling. He also finds that various errors in production function estimates occur due to the omission of the highest level of farm schooling, which finally results in mismeasurement of returns to education.

Taylor and Yunez-Naude (2000) find that education might affect an individual and his choice of economic activities as well as the income generated by the chosen activities. They also find that while schooling levels increase, the returns from schooling shift away from crop production.

Gallacher (2001) finds that education is a significant input in agricultural production. He shows that the absolute magnitude of returns to education depends not only on the geographical location of the firm but on the extent to which the firm is single- or multiple- output. He finds that returns to education obviously depend on firm size. He also shows that if education is held constant, decreasing returns to scale apply, whereas allowing education to vary results in increasing returns to all inputs including education.

Pritchett (2001) shows that the quality of education is just too low to increasecognitive abilities effectively and, ultimately, also productivity. He also finds out that the skills that are provided in formal education are too unspecific to affect agricultural production positively.

Psacharopoulos and Patrinos (2002) examine that wage returns to education vary by levels of schooling, and by gender, returns to primary education are higher for men than for women, and women have higher returns to secondary education than men have.

Knight et al. (2003) examines that a significant reduction of risk aversion if the household head had received at least some schooling. They also show that providing education to farmers not only lets them adopt new technologies earlier, but it may also change their attitude towards relatively risky traditional production technologies.

Dominique van de Walle (2003) studies the impact of education on agricultural productivity in Vietnam. Three major results come out of her study of irrigation and agricultural productivity in Vietnam. First, education of the household head and other family members make a significant contribution to farm profitability. Second, there also seems to exist important complementarities between education and irrigation, thereby giving some indication that education does help Vietnamese farmers make better use of agricultural technology, and third, primary education, but not higher levels of education, has significant impact on farm profitability. Years of schooling are found to have a significant impact on rice productivity, even though it is a small one.

Hassan et al. (2003) finds that education not only improves the mental alertness of an individual but also is responsible for the positive change in the behavior of an individual in all aspects of life. They also find that if the farmers are well educated formally or non-formally, they are able to handle their crop management and production, which leads to increase in overall production of various crops at national level in Pakistan.

Birdsall et al. (2004) suggests that the completion of 6 to 10 years of primary education may constitute critical thresholds for basic competencies such as literacy and numeracy to be acquired on a permanent basis. Moreover, once mastered, these skills may not be sustained over time if they are not required in one's daily environment**.**

Minh-Phuong Ngo (2006) show that the years of schooling are found to have a significant impact on rice productivity, with one additional year of education yielding an increase in rice yields of 1.3 percent in Vietnam. Since basic literacy and numeracy skills, which can only be acquired after completing several years of education, have a large and significant impact on farm efficiency, these low returns to years of schooling are probably due to mis-specifying linearity. In contrast with findings in the empirical literature, primary education is found to have no effect on rice yields. More provoking are findings that, after instrumentation, lower secondary education have no impact on rice yields, whilst large and significant impacts are found for upper secondary and higher education.

Onphandala (2009) finds that the role of farmers' education is quantitatively important in determining the agricultural production in Lao PDR. He finds that the estimated rates of return to schooling for both the upland and the lowland farming of Lao PDR are relatively high, particularly for the rates of return to a completion of primary cycle. He also finds that the adult literacy campaigns, including women's education with their important role in the agricultural labor force, would generate directly the modernization of agricultural practices and the improvements in well-being in the near future.

Yasmeen et al. (2011) finds that educated farmers can plan and cultivate in more efficiently than illiterate farmers, and so income level ultimately determines the change in production. They show that the educated farmers play a crucial role on agricultural production. They also find that the literate farmers are more likely to adopt the use of modern inputs than those who are illiterate.

Grigin (2011) finds that there is a statistically positive relationship between education and wheat output in Turkey. He also finds that there is a great chance that educated farmer contributed positively to agricultural productivity growth, which is just one of the intended aims of the education.

Rehaman et al. (2012) finds that education is one of the most important determinants of agricultural production. They have examined that education makes aware the producer about the latest production techniques, which enables him to increase crop production. They also show that education affects crop production positively in Pakistan. One percent increase in education enrollment leads to 4% increase crop production.

Reimers & Klasen (2012) find that education has a significant positive impact on agricultural productivity worldwide. They find the effect is sizeable, implying that an additional year of schooling for the whole population would increase agricultural productivity by approximately 3.2 percent. They also find that the effect of education is smaller in the poorest countries.

3.3 Research gap

A number of studies have assessed the relation between education and agricultural production (Wu, 1977; Lockheed et al.,1980 ; Jamison & Lau,1982; Philips, 1987; Hassan et al., 2003; Minh-Phuong, 2006; Onphanhdala,2009;Yasmeen et al.,2011; Girgin,2011; Rehaman et al.2012), another number of studies have assessed the impact of education on agricultural production (Singh,1974;Welch,1970;Pudasaini,1983) and another number of studies have assessed the impact of education on rice production (Asadullah & Rahman,2006; Salehin et al.,2009; Haq,2012; Nargis & Lee,2013; Duraisamy,1989) in national and international arena. Most of the studies included aggregate level of education, input cost, cultivable land, family labour and extension service as explanatory variables. But most of them do not include hired labour cost. The general forms of Cobb-Douglas production function are used in most of the studies. These studies have applied to examine the effect of education on agricultural production through classical linear regression model. They did not explain the pitfalls of their model regarding the impact of education on rice production. To have a clear picture of impact of education on rice production in Bangladesh, it is required to enquire deeply. Disaggregate level of education is used as explanatory variable rather than aggregate level of education in this study. The shortcomings of the classical linear regression model have been discussed systematically in this study. As a result, the findings would provide more reliable than other study. That is why, this study demands greater importance in their arena. In this study, the ridge regression has been also applied to explore the impact of education on rice production. This is a scope to do research in this area to fill this gap. To our knowledge, this research is the first of its kind in Bangladesh.

Chapter Four

Methodology of the Study

4.1 Introduction

Research methodology is the philosophy of a research to solve the research problem systematically. It provides an understanding of how research has been conducted and organized in order to obtain information. It also provides the characteristics of the research development, explains the methods to obtain information from respondents and describes how data will be collected and processed. This chapter covers the economic model of impact of education on rice production, sampling technique, methods of data collection, technique and analysis of data and research design.

4.2 Selection of the study area

Shibganj *Upazila* of Bogra district has been purposively selected as the study area for the study. The study has been conducted to determine the impact of education on rice production in three villages namely Chapachil, Paschim Saidpur and Asrafpur in Shibgonj *Upazila* under Bogra district. The Shibgonj *Upazila* comprises of 409 villages (Population census, 2011). The villagers primarily rely on agriculture activities. Therefore, their earnings depend on agricultural activities. Rice is the main agricultural crop in this *Upazila*. That is why; Shibganj has been selected for the study. Comparatively the rice production of this *Upazila* is higher than the other *Upazila* of Bogra district (Bangladesh Bureau of Statistics, Bogra Branch, 2012).

4.3 Methods of data collection

The study is based on primary and secondary data. The primary data have been collected by using a structured questionnaire. Before preparing and applying the questionnaire to the final survey, pre pilot and pilot survey have been done. The pre pilot survey is carried out through the Agricultural Office of Shibganj, concerned Sub-Assistant Agriculture Officer`s (SAAO's), and academics.The pilot survey has been conducted during November 2012 to December 2012. Afterwards, the final survey has been carried out during December 2012 to January 2013.

Secondary data have also been collected from related books, articles, Journals, Unpublished thesis, Population Census of Bangladesh, various issues of Economic Review, Agriculture census, Bangladesh Rice Research Institute (BRRI), Department of Agriculture Extension (DAE), Bangladesh Bureau of Statistics (BBS), Bangladesh Bureau of Statistics Bogra Branch, Ministry of Agriculture, Ministry of Planning, and Internet Sources etc.

4.4 Questionnaire

Three villages in Shibganj *Upazila* of Bogra district are selected purposively. Mainly *aman* and *boro* are cultivated in this area. The data is obtained using a structured questionnaire from face to face interview. The questionnaire is structured in English and translated into Bengali. The questionnaire consists of three major sections. The first section contains personal and socio-economic information. The second section contains the cost and production in this season. The third section includes about extension service.

4.5 Sampling technique of the study

An up to date list of all farmers of the selected villages has been collected from *Upazila* Agriculture Office. The list comprised 551 farmers, which constitute the population. In this study, random sampling technique is employed to collect the data. The numbers of farm household are selected randomly from each village by using determination of sampling formula (Krejcie and Morgan, 1970) for regression analysis. Thus, the sample size for Chapachil is 171, Paschim Saidpur is 96 and Asrafpur is 91 respectively. In this study, the landless farmers have not been considered to show the impact of education on rice production. The determination of sampling technique formula is as follows:

$$s = \frac{\chi^2 NP(1-P)}{d^2(N-1) + \chi^2 P(1-P)}$$

where, s = required sample size, χ^2 = the table value of chi-square for 1 degree of freedom at the desired confidence level(3.841), N = the population size,

P = the population proportion (assumed to be 0.50 since this would provide the maximum sample size), d = the degree of accuracy expressed as a proportion (0.05).

Table 4.1: Distribution of population and sample respondents in three selected villages at Shibganj *Upazila* of Bogra

Name of the village	Total number of farmer	Sample farmer	Percentage (%)
Chapachil	306	171	55%
Paschim Saidpur	127	96	75%
Asrafpur	118	91	77%
Total	551	358	65%

Source: Field survey, December 2012 to January 2013.

4.6 Techniques of data processing

After the collection of data, the next step is data processing. Processing of data involves editing, coding, classification and tabulation.

4.6.1 Editing

After collection of primary data through structural questionnaire, the researcher edited the data to minimize non-sampling errors and to increase the accuracy and consistency of this collected data set.

4.6.2 Coding

It refers the process of assigning of symbols to each response of a category. After completion of editing of data, the researcher himself coded the data efficiently.

4.6.3 Classification

In order to get meaningful relationships among the data the researcher classified the data according to its attributes or based on class intervals.

4.6.4 Tabulation

After completion of editing, coding and classification, the data is organized in tabular forms. Classified and tabulated data has also been presented in the forms of charts and diagrams.

4.7 Analysis of data

The next step is data analysis and the analysis is divided into two major branches namely descriptive analysis and inferential analysis.

4.7.1 Descriptive analysis

Descriptive analysis is used to organizing, summarizing and describing the data or measuring the relationship between two or more variables. Descriptive statistics is also used to analysing the data for frequency, means, standard deviation etc.

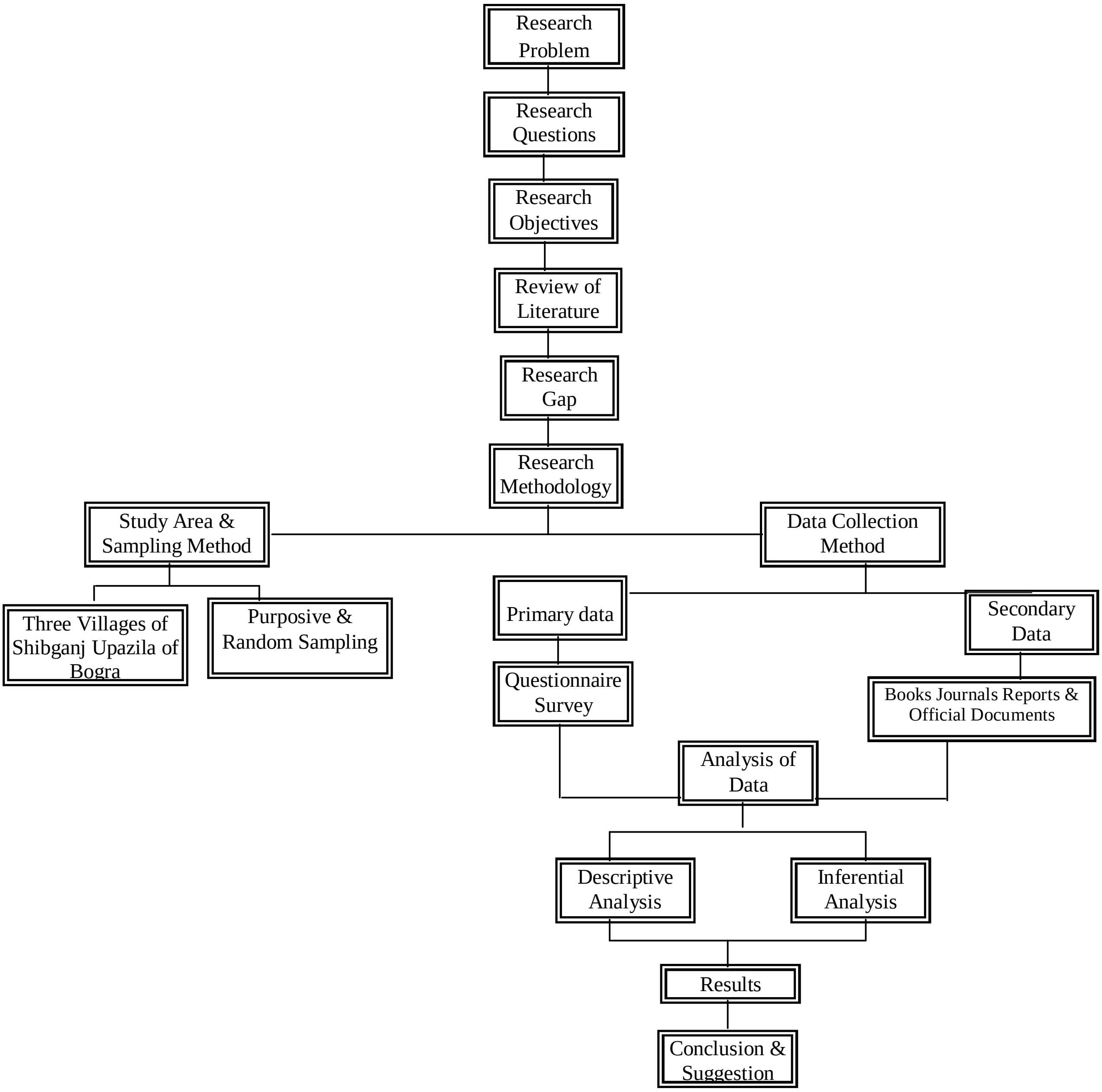
4.7.2 Inferential analysis

Inferential analyses are used to interpreting and to generalizing the findings from sample data analysis. Inferential statistics is also concerned with estimation of parameter and test of hypotheses. For descriptive and inferential analyses, Statistical Package for Social Science (SPSS-11.5 Version), Mathematica (5.0.1 version), Statistica (5.0 version), Eviews (5.0 version), MS Excel, and MS word have been used in this study.

4.8 Research design

A research design is a plan of the proposed research work of the blueprint of the proposed research work. This research consists of four sections. The section one is identification of the problems, and formulation of research questions and objectives strengthening by the literature review. The second section is the collection of primary data and secondary data. The third section is the processing and analyzing of data. The fourth section is conclusion and suggestions.

Research Design



4.9 Reliability and validity

Validity and reliability are two fundamental elements in the evaluation of a measurement instrument. Instruments can be conventional knowledge and survey questionnaires. Reliability is a crucial notion by which questionnaires can be evaluated whether the questionnaires will suffer from measurement error or not. Reliability estimates show the amount of measurement error in a test (Tavakol and Dennick, 2011). Reliability of the questionnaire has been achieved by internal consistency. Internal consistency is measured by employing Cronbach's Alpha test. Cronbach's alpha is a statistic. It is generally used as a measure of internal consistency or reliability of a psychometric instrument. Cronbach`s alpha is defined as (Lopez, 2007):

$$\alpha = \left(\frac{k}{k-1}\right)\left\{1-\left(\frac{\sum_{i}^{k}\operatorname{var}(S_i)}{\operatorname{var}\left(\sum_{i}^{k} S_i\right)}\right)\right\}$$

Where, *k* is the number of items in the instrument and S_i represents the score for item *i*. Cronbach's alpha reliability coefficient normally ranges between 0 and 1. However, actually there is no lower limit to the coefficient. George and Mallery (2003) provide the following rules of thumb

If α > 0.9 → Excellent,

α > 0 .8 → Good,

α > 0 .7→Acceptable,

α > 0.6 → Questionable,

α > 0.5 →Poor and

α < 0 .5 → Unacceptable.

The internal consistency of the questionnaire has also obtained by asking respondents questions more than once during a face-to-face interview (Ali and Noman, 2013). Validity is the accuracy and meaningfulness of inferences, which are based on the research results. Validity of questionnaire is achieved by content validity.

4.10 Empirical theory and method

In this study, the standard method of analysis follows Jamison and Lau (1982). They have used a production function for agriculture output as their basic tool to analyse the effect of education on crops production. They include various explanatory variables in their model particularly area under cultivation, labour input (family labour), education level of household head and extension service. The input cost and hired labour cost for crops production are not included in their model. Input and hired labour is a vital ingredient in any stage of production. In this aspect, we have modified their model through including input cost and hired labour cost. The Cobb-Douglas type production function is used in this study.

$$Y = AK_i^{\beta_1} L_i^{\beta_2} T_i^{\beta_3} e^{\beta_4 D_1 + \beta_5 D_2 + \beta_6 D_3 + \beta_7 D_4 + DExt + \mu_i} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$$

Equation (1) provides nonlinear relationship between output and inputs. So, the nonlinear relationship can be linearized by both side natural logarithms (ln).

$$\ln Y = \ln A + \beta_1 \ln K_i + \beta_2 \ln L_i + \beta_3 \ln T_i + (\beta_4 D_1 + \beta_5 D_2 + \beta_6 D_3 + \beta_7 D_4 + DExt + \mu_i) \ln e$$

$$= \beta_0 + \beta_1 \ln K_i + \beta_2 \ln L_i + \beta_3 \ln T_i + \beta_4 D_1 + \beta_5 D_2 + \beta_6 D_3 + \beta_7 D_4 + DExt + \mu_i \ldots\ldots (2)$$

Where, $\ln A = \beta_0$ and lne =1.

Thus, the model is linear in the parameters $\beta_0, \beta_1, \beta_2, \beta_3, \beta_{4,} \beta_5, \beta_6$ and β_7. So, the model is a linear regression model. So, the fitted model of this study is as follows

$$\ln Y = \beta_0 + \beta_1 \ln K_i + \beta_2 \ln L_i + \beta_3 \ln T_i + \beta_4 D_1 + \beta_5 D_2 + \beta_6 D_3 + \beta_7 D_4 + DExt + \mu_i \ldots\ldots\ldots (3)$$

where, Y_i = total output of rice (kilogram), K_i = input cost (irrigation and others input cost),

L_i = labour cost (family labour and hired labour), T_i = cultivable land (decimal),

D_1 = 1primary education of the farmer (years of schooling)

= 0 otherwise

D_2 = 1secondary education of the farmer (years of schooling)

= 0 otherwise

D_3 =1 higher secondary education of the farmer (years of schooling)

= 0 otherwise

D_4 =1 tertiary education of the farmer (years of schooling)

= 0 otherwise

Ext = extension service (from friends/Neighbours /agriculture officers/others)

D =1 if taken extension service

D = 0 otherwise

μ_i= error term

The error term is assumed random and serially independent having zero mean with finite variance. In order to determine the appropriate technique of estimation, the empirical model is estimated by the ordinary least squares (OLS) method. To better facilitate different diagnostic tests like heteroscedasticity, autocorrelation, multicollinearity are checked in this study.

4.11 Definition of the variables and research hypothesis

Output

Output is defined as the physical output of rice per decimal whereas physical output is defined as the total production of rice cultivated area that is expressed in term of kilogram per decimal.

Input cost

Input cost consists of seeds, seedbed preparation, plough units, irrigation, organic and inorganic fertilizers, insecticides, fungicides, herbicides, harvesting and threshing cost.

Null hypothesis H_0: There is no relation between input cost and rice production.

Alternative hypothesis H_1: There is a relation between input cost and rice production.

Labour cost

Labour cost is measured in man-days of eight hours. There are two types of labour cost in rice production such as hired labour cost and family labour cost.

Null hypothesis H_0: There is no relation between labour cost and rice production.

Alternative hypothesis H_1: There is a relation between labour cost and rice production.

Cultivable land

Cultivable land that is used by ploughing, sowing, and raising crops is expressed as decimal.

Null hypothesis H_0: There is no relation between cultivable land and rice production.

Alternative hypothesis H_1: There is a relation between cultivable land and rice production.

Education

Year of schooling may be represented as a level of education. It is defined as the number of academic years that a person has taken his lesson in educational institutions in this study. Level of education can be divided into five categories. These are illiterate, primary, secondary, higher secondary and tertiary.

Null hypothesis H_0: There is no relation between education and rice production.

Alternative hypothesis H_1: There is a relation between education and rice production.

Illiterate

People who can neither read nor write can be defined as illiterate. Illiterate also refers to someone who has not had any form of formal education at all.

Null hypothesis H_0: There is no relation between illiterate person and rice production.

Alternative hypothesis H_1: There is a relation between illiterate person and rice production.

Primary education

Person who obtains primary education consisting of five years of schooling from a formal school is called primary educated person.

Null hypothesis H_0: There is no relation between primary education and rice production.

Alternative hypothesis H_1: There is a relation between primary education and rice production.

Secondary education

The secondary level of education comprises five years of formal schooling.

Null hypothesis H_0: There is no relation between secondary education and rice production.

Alternative hypothesis H_1: There is a relation between secondary education and rice production.

Higher secondary

The higher secondary level of education comprises of two years of formal education.

Null hypothesis H_0: There is no relation between higher secondary education and rice production.

Alternative hypothesis H_1: There is a relation between higher secondary education and rice production.

Tertiary education

Tertiary education is defined as people who hold education more than higher secondary level.

Null hypothesis H_0: There is no relation between tertiary education and rice production.

Alternative hypothesis H_1: There is a relation between tertiary education and rice production.

Extension service

The contact between agricultural extension agents or officers as well as farmers is introduced as a measure of the availability of information about new and improved inputs.

Null hypothesis H_0: There is no relation between extension service and rice production.

Alternative hypothesis H_1: There is a relation between extension service and rice production.

4.12 *Chi-square* test (χ^2)

The *Chi-square test* is often used to determine whether the variables are statistically independent or not if they are associated. The *Chi-square* test determines if there is dependence (association) between the two classification variables. A *cross tabulation* is a joint frequency distribution of cases based on two or more categorical variables. Displaying a distribution of cases by their values on two or more variables is known as contingency table analysis. The joint frequency distribution can be analyzed with the *chi square* statistic to determine whether the variables are statistically independent or if they are associated. The chi-square statistic compares the observed count in each table cell to the count which would be expected under the assumption of no association between the row and column classifications. The chi-square statistic may be used to test the hypothesis of no association between two or more groups, populations, or criteria.

4.13 Analysis of variance (ANOVA)

The ANOVA is a very powerful technique of statistical analysis. The term ANOVA describes a technique whereby the total variation embedded in data is being analyzed or divided into meaningful components due to independent causes. The purpose of ANOVA is to test for significant differences between means of the groups (i.e., between the group and within the group).

4.14 Regression analysis

As the main objective of this study is to assess the impact of education on rice production, the cause effect analysis is suitable for achieving this objective. In doing so, regression analysis has been applied in this study. Regression analysis has become one of the most widely used statistical tools for analysing multifactor data. It is appealing because it provides a conceptually simple method for investigating functional relationship among variables.

4.14.1 Types of regression

Regressions are of two types. These are

i) Simple regression

(ii) Multiple regressions

i) **Simple regression**

A statistical analysis utilizes one quantitative independent variable to predict the quantitative dependent variable. When one independent variable is used in a regression, it is called a simple regression.

In a simple linear regression model $Y = \beta_0 + \beta_1 X + \mu$, β_0 is the regression interception. The estimate β_0 determines the level of the fitted line. i.e., it indicates the distance of the line directly above or below the origin. β_1 is called the regression coefficient. The slope of the line measured by the estimates of β_1 that gives the average amount of change of Y_i per unit change in the values of X_i. The sign of estimate of β_1 indicates the type of relationship between Y_i and X_i. μ is a stochastic error term. Y_i is the dependent variable and X_i is independent variable.

Let the linear regression model is

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots\dots\dots + \beta_{k-1} X_{k-1i} + \mu_i \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \text{(a)}$$

In the above model Y_i and X_{ji} (i= 1,2,..,n and j=1,2,…,k-1) are the dependent and predictor or explanatory variables respectively. μ_i's are stochastic disturbance term. β_0 and β_j are the intercept term and regression coefficient respectively.

ii) **Multiple regressions**

Multiple regression analysis is a mathematical measure of the average relationship between two or more predictor variables. In multiple regression analysis there are two types of variables one is explained variable or predicted variable or dependent variable and the other is call independent variable.When two or more independent variables are used, it is called a multiple regression.

The general purpose of multiple regressions is to learn more about the relationship between several independent or predictor variables and a dependent or criterion variable. Multiple regression analysis is viewed here as iterative processes, a process in which the outputs are used to diagnosing, validating, criticizing and possibly modifying the inputs.

The data consist of n observations on a dependent or response variable Y_i and p predictor or explanatory variables, X_1, X_2, ………,Xp. The relationship between Y and X_1, X_2… Xp is formulated as a linear model

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots\ldots + \beta_p X_p + \mu \ldots\ldots\ldots\ldots (b)$$

Where, Y_i = dependent variable, X's = independent variables, β's = regression coefficient and μ is a random disturbance or error term.

4.14.2 Diagnosis of regression

Before analyzing the regression results, the classical linear regression assumptions are discussed sequentially.

4.15 Tests for normality of the regression model

Normality tests are used to determine whether a data set is well modeled by a normal distribution or not. Normal distribution has a unique place in the theoretical and applied statistics. The assumption that variables like input cost, labour cost, and cultivable land, five types of level of education and extension service follow normal distribution occurs repeatedly in statistical tests of significance. Consequences of violating the assumption vary from relatively severe for inferences on variables. Techniques for assessing the truth of the assumption that variables follow normal distribution are provided below.

The characteristics of a normal distribution are:

i) the distribution is symmetrical, i.e., mean = median = mode and the coefficient of skewness is equal to zero.

ii) the coefficient of kurtosis is equal to three.

There are several methods of assessing whether data are normally distributed or not. They fail into two board categories: graphical and statistical (Ghasemi and Zahediasl, 2012). The most common are:

(a) Graphical

(i) Quantile-Quantile (Q-Q) plots

(ii) Probability- Probability (P-P) plots

(b) Statistical

(i) Kolmogorov-Smirnov test

(ii) Lilliefors corrected K-S test

(iii) Shapiro-Wilks test

(iv)Anderson-Darling test

(v) D`Agostino-Pearson test

(vi) D`Agostino skewness test

(vii) Jarque-Bera test

4.15.1 Jarque–Bera (JB) test of normality:

The JB test of normality is an asymptotic or large-sample test. It is also based on the OLS residuals. This test first computes the skewness and kurtosis measures of the OLS residuals and uses the following test statistic:

$$JB = n\left[\frac{S^2}{6} + \frac{(K-3)^2}{24}\right]$$

where *n* = sample size, *S* = skewness coefficient, and *K* = kurtosis coefficient.

For a normally distributed variable, *S* = 0 and *K* = 3. Therefore, the JB test of normality is a test of the joint hypothesis that *S* and *K* are 0 and 3, respectively. In that case the value of the JB statistic is expected to be 0. Under the null hypothesis that the residuals are normally distributed and alternative hypothesis that the residuals are not normally distributed, Jarque and Bera showed that asymptotically (i.e., in large samples) the JB statistic follows the *chi-square* distribution with 2 df. If the computed *p* value of the JB statistic in an application is sufficiently low, which will happen if the value of the statistic is very different from 0, one can reject the hypothesis that the residuals are normally distributed. But if the *p* value is reasonably high, which will happen if the value of the statistic is close to zero, we do not reject the normality assumption. If the *p*-value of Jarque-Bera statistics is less than 5 percent (0.05) we can reject null and accept the alternative that is residuals (u) are not normally distributed.

4.16 Heteroscedaticity

One of the important assumptions of the classical linear regression model is that the variance of each disturbance term u_i , conditional on the chosen values of the explanatory variables, is some constant number equal to σ^2. This is the assumption of homoscedasticity**,** or *equal* (homo) *spread* (scedasticity), that is, *equal variance*. Symbolically,

$E(u_i^2) = \sigma^2$, i=1,2,......n

Heteroscedasticity is a statistical term used to describe the behavior of a sample's variance and standard deviation. The variance of the error term is constant (Homoscedasticity). If the error terms do not have constant variance, they are called to be heteroscedastic. It occurs once the variance of the error terms differ across observations. The problem of heteroscedasticity is likely to be more common in cross-sectional than in time series data. In cross-sectional data usually deals with members of a population at a given point in time. Heteroscedasticity does not affect the parameter estimates. The problem with heteroscedasticity is that *t*-statistics does not trusted because our estimates of the standard errors are biased.

4.16.1 Remedies of heteroscedaticity

In general, there are two solutions if the model suffers from heteroscedasticity.

(i) If E (μ_i^2) = σ_i^2 then heteroscedasticity is present. Given values of heteroscedasticity can be corrected by using weighted least squares (WLS) as a special case of generalized least squares (GLS). Weighted least squares are the OLS method of estimation applied to the transformed model.

(ii) If σ_i^2 is not known a priori, then heteroscedasticity is corrected by hypothesizing a relationship between the error variance and one of the explanatory variables.

4.16.2 White heteroscedasticity test

The White Test is a test for detection heteroscedasticity in OLS residuals. In other word, this test can be a test of heteroscedasticity or specification error or both. It has been argued that if no cross-product terms are present in the White Test procedure, then it is a test of pure heteroscedasticity. If cross-product terms are present, then it is a test of both heteroscedasticity and specification bias. The null hypothesis of the White Test is that there is no heteroscedasticity. The test statistic is computed by an auxiliary regression of the squared residuals on all possible cross products of the regressors. The number of observations times the R^2 from the test regression is used to compute the White Test statistic.

4.17 Autocorrelation

Autocorrelation may be defined as correlation between members of series of observations ordered in time (as in time series data) or space (as in cross-sectional data). In the regression context, the classical linear regression model assumes that such autocorrelation does not exist in the disturbances u_i. Symbolically,

$$E(u_i u_j) = 0, \quad i \neq j$$

Put simply, the classical model assumes that the disturbance term relating to any observation is not influenced by the disturbance term relating to any other observation.

4.17.1 Remedies of autocorrelation

If the model suffers from autocorrelation, then the OLS estimators will not be efficient. There are three solutions.

(i) to change the specification of the model, the error term is not auto- correlated.

(ii) to use an estimator that accounts for autocorrelation, e.g. Nonlinear Least Squares or Cochrane - Orcutt iterative method.

(iii) to estimate the parameters of the model with the OLS and account for autocorrelation while calculating standard errors.

4.17.2 Breusch-Godfrey serial correlation (Lagrange multiplier-LM) 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 allows for (1) non-stochastic regressors, such as the lagged values of the regress and; (2) higher-order autoregressive schemes., the BG test which is also known as the LM test.

In ordinary least squares (OLS) regression, time series residuals are often found to be serially correlated with their own lagged values. Serial correlation means (a) OLS is no longer an efficient linear estimator, (b) standard errors are incorrect and generally overstated, and (c) OLS estimates are biased and inconsistent if, a lagged dependent variable is used as a regressor. This test is an alternative to the Q-Statistic for testing for serial correlation. It is available for residuals from OLS, and the original regression may include autoregressive (AR) terms. Unlike the Durbin-Watson Test, the Breusch-Godfrey Test may be used to test for serial correlation beyond the first order, and is valid in the presence of lagged dependent variables. The null hypothesis of the Breusch-Godfrey Test is that there is no serial correlation up to the specified number of lags. The Breusch-Godfrey Test regresses the residuals on the original regressors and lagged residuals up to the specified lag order. The number of observations multiplied by R^2 is the Breusch-Godfrey Test statistic.

4.17.3 Durbin–Watson statistic

The most celebrated test for detecting serial correlation is developed by statisticians Durbin and Watson. It is popularly known as the Durbin–Watson *d* statistic. In statistics, the Durbin–Watson statistic is a test statistic used to detect the presence of autocorrelation (a relationship between values separated from each other by a given time lag) in the residuals (prediction errors) from a regression analysis. It is named after James Durbin and Geoffrey Watson. However, John von Neumann (Von Neumann, 1941) derived the small sample distribution of this ratio in a path-breaking article. Durbin and Watson (1950, 1951) applied this statistic to the residuals from least squares regressions, and developed bounds tests for the null hypothesis that the errors are serially independent (not autocorrelated) against the alternative that they follow a first order autoregressive process. Later, John Denis Sargan and Alok Bhargava developed several von Neumann-Durbin-Watson type test statistics for the null hypothesis that the errors on a regression model follow a process with a unit root against the alternative hypothesis that the errors follow a stationary first order autoregression (Sargan and Bhargava, 1983).

4.18 Multicollinearity

In regression analysis, there is an important assumption is that the explanatory variables are independent to each other i.e. there is no relationship between the explanatory variables. However, in most applications of regression, the explanatory variables are related to each other. This problem is called multicollinearity problem. In other words, it can be said that multicollinearity means the existence of a perfect or exact linear relationship among some or all explanatory variables of regression model. Collinearity refers to the existence of a single linear relationship between/ among the variables.

4.18.1 Sources of multicollinearity

There are several sources of multicollinearity. Multicollinearity may be due to the following factors:

i) the data collection method employed

ii) constraints on the model or in the population being sampled

iii) model specification

iv) an over defined model.

4.18.2 Types of multicollinearity

There are two types of multicollinearity. These are

i) perfect multicollinearity and

ii) near multicollinearity.

i) **Perfect multicollinearity**

If the explanatory variables of a regression model are exactly linearly related, this situation is defined as the problem of exact multicollinearity.

ii) **Near multicollinearity**

If the explanatory variables are nearly linearly related, then this situation is defined as the problem of near multicollinearity.

4.18.3 Techniques for detecting multicollinearity

There are many techniques to detect multicollinearity. In this study, Variance Inflation Factor (VIF) has been used to detect multicollinearity.

I) Examination of correlation matrix (ECM),

II) Variance inflation factor (VIF),

III) Eigenvalues and condition number (ECN) and

IV) Eigen values decomposition (EVD).

4.18.4 Variance inflation factor (VIF)

The variance inflation (inflating) factor (VIF) is often used to test the extent of multicollinearity. It is defined as:

$$VIF_i = \frac{1}{1-R_i^2} = \frac{1}{TOL_i}$$

where R_i^2 is the R^2 in the regression of x_i on all the other independent variables (Greene, 2003). VIF shows how the presence of multicollinearity inflates the variance of an estimator. As R_i^2 approaches 1, VIF approaches infinity. In the absence of any multicollinearity, R_i^2 will be close to zero and VIF will approach unity. The inverse of VIF is called tolerance (TOL). If VIF is very high TOL will be very low. In the literature, VIF and TOL are used interchangeably (Gujarati 2003). As a rule of thumb, multicollinearity may not be a serious issue if VIF does not exceed 5. If The VIF is greater than about 5 it indicates potential multicollinearity problem (Rogerson, 2003). Despite the criticism against VIF it is very widely used as tool of detecting multicollinearity.

4.18.5 Remedies of multicollinearity

Multicollinearity is a matter of degree, not a matter of presence or absence. The higher the degree of multicollinearity, the greater the likelihood is of the disturbing consequences of multicollinearity. Various methods have been developed to cope with multicollinearity problems (Gujarati, 2003 and Adnan et al., 2006) among such methods are

a) a priori information

b) combining cross-section and time series data

c) dropping a variable and specification bias

d) transformation of variables

f) additional or new data

g) reducing collinearity in polynomial regressions

h) factor analysis

i) ridge regression,

j) principal component regression,

k) partial least squares regression,

4.18.6 Ridge regression

Ridge regression provides another alternative estimation method that may be used to advantage when the predictor variables are highly intercorrelated. There are a number of alternative ways to define and compute ridge estimates. Ridge estimates of the regression coefficients may be obtained by solving a slightly altered form of the normal equations. Assume that the standardized form of the regression model is given as:

$$Y = \beta_0 X_1 + \beta_1 X_2 + \beta_2 X_3 + \cdots\cdots\cdots + \beta_n X_n + \mu \cdots\cdots\cdots\cdots \text{..................................(i)}$$

The estimating equations for the ridge regression coefficients are

$$\begin{array}{l} (1+k)\beta_1 + r_{12}\beta_2 + \ldots + r_{1n}\beta_n = r_{1y} \\ r_{21}\beta_1 + (1+k)\beta_2 + \ldots + r_{2n}\beta_n = r_{2y} \\ \quad . \qquad . \qquad . \qquad . \\ \quad . \qquad . \qquad . \qquad . \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\text{(ii)} \\ \quad . \qquad . \qquad . \qquad . \\ r_{n1}\beta_1 + r_{n2}\beta_2 + \ldots + (1+k)\beta_n = r_{ny} \end{array}$$

Where r_{ij} is the correlation between the i^{th} and j^{th} predictor variables and r_{iy} is the correlation between the i^{th} predictor variable and the response variable Y, the solution to (ii), $\hat{\beta}_1, \hat{\beta}_2, \ldots\ldots\ldots\ldots\ldots, \hat{\beta}_n$ is the set of estimated ridge regression coefficients.

The essential parameter that distinguishes ridge regression from OLS is *k*. Note that when $k = 0$, the $\hat{\beta}$'s are the OLS estimates. The parameter *k* may be referred to as the bias parameter. As *k* increases from zero, bias of the estimates increases.

The ridge analysis is to select a value of *k* and to obtain the corresponding estimates of the regression coefficients. If multicollinearity is a serious problem, the ridge estimators will vary dramatically as k is slowly increased from zero. Several methods have been suggested for the choice of k. These methods include:

1. Fixed point

Hoerl, Kennard, and Baldwin (1975) suggest estimating k by

$$k = \frac{n\hat{\sigma}^2(0)}{\sum_{j=1}^{n}[\hat{\beta}_j(0)]^2} \quad \text{(iii)}$$

Where $\hat{\beta}_1(0)$, $\hat{\beta}_2(0)$.......... $\hat{\beta}_n(0)$ are the least squares estimates of $\beta_1, \beta_2, ..., \beta_n$ when the model in equation (i) is fitted to the data (i.e., when k = 0), and $\hat{\sigma}^2(0)$ is the corresponding residual mean square.

2. Iterative method

Hoerl and Kennard (1976) propose the following iterative procedure for selecting k. Start with the initial estimate of k in equation (iii). Denote this value by k_0. Then calculate

$$k_1 = \frac{n\hat{\sigma}^2(0)}{\sum_{j=1}^{n}[\hat{\beta}_j(k_0)]^2} \quad \text{(iv)}$$

Then use k_1 to calculate k_2 as

$$k_2 = \frac{n\hat{\sigma}^2(0)}{\sum_{j=1}^{n}[\hat{\beta}^2(k_1)]^2} \quad \text{(v)}$$

Repeat this process until the difference between two successive estimates of k is negligible.

3. Ridge trace

The behavior of $\hat{\beta}_j(k)$ as a function of k is easily observed from the ridge trace. The value of k selected is the smallest value for which all the coefficients $\hat{\beta}_j(k)$ are stable.

4.18.7 Ridge regression estimators

Hoerl and Kennard (1970) suggest a class of estimators indexed by a parameter k > 0. The estimator is (for a given value of k)

$$\hat{\beta}(k) = (X'X + kI)^{-1} X'Y = (X'X + kI)^{-1} X'X\hat{\beta} \text{ .... (i)}$$

The expected value of $\hat{\beta}(k)$ is

$$\text{E} [\, \hat{\beta}(k) \,] = (X'X + kI)^{-1} X'X\beta \text{ .... (ii)}$$

and the variance – covariance matrix is

$$\text{V} [\, \hat{\beta}(k) \,] = \sigma^2 (X'X + kI)^{-1} X'X (X'X + kI)^{-1} \text{ .... (iii)}$$

The variance inflation factor, VIF (k), as a function of k is the j^th^ diagonal element of the matrix $(X'X + kI)^{-1} X'X (X'X + kI)^{-1}$.

The residual sum of squares can be written as

$$\begin{aligned} SSE(k) &= \left\{Y - X\hat{\beta}(k)\right\}' \left\{Y - X\hat{\beta}(k)\right\} \\ &= (Y - X\hat{\beta})'(Y - X\hat{\beta}) + \left\{\hat{\beta}(k) - \hat{\beta}\right\}' X'X \left\{\hat{\beta}(k) - \hat{\beta}\right\} \text{ .... (iv)} \end{aligned}$$

The total mean square error is

$$\begin{aligned} MSE(k) &= E\left[\left\{\hat{\beta}(k) - \beta\right\}' \left\{\hat{\beta}(k) - \beta\right\}\right] \\ &= \sigma^2 trace\left[(X'X + kI)^{-1} X'X (X'X + kI)^{-1}\right] + k^2 \beta'(X'X + kI)^{-2} \beta \\ &= \sigma^2 \sum_{j=1}^{n} \frac{\lambda_j}{(\lambda_j + k)^2} + k^2 \beta'(X'X + kI)^{-2} \beta \text{ .... (v)} \end{aligned}$$

Note that the first term on the right-hand side of equation (v) is the sum of the variances of the components of $\hat{\beta}(k)$ (total variance) and the second term is the square of the bias. Hoerl and Kennard (1970) prove that there exists a value of k > 0 such that

$$E\left[\left\{\hat{\beta}(k) - \beta\right\}' \left\{\hat{\beta}(k) - \beta\right\}\right] < E\left[\left(\hat{\beta} - \beta\right)' (\hat{\beta} - \beta)\right]$$

that is, the mean square error of the ridge estimator, $\hat{\beta}(k)$,is less than the mean square error of the OLS estimator, $\hat{\beta}$ (Chatterjee and Hadi,2006). Hoerl and Kennard (1970) suggest that an appropriate value of k may be selected by observing the ridge trace and some complementary summary statistics for $\hat{\beta}(k)$ such as SSE(k) and VIF(k).The value of k selected is the smallest value for which $\hat{\beta}(k)$ is stable.

4.19 Fit of the model

The purpose of Analysis of Variance (ANOVA) is to test for significant differences between means. The name analysis of variance is derived from the fact in order to test for statistical significance between means, we are actually comparing (i.e., analyzing) variances. Most of the variances are to be explained by the factor effect. The ANOVA table provides a formal F- test for the factor effect. The F- statistic is the mean square for the factor divided by the mean square for the error. This statistic follows an F distribution with (k-1) and (N - k) degrees of freedom. If the analysis of variance model results in a significant reduction of variation from the total, the F ratio is higher than expected. Therefore, it is clear that the model is a better fit statistically than the overall response mean.

4.20 Comparison between ordinary least squares (OLS) and ridge regression

The OLS estimate $\hat{\beta}$ of β is obtained by minimizing the residual sum of squares, and is given by:

$$RSS = (Y - X\hat{\beta})'(Y - X\hat{\beta}), \quad \hat{\beta} = (X'X)^{-1}X'Y$$

$$Var(\hat{\beta}) = \hat{\sigma}^2 (X'X)^{-1} \text{and} \, MSE(\hat{\beta}) = \hat{\sigma}^2 trace(X'X)^{-1} = \hat{\sigma}^2 \sum_{i=1}^{p} \frac{1}{\lambda_i} \text{...(1)}$$

where $\hat{\sigma}^2$ is the variance. This estimator $\hat{\beta}$ is an unbiased and has a minimum variance. When the independent variables are highly correlated, $X'X$ is ill-conditioned (singular), and the variance of the OLS estimator becomes large. With multicollinearity, the estimated OLS coefficients may be statistically insignificant even though the R-Square may be high.

In order to prevent these difficulties of OLS, Hoerl and Kennard (1970), suggested the ridge regression as an alternative procedure to the OLS method in regression analysis, especially, when multicollinearity exists. The addition of a small positive number k is to the diagonal elements of $X'X$ causes $X'X$ to be non-singular. Therefore, the ridge solution is given by:

$$\hat{\beta}_R = (X'X + kI)^{-1}X'Y, k \geq 0 \text{ .... (2)}$$

Where k is ridge parameter and I is identity matrix. Values of k lie in the range (0, 1). When k = 0, the ridge estimator become as the OLS.

From equation (2), by taking expectation on both sides, then

$E(\hat{\beta}_R) = A_k \beta$ where $A_k = \left[I + k(X'X)^{-1}\right]^{-1}$

and $Var\left(\hat{\beta}_R\right) = \hat{\sigma}^2 A_k (X'x)^{-1} A_k'$

The ridge estimator $\hat{\beta}_R = \left[I + k(X'X)^{-1}\right]^{-1}\hat{\beta}$ is a linear transformation of the OLS. The sum of the squared residuals is an increasing function of *k*. The mean squares error of ridge estimator is given by:

$$MSE\left(\hat{\beta}_R\right) = E\left[\left(\hat{\beta}_R - \beta\right)'\left(\hat{\beta}_R - \beta\right)\right] = \hat{\sigma}^2 trace\left[A_k (X'X)^{-1} A_k'\right] + \hat{\beta}'(I - A_k)'(I - A_k)\hat{\beta}$$

$$= \hat{\sigma}^2 \sum_{i=1}^{p} \frac{\lambda_i}{(\lambda_i + k)^2} + k^2 \hat{\beta}'(X'X + kI)^{-2}\hat{\beta} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(3)$$

Where, $\lambda_1, \lambda_2, \ldots\ldots\lambda_p$ are the eigenvalues of $X'X$ and the first term of the right hand in equation (3) is the trace of the dispersion matrix of the $\hat{\beta}_R$ and the second term is the square length of the bias vector. There always exists a $k > 0$, such that $\hat{\beta}_R$ has smaller MSE than $\hat{\beta}$, this means that $MSE\left(\hat{\beta}(k)\right) < MSE(\hat{\beta})$ .It indicates that ridge estimator performs better than the OLS estimator does. Ridge regression (RR) provides more informative results that are compared to the classical linear regression method to handle the problem of multicollinearity when predictors are highly correlated. It can be concluded that ridge regression model is better than classical linear regression when the multicollinearity problem exists. This is because it has smaller MSE of estimators, smaller variance for most estimators and has smaller coefficient of determination.

Chapter Five

Socio-Economic Condition of the Respondents in the Study Area: A Descriptive Analysis and Results

5.1 Introduction

The socio-economic condition means a situation, which is related to social and economic indicators. Rigidly interpreted the scope of such indicators can indeed be narrowed down to aspects such as employment, education, household income, age, sanitary facilities etc. The socio-economic condition of the respondents is a very important factor for explaining the impact of education on rice production. The socio-economic conditions of the respondents' are influenced on rice production. Some socio-economic characteristics of the respondents of the study area are presented in this section, which are described below.

5.2 Years of schooling of the respondents

Education plays an important role in making decisions concerning selection of seed, adoption of fertilizer, pesticides, supplying standard quality of inputs in time and contact of agriculture and bank officials for suggestion and advice regarding rice cultivation system and credit (Tetlay et.al. 1991). From Table 5.1, the average education level of the respondent is 6.32 years and the standard deviation of the education level of the respondent is 4.60 years. Maximum education level of the respondent is 16 years and minimum is 0.00 years. In Table 5.1, about 14.5 per cent respondents have no institutional education in the study area. About 33 per cent rice cultivators in the study area are included in the primary level of education. 35.8 per cent respondents fall in the secondary level of education. This group mainly operates the rice cultivation in the study area. About 7.3 per cent respondents included in the higher secondary level of education in the study area. About 9.5 per cent respondents fall in the tertiary level of education in the study area.

Table5.1: Years of schooling of the respondents

Item	Chapachil		Paschim Saidpur		Asrafpur		Study Area	
Years of Schooling	Number of respondent	percentage	Number of respondent	percentage	Number of respondent	percentage	Number of respondent	percentage
1	2	3	4	5	6	7	8	9
Illiterate	28	16.4	15	15.6	9	9.9	52	14.5
Primary	52	30.4	38	39.6	28	30.8	118	33.0
Secondary	64	37.4	32	33.3	32	35.2	128	35.8
Higher Secondary	11	6.4	6	6.3	9	9.9	26	7.3
Tertiary	16	9.4	5	5.2	13	14.3	34	9.5
Minimum	0.00		0.00		0.00		0.00	
Maximum	16.00		16.00		16.00		16.00	
Mean	6.28		5.44		7.30		6.32	
Standard deviation	4.61		4.35		4.68		4.60	

Source: Field survey, December 2012 and January 2013

5.3 Marital status of the respondents

The marital status of the respondents is divided into two categories namely married and unmarried, which is presented in Table 5.2. It is observed from the Table 5.2 only 96.6 per cent of the respondents are married and about 3.4per cent of the respondents are unmarried.

Table 5.2: Marital status of the respondents

Name of village	Married (%)	Unmarried (%)	Total (%)
Chapachil	99.4%	0.6%	100
Paschim Saidpur	94.8%	5.2%	100
Asrafpur	95.6%	4.4%	100
Average percentage	96.6%	3.4%	100

Source: Field survey, December 2012 and January 2013

5.4 Age distribution of the respondents

Rice production is influenced by the age of the respondents. In Table 5.3, the mean of age of the respondent is 39.95 years and the standard deviation of the age of the respondent is 11.24 years. Minimum age of the respondent is 20 years and maximum is 78 years in the study area.

Table 5.3: Age distribution of the respondents

Class Intervals (Years)	Chapachil		Paschim Saidpur		Asrafpur		Study Area	
	No. of cultivators	percentage	No. of cultivators	percentage	No. of cultivators	percentage	No. of Cultivators	Percentage
20-29	23	13.5	26	27.1	14	15.4	63	17.6
30-39	46	26.9	35	36.5	36	39.6	117	32.7
40-49	52	30.4	23	24.0	24	26.4	99	27.7
50-59	37	21.6	6	6.3	12	13.2	55	15.4
60-69	8	4.7	5	5.2	4	4.4	17	4.7
70 and above	5	2.9	1	1.0	1	1.1	7	2.0
Total	171	100	96	100	91	100	358	100
Minimum	23.00		20.00		22.00		20.00	
Maximum	78.00		75.00		70.00		78.00	
Mean	42.31		36.73		38.90		39.95	
Standard Deviation	11.27		11.20		10.29		11.24	

Source: Field survey, December 2012 and January 2013

Table 5.3 also indicates that 17.6 per cent respondents are included in the age group of 20-29 years; 32.7 per cent respondents are included in the age group of 30-39 years; 27.7 per cent respondents fall in the age group of 40-49 years; 15.4 per cent respondents are included in the age group of 50-59 years; 4.7 per cent respondents are included in the age group of 60-69 years and 2.0 per cent respondents are more than70 years old.

5.5 Sanitation facility of the respondents

Sanitation facility is one of the respondents' characteristics to assess socio-economic conditions. The types of sanitation being used by the respondents are an important indicator of access to health care facilities. In Bangladesh, sanitation facilities have improved over the years. About 61.6 per cent of the people of Bangladesh use sanitary latrine (water seal and without water seal) *Kutcha* latrine use 31.4 per cent and open space use only 7 per cent (BBS, 2012). About 46 per cent of the people of Shibganj *Upazila* use sanitary latrine (BBS, 2012). About 61.7 per cent respondents of the study area use sanitary latrine. Table 5.4 shows that 38.87 per cent of the respondents use *pucca* latrine, 22.83 per cent of the respondents use semi *pucca* latrine, 22.37per cent of the respondents use *katcha* latrine and 15.97 per cent use open space in study area.

Table 5.4: Sanitation facility of the respondents

Name of village	*Pucca* latrine	Semi *pucca* latrine	*Kutcha* latrine	Open space	Total
Chapachil	31.6%	25.7%	29.2%	13.5%	100
Paschim Saidpur	53.1%	21.9%	10.4%	14.6%	100
Asrafpur	31.9%	20.9%	27.5%	19.8%	100
Total average	38.87%	22.83%	22.37%	15.97%	100

Source: Field survey, December 2012 and January 2013

5.6 Monthly income distribution of the respondents

Rice production is very much influenced by the income of the respondents. Income is the first and foremost important factor to describe the socio-economic condition of the respondents. From Table 5.5, the mean of the monthly income of the respondent is Tk. 10736.66 and the standard deviation of the monthly income of the respondent is Tk. 10232.84. Minimum monthly income of the respondent is Tk. 1800.00 and maximum is Tk.50000.00.

Tables 5.5 about 46.6 percent of the respondents are included in the income group Tk.0-5000 in the study area. About 21.5 per cent respondents are included in the income group of 5001- Tk.10000; about 8.4 per cent respondents fall in the income group Tk.10001-15000; about 9.8 per cent respondents are included in the income group Tk.15001-20000; about 5.3 per cent respondents are included in the income group of Tk. 20001-25000; about 1.7 per cent respondents are included in the income group Tk. 25001-30000; about1.4 percent respondents are included in the income group of Tk. 30001-35000, about 1.4 percent respondents fall in the income group Tk. 35001-40000.

About 2.8 percent respondents are included in the income group Tk. 40001-45000 and about 1.1 percent respondents are included in the income group of Tk. 45001-50000in the study area.

Table 5.5: Distribution of monthly income of the respondents

Class (Income) (Tk.)	Chapachil		Paschim Saidpur		Asrafpur		Study Area	
	No.of cultivators	percentage	No.of cultivators	percentage	No.of cultivators	percentage	No.of cultivators	percentage
0-5000	89	52.0	33	34.4	45	49.5	167	46.6
5001-10000	30	17.5	38	39.6	9	9.9	77	21.5
10001-15000	14	8.2	12	12.5	4	4.4	30	8.4
15001-20000	17	9.9	7	7.3	11	12.1	35	9.8
20001-25000	12	7.0	3	3.1	4	4.4	19	5.3
25001-30000	1	0.6	0	0	5	5.5	6	1.7
30001-35000	1	0.6	1	1.0	3	3.3	5	1.4
35001-40000	0	0	2	2.1	3	3.3	5	1.4
40001-45000	3	1.8	0	1.0	5	5.5	10	2.8
45001-50000	4	2.3	0	1.0	2	2.2	4	1.1
50001and above	0	0	0	0	0	0	0	0
Total	171	100	96	100	91	100	358	100
Minimum	2400.00		1800.00		2000.00		1800.00	
Maximum	50000.00		36000.00		45000.00		50000.00	
Mean	10073.00		9336.89		13460.44		10736.66	
Std. Deviation	9907.45		6684.06		13128.52		10232.84	

Source: Field survey, December 2012 and January 2013

5.7 Family size of the respondents

The average family size is the average number of persons living in a household. The average family size in Bangladesh declines continuously. The average person per household was 4.4 in 2011, compared to 4.8 in 2001in Bangladesh (BBS, 2012). In Table 5.6, the mean of the family size of the respondent is 4.15 and the standard deviation of the family size of the respondent is 1.33. Minimum family size of the respondent is 2.00 and maximum is 10.00.

Table 5.6: Average family size of the respondents

Name of the Village	No. of the Respondent	Minimum	Maximum	Mean	Standard Deviation
Chapachil	171	2.00	10.00	4.25	1.57
Paschim Saidpu	96	2.00	8.00	4.2	1.03
Asrafpur	91	2.00	7.00	3.90	1.09
Total	358	2.00	10.00	4.15	1.33

Source: Field survey, December 2012 and January 2013

In case of Chapachi the average family size is 4.25 which are lower than the national average family size. In this study, the average family size is the highest in Chapachil. On the other hand, the average family size is the lowest in Asrafpur. The data indicates that the average family size of the respondent (4.15) appears to be lower than the national average family size (4.4) (BBS, 2012).

5.8 Conclusion

In this chapter, the socio-economic condition of the respondent in the study area has been examined. The results show that about 68.8 per cent of the respondents who obtained primary and secondary education engaged in rice production in the study area. About 96.6 per cent of the respondents are married. About 60.4 per cent of the respondent is middle age group that is (30-49) years. Results show that about 61.7 per cent respondents of the study area use sanitary latrine that is near about the national level. Results also show that about 68.1per cent respondents of the study area whose monthly income level is Tk. (0-10000).The average family size of the study area is 4.15 which is below the family size of national level at 4.4(GOB,2011).

Chapter Six

Results and Discussion

6.1 Introduction

In this chapter, the impact of education on rice production has been examined by using descriptive and inferential statistics. Chi square test is used to assess the association between level of education and rice production. Mean differences of the group of variable are obtained analysis of variance. Regression analysis has been employed to estimate the impact of education on rice production in the study area. Both quantitative and dichotomous variables are employed as explanatory variables in this study.

6.2 Chapachil

6.2.1 Descriptive statistics

Table 6.1 shows the variables that are used in estimations and their sample statistics namely maximum and minimum values, mean and standard deviation.

Table 6.1: Descriptive statistics of the variables

Variable	Unit of Measurement	No. of Cultivators	Minimum	Maximum	Mean	Standard Deviation
Output	Kg	171	900.00	7600.00	3283.8304	1648.42780
Yield	kg	171	18.28	26.09	22.2469	1.63351
Input cost	Tk.	171	2550.00	19000.00	9788.2982	4781.62141
Input cost	Tk. (per decimal)	171	50.88	75.76	66.4045	7.11170
Labour cost	Tk.	171	2100.00	15500.00	7728.5965	4000.23651
Labour cost	Tk. (per decimal)	171	40.91	60.61	51.7873	5.85851
Cultivable land	decimal	171	49.00	330.00	148.2164	75.19727
Education	years of schooling	171	0.00	16.00	6.2807	4.61616
Extension service	percentage	Yes=52.6 No=47.4				

Source: Field survey, December 2012 and January 2013

The mean, standard deviation, minimum and maximum of the variables are presented in Table 6.1. In Table 6.1, it is found that the average yield of rice is 22.24 kilograms with maximum average yield of 26.09 kilograms and minimum average yield of 18.28 kilograms. The average value of input cost is Tk. 66.40 with maximum and minimum average value of input cost is Tk. 75.76 and Tk. 50.88 respectively. The average value of labour cost is Tk. 51.78 with the maximum and minimum average value of labour cost is Tk. 60.61 and Tk. 40.91 respectively. The average of cultivable land is 148.21 decimal with the maximum and minimum of the cultivable land is 330 decimal and 49 decimal respectively. Table 6.1, the average level of education of the respondent is 6.28 years and the standard deviation of the education level of the respondent is 4.61 years. Maximum education level of the respondent is 16 years and minimum is 0.00 years. Maximum and minimum education level shows a wide variation of the respondents. About 52.6 per cent respondents of the study area are taken agricultural extension service from Sub Assistance Agriculture Officers and rest of 47.4 per cent do not take one.

6.2.2 Results of *chi-square* test (χ^2)

Chi square test is used to assess the association between different level of education and rice production in the study area. The null hypothesis of this test is that there is no relationship between different level of education and rice production. Alternative hypothesis: There is a relationship between different level of education and rice production.

Table 6.2: Impact of education on rice production in the study area

Yield (kg)	Illiterate	Primary	Secondary	Higher Secondary	Tertiary	Total
18-23	28	48	39	1	7	123
23-28	0	4	25	10	9	48
Total	28	52	64	11	16	171
$\chi^2 = 53.257$ df=4 *p* value =0.000						

Source: Field survey, December 2012 and January 2013

Table 6.2 shows the impact of education on rice production in the study area. In Table 6.2, the calculated value of χ^2 is 53.257 and the critical value of χ^2 for 4 degrees of freedom at 0.1% level of significance is 18.467. Since the calculated value of χ^2 is greater than the tabulated value, the null hypothesis can be rejected. So, the alternative hypothesis is accepted at the 0.1% level of significance. It can be said that there is a relationship between the two variables. So, there is evidence of a relationship between rice production and education.

Table 6.3: Impact of illiterate farmers on rice production

Yield(kg)	Illiterate	Others	Total
18-23	28	95	123
23-28	0	48	48
Total	28	143	171
$\chi^2 = 13.066$ df=1 p value =0.000			

Source: Field survey, December 2012 and January 2013

Table 6.3 shows the impact of illiterate farmers on rice production in the study area. In Table 6.3, the calculated value of χ^2 is 13.066 and the critical value of χ^2 for 1 degrees of freedom at 0.1 % level of significance is 10.827. Since 13.066>10.827, the null hypothesis can be rejected, and the alternative hypothesis is accepted at the 0.1% level of significance. That is to say, there is relationship between the two variables. So, there is evidence of a relationship between rice production and illiterate farmer. This is because; the experience of the illiterate farmers is higher than others.

Table 6.4: Experience of the respondent

Level of Education	No. of Cultivators	Percentage	Mean	Standard Deviation	Minimum	Maximum
Illiterate	28	16.4	36.7500	11.8309	17.00	58.0
Primary	52	30.4	28.7885	11.5933	12.00	58.0
Secondary	64	37.4	24.3750	11.7911	5.00	58.0
Higher Secondary	11	6.4	11.9091	7.1617	4.00	26.0
Tertiary	16	9.4	12.7500	7.9958	2.00	30.0
Total	171	100.0	25.8538	13.2094	2.00	58.0

Source: Field survey, December 2012 and January 2013

Table 6.4 shows that the average experience of illiterate rice farmer is 36.75 years, which is higher than others. The lowest average experience is 11.90 years, which consist on higher secondary level.

Table 6.5: Impact of primary education on rice production

Yield(kg)	Primary level	Others	Total
18-23	48	75	123
23-28	4	44	48
Total	52	119	171
$\chi^2 = 15.368$ df=1 p value =0.000			

Source: Field survey, December 2012 and January 2013

Table 6.5 shows the impact of Primary education on rice production. In Table 6.5, the calculated value of χ^2 is 15.368 and the critical value of χ^2 for 1degrees of freedom at 0.1% level of significance is 10.827. Since 15.368>10.827, the null hypothesis can be rejected, and the alternative hypothesis is accepted. It can be said that there is a relationship between the two variables. At the 0.1% level of significance, there is evidence of a relationship between rice production and primary of education.

Table 6.6: Impact of secondary education on rice production

Yield(kg)	Secondary level	Others	Total
18-23	39	84	123
23-28	25	23	48
Total	64	107	171
$\chi^2 = 6.121$ df=1 p value =0.013			

Source: Field survey, December 2012 and January 2013

Table 6.6 shows the impact of secondary education on rice production. In Table 6.6, the calculated value of χ^2 is 6.121 and the critical value of χ^2 for 1 degrees of freedom at 5% level of significance is 3.841. Since 6.121>3.841, the null hypothesis can be rejected, and the alternative hypothesis is accepted. It can be said that there is a relationship between the two variables. At the 5% level of significance, there is evidence of a relationship between rice production and secondary level of education.

Table 6.7: Impact of higher secondary level of education on rice production

Yield(kg)	Higher Secondary Level	Others	Total
18-23	1	122	123
23-28	10	38	48
Total	11	160	171
$\chi^2 = 22.992$ df=1 p value = 0.000			

Source: Field survey, December 2012 and January 2013

Table 6.7 shows the impact of higher secondary education on rice production. In Table 6.7, the calculated value of χ^2 is 22.992 and the critical value of χ^2 for 1 degrees of freedom at

0.1% level of significance is 10.827. Since 22.992>10.827, the null hypothesis can be rejected, and the alternative hypothesis is accepted at the 0.1% level of significance. That is to say, there is a relationship between the two variables. So, there is evidence of a relationship between rice production and higher secondary level of education.

Table 6.8: Impact of tertiary level of education on rice production

Yield(kg)	Tertiary level	Others	Total
18-23	7	116	123
23-28	9	39	48
Total	16	155	171
$\chi^2 = 6.942$ df=1 p value =0.008			

Source: Field survey, December 2012 and January 2013

Table 6.8 shows the impact of tertiary level education on rice production in the study area. In Table 6.8, the calculated value of χ^2 is 6.942 and the critical value of χ^2 for 1 degrees of freedom at 1 % level of significance is 6.635. Since 6.942>6.635, the null hypothesis can be rejected, and the alternative hypothesis is accepted at the 1% level of significance. That is to say, there is a relationship between the two variables. So, there is evidence of a relationship between rice production and tertiary level of education.

6.2.3 Analysis of variance (ANOVA)

One way ANOVA is used to examine mean differences between two or more groups. It is a bivariate test with one independent variable and one dependent variable. The independent variable must be categorical and the dependent variable must be continuous. This analysis is based on 171 randomly selected variables with no missing observations.

Table 6.9: Descriptive of the independent variable

Yield Level of education	N	Mean	Standard Deviation	Standard Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Illiterate	28	19.6514	0.85376	0.16135	19.3204	19.9825	18.28	20.90
Primary	52	22.3062	0.82658	0.11463	22.0761	22.5363	20.20	24.24
Secondary	64	22.7209	1.13306	0.14163	22.4378	23.0039	19.70	26.09
Higher Secondary	11	23.9542	1.08523	0.32721	23.2252	24.6833	21.60	25.51
Tertiary	16	23.5263	1.73208	0.43302	22.6034	24.4493	20.79	25.76
Total	171	22.2469	1.63351	0.12492	22.0003	22.4935	18.28	26.09

Source: Field survey, December 2012 and January 2013

The output includes in Table 6.9 that provides descriptive statistics for each of the factors for the independent variable. As the report indicates, the means of standardized reading scores for the Illiterate, Primary Education, Secondary Education, Higher Secondary Education and Tertiary Education are different from other groups. The results suggest that these groups may differ with regard to their mean. These differences, especially among the independent variables, warrant an ANOVA. The null hypothesis is that the means of various level of education group are equal.

Table 6.10: ANOVA of yield and level of education

Level of education \ Yield	Sum of Squares	df	Mean Square	*F*	*P* value
Between Groups	261.437	4	65.359	56.454	0.000
Within Groups	192.185	166	1.158		
Total	453.622	170			

Source: Field survey, December 2012 and January 2013

The results of the overall *F* test in the ANOVA summary Table 6.10 can be examined to determine whether group means are different or not. As indicated, the overall *F* test is significant (i.e., p value < 0.05), indicating that means between groups are not equal. Because, they reject the null hypothesis.

6.2.4 Empirical results

The empirical results of the production function in equation (3) presented in Table 6.11. In Table 6.11, the findings show that the input cost of production is insignificant and the coefficient of input cost of production is -0.01119. The results indicate that as input cost of production increases by Tk.1 with output decreases by -0.01119 kilogram. The labour cost of production is statistically insignificant. The coefficient of labour cost of production is -0.01393.The results indicate that if the labour cost of production increases by Tk.1, then the total output decreases by-0.01393 kilogram. The cultivable land is statistically highly significant. The coefficient of cultivable land is 1.018699.The results indicate that the cultivable land increases by 1 decimal, total production increases by 1.018699 kilogram per decimal.

Table 6.11: Empirical results of multiple regressions

	$\hat{\beta}$	St. Error	t	P value	Eigenvalue	Tolerance	VIF
1	2	3	4	5	6	7	8
Intercept	3.109359*	0.103653	29.997	0.000	5.462153	-	-
Input cost (k)	-0.01119	0.043358	-0.258	0.796	1.017213	0.012511	79.927
Labor cost (L)	-0.01393	0.041858	-0.332	0.739	1.00009	0.012337	81.055
Cultivable Land (T)	1.018699*	0.025779	39.516	0.000	1.000001	0.036879	27.115
Primary (D_1)	0.081445*	0.008863	9.189	0.000	0.413662	0.400685	2.495
Secondary (D_2)	0.100907*	0.008728	11.560	0.000	0.101371	0.373343	2.678
Higher Secondary (D_3)	0.129872*	0.013291	9.771	0.000	0.005349	0.626438	1.596
Tertiary (D_4)	0.12724*	0.011717	10.859	0.000	0.000136	0.572042	1.748
Extension Service (S)	0.076754*	0.006119	12.542	0.000	0.000024	0.713472	1.401
R^2		0.9960					
Adjusted R^2		0.9958					
Mean square error		54.37					

Source: Field survey, December 2012 and January 2013

* Highly significant

The coefficient of illiterate cultivator is 3.109359, which is highly significant. This is because, if the farmers experience increases, their total output increases. In this study, the level of experience is the highest of illiterate rice cultivators. The coefficient of primary education is (3.109359+0.081445) = 3.190804, which is highly significant. It indicates that if the primary education of farmer increases, their total output increases by 3.190804 kilogram. The coefficient of secondary education is (3.190804+0.100907) =3.291711, which is highly significant. If the secondary education of farmer increases, their total output increases by 3.291711 kilogram. The coefficient of higher secondary education is (3.291711+0.129872) =3.421583, which is highly significant. If the higher secondary education of farmer increases, their total output increases by 3.421583 kilogram. The coefficient of tertiary education is (3.421583+0.12724) =3.548823, which is highly significant. If the tertiary education of farmer increases, their total output increases by 3.548823 kilogram. The coefficient of extension service is 0.076754and it is statistically significant. It indicates that if the extension service increases, their total output increases by 0.076754 kilogram.

In Table 6.11, two variables of this model provide insignificant results and opposite sign. So, this model might suffer from multicollinearity problem.

6.2.5 Reliability and validity

To ensure the reliability of the questionnaire Cronbach's alpha test has been used in this study. The result of Cronbach's alpha test is given in Table 6.12.

Table 6.12 Test of reliability

Number of observation	Number of items	Cronbach`s Alpha
171	9	0.7850

Source: Field survey, December 2012 and January 2013

In Table 6.12, it is observed that Cronbach`s alpha is 0.7850 which indicates a high level of internal consistency for our scale with this specific sample.

In this study, variables and questions are drawn from literature, which ensured the validity of the questionnaire (Ali and Noman, 2013).

6.2.6 Diagnostic of the model

To check the reliability of the above results, the diagnosis of normality, multicollinearity, heteroscedasticity and autocorrelation are essential. For our postulated model, following the rule of thumb multicollinearity is not a troublesome problem. Again, to judge the validity of the above-mentioned results, though not predictable for cross-section data, the test for presence or absence of autocorrelation or serial correlation and heteroscedasticity have been conducted.

6.2.7 Normality test of the model

Normality is achieved by Jarque-Bera test. The null hypothesis of the JB test is residuals that are normally distributed and alternative hypothesis is not normally distributed.

6.13: Jarque-Bera normality test

Skewness	Kurtosis	Jarque-Bera statistic	*P* value
0.047	3.047	0.081095	0.960264

Source: Field survey, December 2012 and January 2013

In Table 6.13, the Jarque-Bera statistics is 0.081095 and the corresponding *p* value is 0.960264. Since the *p* value is more than 5 per cent, we accept null hypothesis meaning that the population residual is normally distributed which fulfills the assumption of a good regression line.

6.2.8 Fit of the model

Table 6.14: Analysis of variance (ANOVA)

	Sums of Squares	df	Mean Squares	*F*	*P* value
Regression	46.24969064	8	5.781211	5075.568	0.000
Residual	0.184522456	162	0.001139		
Total	46.43421309	170			

Source: Field survey, December 2012 and January 2013

Table 6.14, ANOVA summarizes how much of the variance in the data (total sum of squares) are accounted for by the factor effect (factor sum of squares) and how much is random error (residual sum of squares). From Table 6.14, *F* value is 5075.568and the *p* value is 0.000. This indicates that the results obtained from regression output are highly significant. Therefore, it is clear that the model is a better fit statistically.

6.2.9 Heteroscedasticity

Heteroscedasticity is obtained by white heteroscedasticity test. The null hypothesis of white test is that there is no heteroscedasticity and alternative hypothesis is that there is heteroscedasticity in the error term.

Table 6.15: White heteroscedasticity test

		P value
F Statistic	2.301845	0.012124
Obser* R^2	23.49047	0.015061

Source: Field survey, December 2012 and January 2013

It can be observed from Table 6.15 there is no heteroscedasticity in the error term of the model. The result is confirmed by White heteroscedaticity test. Obser*R^2=23.49047which has, asymptotically, a chi-square distribution with 44 df (Gujarati, 2003). The 5% critical chi-square value for 44 df is 60.481. Since the calculated value of chi-square is less than the critical value at 5% level of significance, it can be said that there is no heteroscedasticity in the error term of the model.

6.2.10 Autocorrelation

The Breusch-Godfrey serial correlation and Durbin-Watson statistic have been used to test for presence of serial correlation among the residuals. The null hypothesis of the Breusch-Godfrey serial correlation LM test is that there is no serial correlation in the residuals and alternative hypothesis there is serial correlation in the residuals. From Table 6.16, the*p*-value (0.291848) of Obs*R-squared is more than 5 percent ($p > 0.05$), we cannot reject null hypothesis meaning that residuals (u) are not serially correlated which is desirable. So, it can be said that the model is free from autocorrelation. The value of the Durbin-Watson statistic ranges from 0 to 4(Gujarati, 2003). As a general rule of thumb, the residuals are not correlated if the Durbin-Watson statistic is approximately 2 and an acceptable range is 1.50 to 2.50 (Alam et.al, 2013).

Table 6.16: Breusch-Godfrey serial correlation LM and Durbin-Watson tests

		P value
F Statistics	1.169142	0.313273
Obser**R-squared*	2.463046	0.291848
d Statistic (DW)	2.015	

Source: Field survey, December 2012 and January 2013

In Table 6.16, the value of *d* statistic is 2.015, which are about to 2. It indicates that there is no serial correlation.

6.2.11 Multicollinearity

Table 6.11 shows that there are three of eigenvalues close to zero and three VIF`s values more than 5. These results indicate that the model suffers from multicollinearity. It can also be found that the value of R^2 and adjusted R^2 are very high.

6.2.12 Results of ridge regression

Ridge regression has been applied to overcome the problem of multicollinearity. The ridge regression results have been shown in Table 6.17.
All VIF values are less than 5 which are shown in Table 6.17. These results indicate that this model is free from multicollinearity problems. It also shows the different results between Table 6.11 and Table 6.17. All variables are statistically significant except higher secondary level of education in Table 6.17.
The coefficient of input cost of production is 0.254384 and it is statistically highly significant. The results indicate that as input cost of production increases by Tk.1, output increases by 0.254384 kilogram.The coefficient of labour cost of production is 0.247299 it is statistically highly significant. The results indicate that if the labour cost of production increases by Tk.1, then output increases by 0.247299 kilogram. The coefficient of cultivable land is 0.42456 it is statistically highly significant. The results indicate that the cultivable land increases by 1 decimal, total production increases by 0.42456 kilogram per decimal.

The coefficient of illiterate farmer is 1.326683, which is highly significant. This is because, if the farmers experience increases, their total output increases. In this study, the level of experience is the high of illiterate rice farmers. The coefficient of primary education is (1.326683+0.056527) = 1.38321, which is significant. It indicates if the primary education of farmer increases, their total output increases by 1.38321 kilogram. The coefficient of secondary education is (1.38321+0.076071) =1.459281, which is highly significant. If the secondary education of farmer increases, their total output increases by 1.459281 kilogram.

Table 6.17: Empirical results of ridge regression

	$\hat{\beta}$	St. Error	t	P value	Tolerance	VIF
1	2	3	4	5	6	7
Intercept	1.326683*	0.208744	6.355565	0.0000	-	-
Input cost(K)	0.254384*	0.042057	6.0485	0.0000	0.207441	4.820647
Labor cost(L)	0.247299*	0.040405	6.120465	0.0000	0.206554	4.841339
Cultivable Land(T)	0.42456*	0.041187	10.30816	0.0000	0.225386	4.43683
Primary (D1)	0.056527**	0.027362	2.065869	0.0404	0.655851	1.524736
Secondary (D2)	0.076071*	0.026733	2.845624	0.0050	0.620887	1.610599
Higher Secondary (D3)	0.06812	0.04448	1.531483	0.1276	0.872623	1.14597
Tertiary (D4)	0.094951**	0.038114	2.491212	0.0137	0.843397	1.185681
Extension Service	0.06182*	0.021028	2.939929	0.0037	0.942637	1.060854
R^2	0.9380					
Adjusted R^2	0.9349					
Mean square error	3.36					
K (Ridge parameter)	0.13000					

Source: Field survey, December 2012 and January 2013

* Highly significant **5% level of significant

The coefficient of higher secondary education is (1.459281 + 0.06812) =1.527401, which is insignificant. If the higher secondary education of farmer increases, their total output increases by 1.527401 kilogram. The coefficient of tertiary education is (1.527401+ 0.094951) =1.622352, which is significant. If the tertiary education of farmer increases, their total output increases by 1.622352 kilogram.

The coefficient of extension service is 0.06182 and it is statistically significant. The results indicate that the extension service increases, total production increases.

6.2.13 Fit of the overall model

Table 6.18: Analysis of variance of ridge regression

	Sums of Squares	df	Mean Squares	F	P value
Regression	43.55563	8	5.444454	306.4015	0.000
Residual	2.878581	162	0.017769		
Total	46.43421	170			

Source: Field survey, December 2012 and January 2013

Table 6.18, analysis of variance of ridge regression summarizes how much of the variance in the data (total sum of squares) are accounted for by the factor effect (factor sum of squares) and how much is random error (residual sum of squares). In Table 6.18, *F* value is 306.40 and the overall results are highly significant.

6.2.14 Comparison between the OLS method and ridge regression method

It has been justified that ridge regression is better than OLS method. Table 6.11 shows the results of OLS. In Table 6.11, coefficient of determination (*R*-squared) is 0.9960, adjusted R^2 is 0.9958, 2nd column shows the OLS estimator of $\hat{\beta}$, Eigen values found column 6th, VIF is in column 8th. Here maximum VIF is 81.055, which indicates greater multicollinearity. It is observed that *R*-squared and adjusted *R*-squared is very high, and least squares estimates are unstable. The predictor variables are correlated so ridge regression techniques can be applied to stable set of correlation.

Table 6.17 shows the results of ridge regression. In Table 6.17, the coefficient of determination(*R*-squared) is 0.9380, adjusted R^2 is 0.9349, 2nd column shows, the ridge estimator of $\hat{\beta}_R$, VIF is in column 7th. It also observed that *R*-squared and adjusted *R*-squared is less than OLS and ridge estimates are stable than OLS estimates.

It is found from Table 6.11 and Table 6.17, the tolerance of OLS estimates is less than the tolerance of ridge estimates. As a result, the VIF for ridge estimates is less than the VIF for OLS estimates. These results indicate that the ridge regression method is better than OLS as it is clear from the Table 6.11 and Table 6.17.

It is estimated $\hat{\beta}$ using OLS estimator and estimate $\hat{\beta}_R$ using ridge estimator with different choices of *k* from a grid (0.01, 0.02..., and 0.13). It is computed mean square error for OLS estimator and mean square error for ridge regression estimator. From Table 6.11 and Table 6.17 MSE for OLS is greater than the MSE for ridge regression. This result indicates that ridge estimator performs better than the OLS estimator does.

6.2.15 Discussion

Table 6.17, the input cost of production has positive and significant (*p* value = 0.0000) effect on rice production. If the input cost of production increases, output increases. The same results in line with Appleton & Balihuta (1996) and Weir (1999). Table 6.17 shows that the labour cost of production has positive and significant (*p* value = 0.0000) effect on rice production. The results indicate that if the labour cost of production increases, and then output increases. The findings were consistent with studies by Cotlear (1986), Appleton & Balihuta (1996), Yang (1997) and Weir (1999). Table 6.17 reveals that the cultivable land has positive and significant (*p* value = 0.0000) effect on rice production. The results indicate that the cultivable land increases, then total production increases. The same results were found by Cotlear (1986), Appleton & Balihuta (1996), Yang (1997), Weir (1999) and Rehman et al. (2012).

Table 6.2 the *Chi-Square* results show that there is a significant (*p* value = 0.000) relationship between education and rice production in the study area. In Table 6.10 the analysis of variance (ANOVA) also shows that there are significant (*p* value = 0.000) relationships among the means of the different level of education. ANOVA results suggest that the different level of educational group may differ with regard their mean scores in Table 6.9. YEARS of schooling are found to be a significant impact on rice production, with one additional year of education yielding an increase in rice yields.

Table 6.5 the *Chi-Square* result shows that there is a significant (*p* value = 0.000) relationship between primary education and rice yield. Similar results were found in ANOVA. In Table 6.17 ridge regression result also shows that the primary education has positive and significant (*p* value = 0.0404) effect on rice production. So, the level of primary education of the rice cultivators increases, their total output increases. The similar results were found by Singh (1974), Dominique van de Walle (2003),Onphanhdala (2009) and Haq (2012).

Table 6.6 the *Chi-Square* results shows that there is a significant (*p* value = 0.013) relationship between secondary education and rice yield. ANOVA also supports this result. Ridge regression results show that the secondary education has positive and significant (*p* value = 0.0050) impact on rice production. So, the level of secondary education of the rice cultivators increases, their total output increases. The similar results were found by Singh (1974) and Asadullah & Rahman (2006).

Table 6.7 the *Chi-Square* results shows that there is a significant (*p* value = 0.000) relationship between higher secondary education and rice yield. ANOVA also confirm this result. Ridge regression results reveal that higher secondary education has positive and (*p* value = 0.1276) insignificant impact on rice production. So, the level of higher secondary education of the rice cultivators increases, their total output increases also.

Table 6.8 the *Chi-Square* results shows that there is a significant (*p* value = 0.008) relationship between tertiary education and rice yield. Similar results were found in ANOVA. Ridge regression results show that tertiary education has positive and significant (*p* value = 0.0137) effect on rice production. So, the level of tertiary education of the rice cultivators increases, their total output increases. The similar results were found by Pudasaini (1983) and Gemmell (1996).

In Table 6.17, shows that the extension service has positive and significant (*p* value = 0.0037) effects in improving rice productivity. It is clarified that the more the extension contacts between extension agents and farmers leads to the higher the productivity. The similar results were found by Huffman (1974),Haq (2011) and Nargis & Lee (2013).

6.3 Paschim Saidpur

6.3.1 Descriptive statistics

Table 6.19 shows the variables that are used in estimations and their sample statistics namely maximum and minimum values, mean and standard deviation.

The mean, standard deviation, minimum and maximum of the variables have been presented in Table 6.19. From Table 6.19, it is found that the average yield of rice is 22.55 kilograms with maximum average yield of 26.77 kilograms and minimum average yield of 18.29 kilograms. The average value of input cost is Tk. 68.45 with maximum and minimum average value of input cost is Tk. 75.76 and Tk. 51.52 respectively. The average value of labour cost is Tk. 53.94 with the maximum and minimum average value of labour cost is Tk. 60.00 and Tk. 42.42 respectively. The average of cultivable land is 148.76 decimal with the maximum and minimum of the cultivable land is 330 decimal and 49 decimal respectively.

Table 6.19: Descriptive statistics of the variables

Variable	Unit of Measurement	No. of Cultivators	Minimum	Maximum	Mean	Standard Deviation
Output	Kg	96	950.00	7400.00	3349.4792	1818.48744
Yield	kg	96	18.29	26.77	22.5571	1.80761
Input cost	Tk.	96	2825.00	18500.00	9919.7292	4848.35533
Input cost	Tk. (per decimal)	96	51.52	75.76	68.4541	6.88155
Labour cost	Tk.	96	2150.00	15800.00	7907.2500	4064.15996
Labour cost	Tk. (per decimal)	96	42.42	60.00	53.9425	4.89606
Cultivable land	decimal	96	49.00	330.00	148.7604	80.75144
Education	years of schooling	96	0.00	16.00	5.4479	4.35979
Extension service	percentage	Yes = 63.5 No = 36.5				

Source: Field survey, December 2012 and January 2013

In Table 6.19, the average level of education of the respondent is 5.45 years and the standard deviation of the education level of the respondent is 4.35 years. Maximum education level of the respondent is 16 years and minimum is 0.00 years. Maximum and minimum education level shows a wide variation of the respondents. About 63.5 per cent respondents of the study area are taken agricultural extension service from Sub Assistance Agriculture Officers and rest of 36.5 per cent do not take one.

6.3.2 Results of ***chi-square*** test (χ^2)

Chi square test is used to assess the association between different level of education and rice production in the study area. The null hypothesis of this test is that there is no relationship between different level of education and rice production. Alternative hypothesis: There is a relationship between different level of education and rice production.

Table 6.20: Impact of education on rice production in the study area

Yield (kg)	Illiterate	Primary	Secondary	Higher Secondary	Tertiary	Total
18-23	15	26	14	3	2	60
23-28	0	12	18	3	3	36
Total	15	38	32	6	5	96
$\chi^2 = 15.848$ df=4 p value =0.003						

Source: Field survey, December 2012 and January 2013

Table 6.20 shows the impact of education on rice production in the study area. In Table 6.20, the calculated value of χ^2 is 15.848 and the critical value of χ^2 for 4 degrees of freedom at 1% level of significance is 13.277. Since the calculated value of χ^2 is greater than the tabulated value, the null hypothesis can be rejected. So, the alternative hypothesis is accepted at the 1% level of significance. It can be said that there is a relationship between the two variables. So, there is evidence of a relationship between rice production and education.

Table 6.21: Impact of illiterate farmers on rice production

Yield(kg)	Illiterate	Others	Total
18-23	15	45	60
23-28	0	36	36
Total	15	81	96
$\chi^2 = 10.667$ df=1 *p* value =0.002			

Source: Field survey, December 2012 and January 2013

Table 6.21 shows the impact of illiterate farmers on rice production in the study area. In Table 6.21, the calculated value of χ^2 is 10.667 and the critical value of χ^2 for 1 degrees of freedom at 0.2 % level of significance is 9.550. Since 10.667 > 9.550, the null hypothesis can be rejected, and the alternative hypothesis is accepted at the 0.2% level of significance. So, there is evidence of a relationship between rice production and illiterate farmer. This is because the experience of the illiterate farmers is higher.

Table 6.22: Experience of the respondent

Level of Education	No. of cultivators	Percentage	Mean	Standard Deviation	Minimum	Maximum
Illiterate	15	15.6	21.3333	9.6337	7.00	40.0
Primary	38	39.0	19.4474	11.1880	5.00	56.0
Secondary	32	33.3	20.4375	12.3391	4.00	50.0
Higher Secondary	6	6.3	9.1667	4.7504	4.00	16.0
Tertiary	5	5.2	2.8000	1.3038	2.00	5.0
Total	96	100	18.5625	11.6074	2.00	56.0

Source: Field survey, December 2012 and January 2013

Table 6.22 shows that the average experience of illiterate rice farmer is 21.33 years, which is higher than others. The lowest average experience is 2.80 years, which consist on tertiary level.

Table 6.23: Impact of primary education on rice production

Yield(kg)	Primary level	Others	Total
18-23	26	34	60
23-28	12	24	36
Total	38	58	96
$\chi^2 = 0.941$ df=1 p value =0.332			

Source: Field survey, December 2012 and January 2013

Table 6.23 shows the impact of Primary education on rice production. In Table 6.23, the calculated value of χ^2 is 0.941 and it is insignificant. So, the null hypothesis can be accepted. It can be said that there is no relationship between the two variables. Because the experience of primary educated farmers of this village is 19.45 years.

Table 6.24: Impact of secondary education on rice production

Yield(kg)	Secondary level	Others	Total
18-23	14	46	60
23-28	18	18	36
Total	32	64	96
$\chi^2 = 7.200$ df=1 p value =0.007			

Source: Field survey, December 2012 and January 2013

Table 6.24 shows the impact of secondary education on rice production. In Table 6.24, the calculated value of χ^2 is 7.200 and the critical value of χ^2 for 1 degrees of freedom at 1% level of significance is 6.635. Since 7.200>6.635, the null hypothesis can be rejected, and the alternative hypothesis is accepted. It can be said that there is a relationship between the two variables. At the 1% level of significance, there is evidence of a relationship between rice production and secondary level of education.

Table 6.25: Impact of higher secondary level of education on rice production

Yield(kg)	Higher Secondary Level	Others	Total
18-23	3	57	60
23-28	3	33	36
Total	6	90	96
$\chi^2 = 0.427$ df=1 p value = 0.514			

Source: Field survey, December 2012 and January 2013

Table 6.25 shows the impact of higher secondary education on rice production. In Table 6.25, the calculated value of χ^2 is 0.427 and it is insignificant. So, the null hypothesis can be accepted. That is to say, there is no relationship between the two variables. This is because the experience of higher secondary educated farmers of this village is 9.17 years.

Table 6.26: Impact of tertiary level of education on rice production

Yield(kg)	Tertiary level	Others	Total
18-23	2	58	60
23-28	3	33	36
Total	5	91	96
$\chi^2 = 1.139$ df=1 p value =0.286			

Source: Field survey, December 2012 and January 2013

Table 6.26 shows the impact of tertiary level education on rice production in the study area. In Table 6.26, the calculated value of χ^2 is 1.139 and it is insignificant. So, the null hypothesis can be accepted. That is to say, there is no relationship between the two variables. This is because the experience of tertiary educated farmers of this village is 2.80 years.

6.3.3 Analysis of variance

One way ANOVA is used to examine mean differences between two or more groups. It is a bivariate test with one independent variable and one dependent variable. The independent variable must be categorical and the dependent variable must be continuous. This analysis is based on 96 randomly selected variables with no missing observations.

Table 6.27: Descriptive of the independent variable

Yield / Level of education	N	Mean	Standard Deviation	Standard Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Illiterate	15	19.3298	0.78147	0.20177	18.8971	19.7626	18.29	20.74
Primary	38	22.8376	0.69033	0.11199	22.6107	23.0645	21.21	24.44
Secondary	32	23.3989	1.41535	0.25020	22.8886	23.9092	20.27	26.01
Higher Secondary	6	23.4005	1.06004	0.43276	22.2881	24.5130	22.26	24.85
Tertiary	5	23.7069	2.43589	1.08936	20.6824	26.7315	20.79	26.77
Total	96	22.5571	1.80761	0.18449	22.1908	22.9233	18.29	26.77

Source: Field survey, December 2012 and January 2013

The output includes in Table 6.27 that provides descriptive statistics for each of the factors for the independent variable. As the report indicates, the means of standardized reading scores for the Illiterate, Primary Education, Secondary Education, Higher Secondary Education and Tertiary Education are different from other groups. The results suggest that these groups may differ with regard to their mean. These differences, especially among the independent variables, warrant an ANOVA. The null hypothesis is that the means of various level of education group are equal.

Table 6.28: ANOVA of yield and level of education

Level of education \ Yield	Sum of Squares	df	Mean Square	*F*	*P* value
Between Groups	192.774	4	48.193	37.281	0.000
Within Groups	117.635	91	1.293		
Total	310.409	95			

Source: Field survey, December 2012 and January 2013

The results of the overall F test in the ANOVA summary Table 6.28 can be examined to determine whether group means are different or not. As indicated, the overall F test is significant (i.e., p value < 0.05), indicating that means between groups are not equal. Because, they reject the null hypothesis.

6.3.4 Empirical results

The empirical results of the production function in equation (3) presented in Table 6.29. In Table 6.29, the findings show that the input cost of production is insignificant and the coefficient of input cost of production is 0.069843. The results indicate that as input cost of production increases by Tk.1 with output increases by 0.069843 kilogram. The labour cost of production is statistically insignificant .The coefficient of labour cost of production is

-0.07802.The results indicate that if the labour cost of production increases by Tk.1, then the total output decreases by -0.07802 kilogram. The cultivable land is statistically highly significant. The coefficient of cultivable land is 0.992606.The results indicate that the cultivable land increases by 1 decimal, total production increases by 0.992606 kilogram per decimal.

Table 6.29: Empirical results of multiple regressions

	$\hat{\beta}$	St. Error	t	P value	Eigenvalue	Tolerance	VIF
1	2	3	4	5	6	7	8
Intercept	3.047578*	0.200495	15.20	0.000	5.580373	-	-
Input cost(k)	0.069843	0.074718	0.934	0.352	1.003256	0.009608	104.083
Labor cost(L)	-0.07802	0.076667	-1.017	0.311	1.000265	0.008507	117.553
Cultivable Land(T)	0.992606*	0.043055	23.05	0.000	1.000003	0.025503	39.2103
Primary(D_1)	0.114214*	0.013914	8.208	0.000	0.331189	0.308103	3.24566
Secondary(D_2)	0.142338*	0.014616	9.738	0.000	0.079328	0.300504	3.32773
Higher Secondary(D_3)	0.136665*	0.021152	6.461	0.000	0.005481	0.54419	1.83759
Tertiary(D_4)	0.16098*	0.020533	7.840	0.000	0.000088	0.685393	1.45901
Extension Service(S)	0.074288*	0.010256	7.243	0.000	0.000016	0.585474	1.70801
R^2		0.9959					
Adjusted R^2		0.9955					
Mean square error		96.698					

Source: Field survey, December 2012 and January 2013

* Highly significant

The coefficient of illiterate farmer is 3.047578, which is highly significant. This is because, if the farmers experience increases, their total output increases. In this study, the level of experience is the highest of illiterate rice farmer. The coefficient of primary education is (3.047578 + 0.114214) = 3.161792, which is highly significant. It indicates that if the primary education of farmer increases, their total output increases by 3.161792 kilogram. The coefficient of secondary education is (3.161792 + 0.142338) =3.30413, which is highly significant. If the secondary education of farmer increases, their total output increases by 3.30413 kilogram. The coefficient of higher secondary education is (3.30413 + 0.136665) = 3.440795, which is highly significant. If the higher secondary education of farmer increases, their total output increases by 3.440795 kilogram. The coefficient of tertiary education is (3.440795 + 0.16098) =3.601775, which is highly significant. If the tertiary education of farmer increases, their total output increases by 3.601775 kilogram. The coefficient of extension service is 0.074288 and it is statistically highly significant. It indicates that if the extension service increases, their total output increases by 0.074288 kilogram.

In Table 6.29, two variables of this model provide insignificant results and one is opposite sign. So, this model might suffer from multicollinearity problem.

6.3.5 Reliability and validity

To ensure the reliability of the questionnaire Cronbach's alpha test has been used in this study. The result of Cronbach's alpha test is given in Table 6.30.

Table 6.30: Test of reliability

Number of observation	Number of items	Cronbach`s Alpha
96	9	0.7973

Source: Field survey, December 2012 and January 2013

In Table 6.30, it is observed that Cronbach`s alpha is 0.7973 which indicates a high level of internal consistency for our scale with this specific sample.

In this study, variables and questions are drawn from literature, which ensured the validity of the questionnaire (Ali and Noman, 2013).

6.3.6 Diagnostic of the model

To check the reliability of the above results, the diagnosis of normality, multicollinearity, heteroscedasticity and autocorrelation are essential. For our postulated model, following the rule of thumb multicollinearity is not a troublesome problem. Again, to judge the validity of the above-mentioned results, though not predictable for cross-section data, the test for presence or absence of autocorrelation or serial correlation and heteroscedasticity have been conducted.

6.3.7 Normality test

Normality is achieved by Jarque-Bera test. The null hypothesis of the JB test is that the residuals are normally distributed and alternative hypothesis is not normally distributed.

6.31: Jarque-Bera normality test

Skewness	Kurtosis	Jarque-Bera statistic	*P* value
0.232	2.905	0.900345	0.637518

Source: Field survey, December 2012 and January 2013

In Table 6.31, the Jarque-Bera statistics is 0.900345 and the corresponding *p*- value is 0.637518. Since the *p* value is more than 5 per cent, we accept null hypothesis meaning that the population residual is normally distributed which fulfills the assumption of a good regression line.

6.3.8 Fit of the model

Table 6.32: Analysis of variance (ANOVA)

	Sums of Squares	df	Mean Squares	*F*	*P* value
Regression	29.24732	8	3.655915	2669.48	0.000
Residual	0.119149	87	0.00137		
Total	29.36647	95			

Source: Field survey, December 2012 and January 2013

Table 6.32, ANOVA summarizes how much of the variance in the data (total sum of squares) are accounted for by the factor effect (factor sum of squares) and how much is random error (residual sum of squares). From Table 6.32, *F* value is 2669.48 and the *p* value is 0.000. This indicates that the results obtained from regression output are highly significant. Therefore, it is clear that the model is a better fit statistically.

6.3.9 Heteroscedasticity

Heteroscedasticity is obtained by white heteroscedasticity test. The null hypothesis of white test is that there is no heteroscedasticity and alternative hypothesis is that there is heteroscedasticity in the error term.

Table 6.33: White heteroscedasticity test

		P value
F Statistic	1.592681	0.115773
Obser* R^2	16.56698	0.121353

Source: Field survey, December 2012 and January 2013

As can be observed from Table 6.33 there is no heteroscedasticity in the error term of the model. The result is confirmed by White heteroscedaticity test. Obser*R^2=16.56698which has, asymptotically, a chi-square distribution with 44 df (Gujarati, 2003). The 20% critical chi-square value for 44 df is 51.639. Since the calculated value of chi-square is less than the critical value at 20% level of significance, it can be said that there is no heteroscedasticity in the error term of the model.

6.3.10 Autocorrelation

The Breusch-Godfrey serial correlation and Durbin-Watson statistic have been used to test for presence of serial correlation among the residuals. The null hypothesisof the

Breusch-Godfrey serial correlation LM test is that there is no serial correlation in the residuals and alternative hypothesis is that there is serial correlation in the residuals. From Table 6.34, the*p*-value (0.180852) of Obs*R-squared is more than 5 percent ($p > 0.05$), we cannot reject null hypothesis meaning that residuals (u) are not serially correlated which is desirable. So, it can be said that the model is free from autocorrelation. The value of the Durbin-Watson statistic ranges from 0 to 4(Gujarati, 2003). As a general rule of thumb, the residuals are not correlated if the Durbin-Watson statistic is approximately 2 and an acceptable range is 1.50 to 2.50 (Alam et.al, 2013).

Table 6.34: Breusch-Godfrey serial correlation LM and Durbin-Watson tests

		P value
F Statistics	1.570064	0.214005
Obser**R-squared*	3.420147	0.180852
d Statistic (DW)	1.833	

Source: Field survey, December 2012 and January 2013

In Table 6.34, the value of *d* statistic is 1.833, which is about to 2. It indicates that there is no serial correlation.

6.3.11 Multicollinearity

Table 6.29 shows that there are three of eigenvalues close to zero and three VIF`s values more than 5. These results indicate that the model suffers from multicollinearity. It can also be found that the value of R^2 and adjusted R^2 are very high.

6.3.12 Results of ridge regression

Ridge regression has been applied to overcome the problem of multicollinearity. The ridge regression results have been shown in Table 6.35.

All VIF values are less than 5 which is shown in Table 6.35. These results indicate that this model is free from multicollinearity problems. It also shows the different results between Table 6.29 and Table 6.35. All variables are statistically significant except higher secondary education and extension service in Table 6.35.

The coefficient of input cost of production is 0.293052and it is statistically highly significant. The results indicate that as input cost of production increases by Tk.1, output increases by 0.293052kilogram.The coefficient of labour cost of production is 0.286722it is statistically highly significant. The results indicate that if the labour cost of production increases by Tk.1, then output increases by 0.286722kilogram. The coefficient of cultivable land is 0.390897it is statistically highly significant. The results indicate that the cultivable land increases by 1 decimal, total production increases by 0.390897kilogram per decimal.

Table 6.35: Empirical results of ridge regression

	$\hat{\beta}$	St. Error	*t*	*P* value	Tolerance	VIF
1	2	3	4	5	6	7
Intercept	0.770394**	0.322579	2.388236	0.019	-	-
Input cost(K)	0.293052*	0.060723	4.826047	0.000	0.203011	4.925845
Labor cost(L)	0.286722*	0.058899	4.86807	0.000	0.201155	4.971282
Cultivable Land(T)	0.390897*	0.055346	7.062747	0.000	0.215388	4.642775
Primary (D1)	0.074098***	0.038544	1.922448	0.057	0.560384	1.784491
Secondary (D2)	0.104254**	0.040375	2.58211	0.011	0.549592	1.819532
Higher Secondary (D3)	0.061121	0.063665	0.960033	0.339	0.838314	1.19287
Tertiary (D4)	0.10938***	0.06542	1.671969	0.098	0.942265	1.061273
Extension Service	0.050398	0.031237	1.613376	0.110	0.88075	1.135396
R^2	0.9433					
Adjusted R^2	0.9381					
Mean square error	6.275					
k(Ridge parameter)	0.13000					

Source: Field survey, December 2012 and January 2013

* Highly significant **5% level of significant***10% level of significant

The coefficient of illiterate farmer is 0.770394, which is significant. This is because, if the farmers experience increases, their total output increases. In this study, the level of experience is the high of illiterate rice farmers. The coefficient of primary education is (0.770394+0.074098) = 0.844492, which is significant. It indicates that if the primary education of farmer increases, their total output increases by 0.844492 kilogram. The coefficient of secondary education is (0.844492 +0.104254) =0.948746, which is significant. If the secondary education of farmer increases, their total output increases by 0.948746 kilogram.

The coefficient of higher secondary education is (0.948746 + 0.061121) =1.009867, which is insignificant. If the higher secondary education of farmer increases, their total output increases by 1.009867 kilogram.The coefficient of tertiary education is (1.009867+ 0.10938) =1.119247, which is significant. If the tertiary education of farmer increases, their total output increases by 1.119247 kilogram.

The coefficient of extension service is 0.050398 and it is statistically insignificant. The results indicate that the extension service increases, total production increases.

6.3.13 Fit of the overall model

Table 6.36: Analysis of variance of ridge regression

	Sums of Squares	df	Mean Squares	*F*	*P* value
Regression	27.70363	8	3.462953	181.182	0.000
Residual	1.662841	87	0.019113		
Total	29.36647	95			

Source: Field survey, December 2012 and January 2013

Table 6.36, analysis of variance of ridge regression summarizes how much of the variance in the data (total sum of squares) are accounted for by the factor effect (factor sum of squares) and how much is random error (residual sum of squares). In Table 6.36, *F* value is 181.182 and the overall results are highly significant.

6.3.14 Comparison between the OLS Method and Ridge Regression Method

It has been justified that ridge regression is better than OLS method. Table 6.29 shows the results of OLS. In Table 6.29, coefficient of determination (*R*-squared) is 0.9959, adjusted R^2 is 0.9955, 2nd column shows the OLS estimator of $\hat{\beta}$, Eigen values found column 6th, VIF is in column 8th. Here maximum VIF is 117.553, which indicates greater multicollinearity. It is observed that *R*-squared and adjusted *R*-squared is very high, and least squares estimates are unstable. The predictor variables are correlated so ridge regression techniques can be applied to stable set of correlation.

Table 6.35 shows the results of ridge regression. In Table 6.35, the coefficient of determination(*R*-squared) is 0.9433, adjusted R^2 is 0.9381, 2nd column shows the ridge estimator of $\hat{\beta}_R$, VIF is in column 7^{th}, it also observed that *R*-squared and adjusted *R*-squared are less than OLS, and ridge estimates are stable than OLS estimates.

It is found from Table 6.29 and Table 6.35, the tolerance of OLS estimates is less than the tolerance of ridge estimates. As a result, the VIF for ridge estimates is less than the

VIF for OLS estimates. These results indicate that the ridge regression method is better than OLS as it is clear from the Table 6.29 and Table 6.35.

It is estimated $\hat{\beta}$ using OLS estimator and estimate $\hat{\beta}_R$ using ridge estimator with different choices of k from a grid (0.01, 0.02..., and 0.13). It is computed mean square error for OLS estimator and mean square error for ridge regression estimator. From Table 6.29 and Table 6.35 MSE for OLS is greater than the MSE for ridge regression. This result indicates that ridge estimator performs better than the OLS estimator does.

6.3.15 Discussion

Table 6.35, the input cost of production has positive and significant (p value = 0.0000) effect on rice production. If the input cost of production increases, output increases. The same results in line with Appleton & Balihuta (1996) and Weir (1999). Table 6.35 shows that the labour cost of production has positive and significant (p value = 0.0000) effect on rice production. The results indicate that if the labour cost of production increases, and then output increases. The findings were consistent with studies by Cotlear (1986), Appleton & Balihuta (1996), Yang (1997) and Weir (1999). Table 6.35 reveals that the cultivable land has positive and significant (p value = 0.0000) effect on rice production. The results indicate that the cultivable land increases, then total production increases. The same results were found by Cotlear (1986), Appleton & Balihuta (1996), Yang (1997), Weir (1999) and Rehman et al. (2012).

Table 6.20 the *Chi-Square* results show that there is a significant (p value = 0.003) relationship between education and rice production in the study area. In Table 6.28 the analysis of variance (ANOVA) also shows that there are significant (p value = 0.000) relationships among the means of the different level of education. ANOVA results suggest that the different level of educational group may differ with regard their mean scores in Table 6.27. Years of schooling are found to be a significant impact on rice production, with one additional year of education yielding an increase in rice yields.

Table 6.23 the *Chi-Square* result shows that there is no significant (p value = 0.332) relationship between primary education and rice yield. But in Table 6.28 the analysis of variance (ANOVA) shows that there are significant (p value = 0.000) relationships among the means of the different level of education. In Table 6.35 ridge regression result also shows that the primary education has positive and significant (*p* value = 0.057) effect on rice production. So, the level of primary education of the rice cultivators increases, their total output increases. The similar results were found by Singh (1974), Dominique van de Walle (2003),Onphanhdala (2009) and Haq(2012).

Table 6.24 the *Chi-Square* results shows that there is a significant (*p* value = 0.007) relationship between secondary education and rice yield. ANOVA also supports this result. Ridge regression results show that the secondary education has positive and significant (*p* value = 0.011) impact on rice production. So, the level of secondary education of the rice cultivators increases, their total output increases. The similar results were found by Singh (1974) and Asadullah & Rahman (2006).

Table 6.25 the *Chi-Square* results shows that there is no significant (*p* value = 0.514) relationship between higher secondary education and rice yield. In Table 6.28 the analysis of variance (ANOVA) shows that there are significant (*p* value = 0.000) relationships among the means of the different level of education. In Table 6.35, ridge regression results reveal that higher secondary education has positive but insignificant (*p* value = 0.339) impact on rice production. So, the level of higher secondary education of the rice cultivators increases, their total output increases also.

Table 6.26 the *Chi-Square* results shows that there is no significant (*p* value = 0.286) relationship between tertiary education and rice yield. In Table 6.28 the analysis of variance (ANOVA) also shows that there are significant (*p* value = 0.000) relationships among the means of the different level of education. In Table 6.35, ridge regression results show that tertiary education has positive and significant (*p* value = 0.098) effect on rice production. So, the level of tertiary education of the rice cultivators increases, their total output increases. The similar results were found by Pudasaini (1983) and Gemmell (1996).

In Table 6.35, shows that the extension service has positive but insignificant (*p* value = 0.110) effects in improving rice productivity. It is clarified that the more the extension contacts between extension agents and farmers leads to the higher the productivity.

6.4 Asrafpur

6.4.1 Descriptive statistics

Table 6.37 shows the variables that are used in estimations and their sample statistics namely maximum and minimum values, mean and standard deviation.

Table 6.37: Descriptive statistics of the variables

Variable	Unit of Measurement	No. of cultivators	Minimum	Maximum	Mean	Standard Deviation
Output	Kg	91	1100.00	7500.00	3756.0440	1789.81044
Yield	kg	91	18.26	26.77	23.0511	1.91069
Input cost	Tk.	91	2550.00	19525.00	11070.4835	4860.53051
Input cost	Tk. (per decimal)	91	51.00	75.76	68.4001	7.40649
Labour cost	Tk.	91	2200.00	15500.00	8872.4725	4083.85550
Labour cost	Tk. (per decimal)	91	42.42	60.00	54.2210	5.27292
Cultivable land	decimal	91	49.00	330.00	165.8242	82.75843
Education	years of schooling	91	0.00	16.00	7.3077	4.68495
Extension service	percentage	Yes = 64.8 No = 35.2				

Source: Field survey, December 2012 and January 2013

The mean, standard deviation, minimum and maximum of the variables are presented in Table 6.37. In Table 6.37, it is found that the average yield of rice is 23.05 kilograms with maximum average yield of 26.77 kilograms and minimum average yield of 18.26 kilograms. The average value of input cost is Tk. 68.40 with maximum and minimum average value of input cost is Tk. 75.76 and Tk. 51.00 respectively. The average value of labour cost is Tk. 54.22 with the maximum and minimum average value of labour cost is Tk. 60.00 and Tk. 42.42 respectively. The average of cultivable land is 165.82 decimal with the maximum and minimum of the cultivable land is 330 decimal and 49 decimal respectively. In Table 6.37, the average level of education of the respondent is 7.31 years and the standard deviation of the education level of the respondent is 4.68 years. Maximum education level of the respondent is 16 years and minimum is 0.00 years. Maximum and minimum education level shows a wide variation of the respondents. About 64.8 per cent respondents of the study area are taken agricultural extension service from Sub Assistance Agriculture Officers and rest of 35.2 per cent do not take one.

6.4.2 Results of *chi-square* test (χ^2)

Chi square test is used to assess the association between different level of education and rice production in the study area. The null hypothesis of this test is there is no relationship between different level of education and rice production. Alternative hypothesis: There is a relationship between different level of education and rice production.

Table 6.38: Impact of education on rice production in the study area

Yield (kg)	Illiterate	Primary	Secondary	Higher Secondary	Tertiary	Total
18-23	9	20	15	4	2	50
23-28	0	8	17	5	11	41
Total	9	28	32	6	13	91
$\chi^2 = 19.914$ df=4 *p* value =0.001						

Source: Field survey, December 2012 and January 2013

Table 6.38 shows the impact of education on rice production in the study area. In Table 6.38, the calculated value of χ^2 is 19.914 and the critical value of χ^2 for 4 degrees of freedom at 0.1% level of significance is 18.467. Since the calculated value of χ^2 is greater than the tabulated value, the null hypothesis can be rejected. So, the alternative hypothesis is accepted at the 0.1% level of significance. It can be said that there is a relationship between the two variables. So, there is evidence of a relationship between rice production and education.

Table 6.39: Impact of illiterate farmers on rice production

Yield(kg)	Illiterate	Others	Total
18-23	9	41	50
23-28	0	41	41
Total	9	82	91
$\chi^2 = 8.190$ df=1 *p* value =0.004			

Source: Field survey, December 2012 and January 2013

Table 6.39 shows the impact of illiterate farmers on rice production in the study area. In Table 6.39, the calculated value of χ^2 is 8.190 and the critical value of χ^2 for 1 degrees of freedom at 1% level of significance is 6.635. Since 8.190 > 6.635, the null hypothesis can be rejected, and the alternative hypothesis is accepted at the 1% level of significance. That is to say, there is relationship between the two variables. So, there is evidence of a relationship between rice production and illiterate farmer. This is because; the experience of the illiterate farmers is higher than others.

Table 6.40: Experience of the respondent

Level of Education	No. of cultivators	Percentage	Mean	Standard Deviation	Minimum	Maximum
Illiterate	9	9.9	20.5556	8.7050	10.00	38.0
Primary	28	30.8	19.8929	9.70811	8.00	52.00
Secondary	32	35.2	10.3011	12.3391	6.00	45.0
Higher Secondary	9	9.9	11.1111	4.7287	6.00	20.0
Tertiary	13	14.3	12.5385	10.2112	2.00	30.0
Total	91	100.0	17.6813	9.92178	2.00	52.00

Source: Field survey, December 2012 and January 2013

Table 6.40 shows that the average experience of illiterate rice farmer is 20.55 years, which is
higher than others. The lowest average experience is 10.30 years, which consist on secondary level.

Table 6.41: Impact of primary education on rice production

Yield(kg)	Primary level	Others	Total
18-23	20	30	50
23-28	8	33	41
Total	28	63	91
$\chi^2 = 4.439$ df=1 p value =0.035			

Source: Field survey, December 2012 and January 2013

Table 6.41 shows the impact of Primary education on rice production. In Table 6.41, the calculated value of χ^2 is 4.439 and the critical value of χ^2 for 1degrees of freedom at 5% level of significance is 3.841. Since 4.439 > 3.841, the null hypothesis can be rejected, and the alternative hypothesis is accepted. It can be said that there is a relationship between the two variables. At the 5% level of significance, there is evidence of a relationship between rice production and primary of education.

Table 6.42: Impact of secondary education on rice production

Yield(kg)	Secondary level	Others	Total
18-23	15	35	50
23-28	17	24	41
Total	32	59	91
$\chi^2 = 1.298$ df=1 p value =0.254			

Source: Field survey, December 2012 and January 2013

Table 6.42 shows the impact of secondary education on rice production. In Table 6.42, the calculated value of χ^2 is 1.298 and it is statistically insignificant. So, the null hypothesis can be accepted. It can be said that there is no relationship between rice production and secondary level of education in this village.This is because the experience of secondary educated farmers of this village is 10.30 years.

Table 6.43: Impact of higher secondary level of education on rice production

Yield(kg)	Higher Secondary Level	Others	Total
18-23	6	44	50
23-28	3	38	41
Total	9	82	91
$\chi^2 = 0.554$ df=1 p value = 0.457			

Source: Field survey, December 2012 and January 2013

Table 6.43 shows the impact of higher secondary education on rice production. In Table 6.43, the calculated value of χ^2 is 0.554 and statistically insignificant. So, the null hypothesis can be accepted. That is to say, there is no relationship between rice production and higher secondary level of education in this village.This is because the experience of higher secondary educated farmers of this village is 11.11 years.

Table 6.44: Impact of tertiary level of education on rice production

Yield(kg)	Tertiary level	Others	Total
18-23	2	48	50
23-28	11	30	41
Total	13	78	91
$\chi^2 = 9.588$ df=1 p value =0.002			

Source: Field survey, December 2012 and January 2013

Table 6.44 shows the impact of tertiary level education on rice production in the study area. In Table 6.44, the calculated value of χ^2 is 9.588 and the critical value of χ^2 for 1 degrees of freedom at 0.2 % level of significance is 9.550. Since 9.588 > 9.550, the null hypothesis can be rejected, and the alternative hypothesis is accepted at the 0.2% level of significance. That is to say, there is a relationship between the two variables. So, there is evidence of a relationship between rice production and tertiary level of education.

6.4.3 Analysis of variance

One way ANOVA is used to examine mean differences between two or more groups. It is a bivariate test with one independent variable and one dependent variable. The independent variable must be categorical and the dependent variable must be continuous. This analysis is based on 91 randomly selected variables with no missing observations.

Table 6.45: Descriptive of the independent variable

Yield Level of education	N	Mean	Standard Deviation	Standard Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Illiterate	9	19.7503	0.83716	0.27905	19.1068	20.3938	18.26	21.21
Primary	28	22.5690	0.97893	0.18500	22.1894	22.9486	20.00	24.24
Secondary	32	23.5360	1.58512	0.28021	22.9645	24.1075	20.54	26.42
Higher Secondary	9	23.2689	1.45614	0.48538	22.1496	24.3882	21.21	25.51
Tertiary	13	25.0301	1.79814	0.49872	23.9435	26.1167	20.72	26.77
Total	91	23.0511	1.91069	0.20029	22.6532	23.4490	18.26	26.77

Source: Field survey, December 2012 and January 2013

The output includes in Table 6.45 that provides descriptive statistics for each of the factors for the independent variable. As the report indicates, the means of standardized reading scores for the Illiterate, Primary Education, Secondary Education, Higher Secondary Education and Tertiary Education are different from other groups. The results suggest that these groups may differ with regard to their mean. These differences, especially among the independent variables, warrant an ANOVA. The null hypothesis is that the means of various level of education group are equal.

Table 6.46: ANOVA of yield and level of education

Level of education \ Yield	Sum of Squares	df	Mean Square	*F*	*P* value
Between Groups	163.432	4	40.858	21.278	0.000
Within Groups	165.134	86	1.920		
Total	328.566	90			

Source: Field survey, December 2012 and January 2013

The results of the overall *F* test in the ANOVA summary Table 6.46 can be examined to determine whether group means are different or not. As indicated, the overall *F* test is significant (i.e., p value < 0.05), indicating that means between groups are not equal. Because, they reject the null hypothesis.

6.4.4 Empirical results

The empirical results of the production function in equation (3) presented in Table 6.47. In Table 6.47, the findings show that the input cost of production is insignificant and the coefficient of input cost of production is -0.04171. The results indicate that as input cost of production increases by Tk.1 with output decreases by -0.04171 kilogram. The labour cost of production is statistically insignificant .The coefficient of labour cost of production is-0.01886. The results indicate that if the labour cost of production increases by Tk.1, then the total output decreases by -0.01886 kilogram. The cultivable land is statistically highly significant. The coefficient of cultivable land is 1.036307.The results indicate that the cultivable land increases by 1 decimal, total production increases by 1.036307 kilogram per decimal.

The coefficient of illiterate farmer is 3.39816, which is highly significant. This is because, if the farmers experience increases, their total output increases. In this study, the level of experience is the highest of illiterate rice cultivators. The coefficient of primary education is (3.39816+ 0.019204) = 3.417364, which is statistically insignificant. It indicates that if the primary education of farmer increases, their total output increases by 3.417364 kilogram.

Table 6.47: Empirical results of multiple regressions

	$\hat{\beta}$	St. Error	*t*	*P* value	Eigenvalue	Tolerance	VIF
1	2	3	4	5	6	7	8
Intercept	3.39816*	0.20196	16.825	0.0000	5.635941	-	-
Input cost(k)	-0.04171	0.10232	-0.4076	0.6845	1.165418	0.007798	128.242
Labor cost(L)	-0.01886	0.11142	-0.1693	0.8659	1.000053	0.006153	162.513
Cultivable Land(T)	1.036307*	0.04695	22.0682	0.0000	0.736965	0.034144	29.2874
Primary(D_1)	0.019204	0.01404	1.36737	0.1752	0.313616	0.526374	1.89978
Secondary(D_2)	0.043938**	0.01426	3.08101	0.0028	0.143514	0.476987	2.09649
Higher Secondary(D_3)	0.023771	0.01631	1.45697	0.1489	0.004353	0.932304	1.07261
Tertiary(D_4)	0.096385*	0.01735	5.55241	0.0000	0.000127	0.599387	1.66837
Extension Service(S)	0.108823*	0.01199	9.06957	0.0000	0.000012	0.673798	1.48412
R^2		0.9930					
Adjusted R^2		0.9923					
Mean square error		180.66					

Source: Field survey, December 2012 and January 2013

* Highly significant**1% level of significant

The coefficient of secondary education is (3.417364 + 0.043938) =3.461302, which is statistically significant. If the secondary education of farmer increases, their total output increases by 3.461302 kilogram. The coefficient of higher secondary education is (3.461302+ 0.023771) = 3.485073, which is insignificant. If the higher secondary education of farmer increases, their total output increases by 3.485073 kilogram. The coefficient of tertiary education is (3.485073 + 0.096385) =3.581458, which is highly significant. If the tertiary education of farmer increases, their total output increases by 3.581458 kilogram. The coefficient of extension service is 0.108823 and it is statistically highly significant. It indicates that if the extension service increases, their total output increases by 0.108823 kilogram.

In Table 6.47, four variables of this model provide insignificant results and two is opposite sign. So, this model might suffer from multicollinearity problem.

6.4.5 Reliability and validity

To ensure the reliability of the questionnaire Cronbach's alpha test has been used in this study. The result of Cronbach's alpha test is given in Table 6.48.

Table 6.48: Test of reliability

Number of observation	Number of items	Cronbach`s Alpha
91	9	0.7651

Source: Field survey, December 2012 and January 2013

In Table 6.48, it is observed that Cronbach`s alpha is 0.7651 which indicates a high level of internal consistency for our scale with this specific sample.

In this study, variables and questions are drawn from literature, which ensured the validity of the questionnaire (Ali and Noman, 2013).

6.4.6 Diagnostic of the model

To check the reliability of the above results, the diagnosis of normality, multicollinearity, heteroscedasticity and autocorrelation are essential. For our postulated model, following the rule of thumb multicollinearity is not a troublesome problem. Again, to judge the validity of the above-mentioned results, though not predictable for cross-section data, the test for presence or absence of autocorrelation or serial correlation and heteroscedasticity have been conducted.

6.4.7 Normality test

Normality is achieved by Jarque-Bera test. The null hypothesis of the JB test is that the residuals are normally distributed and alternative hypothesis is not normally distributed.

6.49: Jarque-Bera normality test

Skewness	Kurtosis	Jarque-Bera statistic	*P* value
0.220	2.712	1.051	0.5911

Source: Field survey, December 2012 and January 2013

In Table 6.49, the Jarque-Bera statistics is 1.051 and the corresponding *p*- value is 0.5911. Since the *p* value is more than 5 per cent, we accept null hypothesis meaning that the population residual is normally distributed which fulfills the assumption of a good regression line.

6.4.8 Fit of the model

Table 6.50: Analysis of variance (ANOVA)

	Sums of Squares	df	Mean Squares	*F*	*P* value
Regression	23.62847	8	2.953558	1467.527	0.000
Residual	0.165034	82	0.002013		
Total	23.7935	90			

Source: Field survey, December 2012 and January 2013

Table 6.50, ANOVA summarizes how much of the variance in the data (total sum of squares) are accounted for by the factor effect (factor sum of squares) and how much is random error (residual sum of squares). From Table 6.50, *F* value is 1467.527and the *p* value is 0.000. This indicates that the results obtained from regression output are highly significant. Therefore, it is clear that the model is a better fit statistically.

6.4.9 Heteroscedasticity

Heteroscedasticity is obtained by white heteroscedasticity test. The null hypothesis of white test is that there is no heteroscedasticity and alternative hypothesis is that there is heteroscedasticity in the error term.

Table 6.51: White heteroscedasticity test

		P value
F Statistic	1.767409	0.074015
Obser* R^2	17.97186	0.082236

Source: Field survey, December 2012 and January 2013

It can be observed from Table 6.51that there is no heteroscedasticity in the error term of the model. The result is confirmed by White heteroscedaticity test. Obser*R^2=17.97186which has, asymptotically, a chi-square distribution with 44 df. The 10% critical chi-square value for 44 df is 56.369. Since the calculated value of chi-square is less than the critical value at 5% level of significance, it can be said that there is no heteroscedasticity in the error term of the model.

6.4.10 Autocorrelation

The Breusch-Godfrey serial correlation and Durbin-Watson statistic have been used to test for presence of serial correlation among the residuals. The null hypothesis of the Breusch-Godfrey serial correlation LM test is that there is no serial correlation in the residuals and alternative hypothesis is that there is serial correlation in the residuals. From Table 6.52, the*p*-value (0.474839) of Obs*R-squared is more than 5 percent ($p > 0.05$), we cannot reject null hypothesis meaning that residuals (u) are not serially correlated which is desirable. So, it can be said that the model is free from autocorrelation. The value of the Durbin-Watson statistic ranges from 0 to 4(Gujarati, 2003). As a general rule of thumb, the residuals are not correlated if the Durbin-Watson statistic is approximately 2 and an acceptable range is 1.50 to 2.50 (Alam et.al, 2013).

Table 6.52: Breusch-Godfrey serial correlation LM and Durbin-Watson tests

		P value
F Statistics	0.665648	0.516764
Obser**R-squared*	1.489560	0.474839
d Statistic (DW)	2.116	

Source: Field survey, December 2012 and January 2013

In Table 6.52, the value of *d* statistic is 2.116, which is about to 2. It indicates that there is no serial correlation.

6.4.11 Multicollinearity

Table 6.47 shows that there are three of eigenvalues close to zero and three VIF`s values more than 5. These results indicate that the model suffers from multicollinearity. It can also be found that the value of R^2 and adjusted R^2 are very high.

6.4.12 Results of ridge regression

Ridge regression has been applied to overcome the problem of multicollinearity. The ridge regression results have been shown in Table 6.53.

All VIF values are less than 5 which are shown in Table 6.53. These results indicate that this model is free from multicollinearity problems. It also shows the different results between Table 6.47 and Table 6.53. All variables are statistically significant except primary, secondary and higher secondary education in Table 6.53.

The coefficient of input cost of production is 0.249436 and it is statistically highly significant. The results indicate that as input cost of production increases by Tk.1, output increases by 0.249436 kilogram.The coefficient of labour cost of production is 0.260582 it is statistically highly significant. The results indicate that if the labour cost of production increases by Tk.1, then output increases by 0.260582 kilogram. The coefficient of cultivable land is 0.414234 it is statistically highly significant. The results indicate that the cultivable land increases by 1 decimal, total production increases by 0.414234 kilogram per decimal.

Table 6.53: Empirical results of ridge regression

	$\hat{\beta}$	St. Error	t	P value	Tolerance	VIF
1	2	3	4	5	6	7
Intercept	1.347967*	0.337749	3.991028	0.000142	-	-
Input cost(K)	0.249436*	0.062601	3.984504	0.000146	0.207162	4.827137
Labor cost(L)	0.260582*	0.061627	4.228386	0.000060	0.200052	4.998712
Cultivable Land(T)	0.414234*	0.057946	7.148651	0.000000	0.223	4.484306
Primary (D1)	0.017423	0.037301	0.467077	0.641683	0.742068	1.347586
Secondary (D2)	0.039097	0.037326	1.047451	0.297969	0.692414	1.444223
Higher Secondary (D3)	0.041371	0.048025	0.861443	0.391506	1.070036	0.934548
Tertiary (D4)	0.083646**	0.046966	1.780975	0.07862	0.814289	1.228066
Extension Service	0.056097**	0.03325	1.687118	0.095381	0.872571	1.146039
R^2	0.9310					
Adjusted R^2	0.9243					
Mean square error	6.80					
k(Ridge parameter)	0.13000					

Source: Field survey, December 2012 and January 2013

* Highly significant **10% level of significant

The coefficient of illiterate farmer is 1.347967, which is highly significant. This is because, if the farmers experience increases, their total output increases. In this study, the level of experience is the high of illiterate rice farmers. The coefficient of primary education is (1.347967+ 0.017423) = 1.36539, which is insignificant. It indicates that if the primary education of farmer increases, their total output increases by 1.36539 kilogram. The coefficient of secondary education is (1.36539+ 0.039097) = 1.404487, which is insignificant. If the secondary education of farmer increases, their total output increases by 1.404487 kilogram.

The coefficient of higher secondary education is (1.404487 + 0.041371) =1.445858, which is insignificant. If the higher secondary education of farmer increases, their total output increases by 1.445858 kilogram. The coefficient of tertiary education is

(1.445858 + 0.083646) =1.529504, which is significant. If the tertiary education of farmer increases, their total output increases by 1.529504 kilogram.
The coefficient of extension service is 0.056097 and it is statistically significant. The results indicate that the extension service increases, total production increases.

6.4.13 Fit of the overall model

Table 6.54: Analysis of variance of ridge regression

	Sums of Squares	df	Mean Squares	*F*	*P* value
Regression	22.15231	8	2.769039	138.3513	0.000
Residual	1.641192	82	0.020015		
Total	23.7935	90			

Source: Field survey, December 2012 and January 2013

Table 6.54, analysis of variance of ridge regression summarizes how much of the variance in the data (total sum of squares) are accounted for by the factor effect (factor sum of squares) and how much is random error (residual sum of squares). In Table 6.54, *F* value is 138.35 and the overall results are highly significant.

6.4.14 Comparison between the ordinary least squares (OLS) method and ridge regression method

It has been justified that ridge regression is better than OLS method. Table 6.47 shows the results of OLS. In Table 6.47, coefficient of determination (*R*-squared) is 0.9930, adjusted R^2 is 0.9923, 2nd column shows the OLS estimator of $\hat{\beta}$, Eigen values found column 6th, VIF is in column 8th. Here maximum VIF is 162.513, which indicates greater multicollinearity. It is observed that *R*-squared and adjusted *R*-squared is very high, and least squares estimates are unstable. The predictor variables are correlated so ridge regression techniques can be applied to stable set of correlation.
Table 6.53 shows the results of ridge regression. In Table 6.53, the coefficient of determination(*R*-squared) is 0.9310, adjusted R^2 is 0.9243, 2nd column shows the ridge estimator of $\hat{\beta}_R$, VIF is in column 7^{th},it also observed that *R*-squared and adjusted *R*-squared are less than OLS, and ridge estimates are stable than OLS estimates.
It is found from Table 6.47 and Table 6.53, the tolerance of OLS estimates is less than the tolerance of ridge estimates. As a result, the VIF for ridge estimates is less than the

VIF for OLS estimates. These results indicate that the ridge regression method is better than OLS as it is clear from the Table 6.47 and Table 6.53.

It is estimated $\hat{\beta}$ using OLS estimator and estimate $\hat{\beta}_R$ using ridge estimator with different choices of k from a grid (0.01, 0.02..., and 0.13). It is computed mean square error for OLS estimator and mean square error for ridge regression estimator. From Table 6.47 and Table 6.53 MSE for OLS is greater than the MSE for ridge regression. This result indicates that ridge estimator performs better than the OLS estimator does.

6.4.15 Discussion

Table 6.53, the input cost of production has positive and significant (*p* value = 0.000146) effect on rice production. If the input cost of production increases, output increases. The same results in line with Appleton & Balihuta (1996) and Weir (1999). Table 6.53 shows that the labour cost of production has positive and significant (*p* value = 0.000060) effect on rice production. The results indicate that if the labour cost of production increases, and then output increases. The findings were consistent with studies by Cotlear (1986), Appleton & Balihuta (1996), Yang (1997) and Weir (1999). Table 6.53 reveals that the cultivable land has positive and significant (*p* value = 0.00000) effect on rice production. The results indicate that the cultivable land increases, then total production increases. The same results were found by Cotlear (1986), Appleton & Balihuta (1996), Yang (1997), Weir (1999) and Rehman et al. (2012).

Table 6.38 the *Chi-Square* results show that there is a significant (*p* value = 0.001) relationship between education and rice production in the study area. In Table 6.46 the analysis of variance (ANOVA) also shows that there are significant (*p* value = 0.000) relationships among the means of the different level of education. ANOVA results suggest that the different level of educational group may differ with regard their mean scores in Table 6.45. Years of schooling are found to be a significant impact on rice production, with one additional year of education yielding an increase in rice yields.

Table 6.41 the *Chi-Square* result shows that there is a significant (*p* value = 0.035) relationship between primary education and rice yield. Similar results are found in ANOVA. In Table 6.53 ridge regression result shows that the primary education has positive but insignificant (*p* value = 0.641683) effect on rice production. So, the level of primary education of the rice cultivators increases, their total output increases.

Table 6.42 the *Chi-Square* results shows that there is no significant (*p* value = 0.254) relationship between secondary education and rice yield. But in Table 6.46 the analysis of variance (ANOVA) shows that there are significant (*p* value = 0.000) relationships among the means of the different level of education. Ridge regression results show that the secondary education has positive and insignificant (*p* value = 0.297969) impact on rice production. So, the level of secondary education of the rice cultivators increases, their total output increases.

Table 6.43 the *Chi-Square* results shows that there is no significant (*p* value = 0.457) relationship between higher secondary education and rice yield. But in Table 6.46 the analysis of variance (ANOVA) shows that there are significant (*p* value = 0.000) relationships among the means of the different level of education. Ridge regression results reveal that higher secondary education has positive and (*p* value = 0.391506) insignificant impact on rice production. So, the level of higher secondary education of the rice cultivators increases, their total output increases also.

Table 6.44 the *Chi-Square* results shows that there is a significant (*p* value = 0.002) relationship between tertiary education and rice yield. Similar results are found in ANOVA. Ridge regression results show that tertiary education has positive and significant (*p* value = 0.07862) effect on rice production. So, the level of tertiary education of the rice cultivators increases, their total output increases. The similar results were found by Pudasaini (1983) and Gemmell (1996).

In Table 6.53, shows that the extension service has positive and significant (*p* value = 0.095381) effects in improving rice productivity. It is clarified that the more the extension contacts between extension agents and farmers leads to the higher the productivity. The similar results were found by Huffman (1974),Haq (2011) and Nargis & Lee (2013).

6.5 Overall results and discussion

6.5.1 Descriptive statistics

Table 6.55 shows the variables that are used in estimations and their sample statistics namely maximum and minimum values, mean and standard deviation.

The mean, standard deviation, minimum and maximum of the variables have been presented in Table 6.55. In Table 6.55, it is found that the average yield of rice is 22.53 kilograms with maximum average yield of 26.77 kilograms and minimum average yield of 18.26 kilograms. The average value of input cost is Tk. 67.46 with maximum and minimum average value of input cost is Tk. 75.76 and Tk. 50.88 respectively.

Table 6.55: Descriptive statistics of the variables

Variable	Unit of Measurement	No. of cultivators	Minimum	Maximum	Mean	Standard Deviation
Output	Kg	358	900.00	7600.00	3421.4665	1737.96792
Yield	Kg(per decimal)	358	18.26	26.77	22.5345	1.77997
Input cost	Tk.	358	2550.00	19525.00	10149.4609	4836.47994
Input cost	Tk.(per decimal)	358	50.88	75.76	67.4614	7.17910
Labour cost	Tk.	358	2100.00	15800.00	8067.2654	4055.48180
Labour cost	Tk.(per decimal)	358	40.91	60.61	52.9838	5.57272
Cultivable land	decimal	358	49.00	330.00	152.8380	78.81766
Education	years of schooling	358	0.00	16.00	6.3184	4.60352
Extension service	percentage	Yes=58.7 No=41.3				

Source: Field survey, December 2012 and January 2013

The average value of labour cost is Tk. 52.98 with the maximum and minimum average value of labour cost is Tk. 60.61 and Tk. 40.91 respectively. The average of cultivable land is 152.83 decimal with the maximum and minimum of the cultivable land is 330 decimal and 49 decimal respectively. In Table 6.55, the average level of education of the respondent is 6.31 years and the standard deviation of the education level of the respondent is 4.60 years. Maximum education level of the respondent is 16 years and minimum is 0.00 years. Maximum and minimum education level shows a wide variation of the respondents. About 58.7 per cent respondents of the study area are taken agricultural extension service from Sub Assistance Agriculture Officers and rest of 41.3 per cent do not take one.

6.5.2 Results of *chi-square* (χ^2) test

Chi square test is used to assess the association between different level of education and rice production in the study area. The null hypothesis of this test is that there is no relationship between different level of education and rice production. Alternative hypothesis: There is a relationship between different level of education and rice production.

Table 6.56: Impact of education on rice production in the study area

Yield (kg)	Illiterate	Primary	Secondary	Higher Secondary	Tertiary	Total
18-23	52	94	68	8	11	233
23-28	0	24	60	18	23	125
Total	52	118	128	26	34	358
$\chi^2 = 76.487$ df=4 *p* value =0.000						

Source: Field survey, December 2012 and January 2013

Table 6.56 shows the impact of education on rice production in the study area. In Table 6.56, the calculated value of χ^2 is 76.487 and the critical value of χ^2 for 4 degrees of freedom at 0.1% level of significance is 18.467. Since the calculated value of χ^2 is greater than the tabulated value, the null hypothesis can be rejected. So, the alternative hypothesis is accepted at the 0.01% level of significance. It can be said that there is a relationship between the two variables. So, there is evidence of a relationship between rice production and education.

Table 6.57: Impact of illiterate farmers on rice production

Yield(kg)	Illiterate	Others	Total
18-23	52	181	233
23-28	0	125	125
Total	52	306	358
$\chi^2 = 32.638$ df=1 *p* value =0.000			

Source: Field survey, December 2012 and January 2013

Table 6.57 shows the impact of illiterate farmers on rice production in the study area. In Table 6.57, the calculated value of χ^2 is 32.638 and the critical value of χ^2 for 1 degrees of freedom at 0.1 % level of significance is 10.827. Since 32.638 > 10.827, the null hypothesis can be rejected, and the alternative hypothesis is accepted at the 0.1%

level of significance. That is to say, there is relationship between the two variables. So, there is evidence of a relationship between rice production and illiterate rice farmer. This is because; the experience of the illiterate farmers is higher than others.

Table 6.58: Experience of the respondent

Level of Education	No. of cultivators	Percentage	Mean	Standard Deviation	Minimum	Maximum
Illiterate	52	14.5	29.5000	13.1931	7.00	58.0
Primary	118	33.0	23.6695	11.86466	5.00	58.0
Secondary	128	35.8	22.0156	11.7453	4.00	58.0
Higher Secondary	26	7.3	11.0000	5.7758	4.00	26.0
Tertiary	34	9.5	11.2059	8.9299	2.00	30.0
Total	358	100	21.8212	12.6005	2.00	58.0

Source: Field survey, December 2012 and January 2013

Table 6.58 shows that the average experience of illiterate rice farmer is 29.50 years, which is
higher than others. The lowest average experience is 11.00 years, which consist on higher secondary level.

Table 6.59: Impact of primary education on rice production

Yield(kg)	Primary level	Others	Total
18-23	94	139	233
23-28	24	101	125
Total	118	240	358
$\chi^2 = 16.459$ df=1 p value =0.000			

Source: Field survey, December 2012 and January 2013

Table 6.59 shows the impact of Primary education on rice production. In Table 6.59, the calculated value of χ^2 is 16.459 and the critical value of χ^2 for 1degrees of freedom at 0.1% level of significance is 10.827. Since 16.459 >10.827, the null hypothesis can be rejected, and the alternative hypothesis is accepted. It can be said that there is a relationship between the two variables. At the 0.1% level of significance, there is evidence of a relationship between rice production and primary of education.

Table 6.60: Impact of secondary education on rice production

Yield(kg)	Secondary level	Others	Total
18-23	68	165	233
23-28	60	65	125
Total	128	230	358
$\chi^2 = 12.538$ df=1 p value =0.000			

Source: Field survey, December 2012 and January 2013

Table 6.60 shows the impact of secondary education on rice production. In Table 6.60, the calculated value of χ^2 is 12.538 and the critical value of χ^2 for 1 degrees of freedom at 0.1% level of significance is 10.827. Since 12.538 >10.827, the null hypothesis can be rejected, and the alternative hypothesis is accepted. It can be said that there is a relationship between the two variables. At the 0.1% level of significance, there is evidence of a relationship between rice production and secondary level of education.

Table 6.61: Impact of higher secondary level of education on rice production

Yield(kg)	Higher Secondary Level	Others	Total
18-23	10	223	233
23-28	16	109	125
Total	26	332	358
$\chi^2 = 8.744$ df=1 p value = 0.003			

Source: Field survey, December 2012 and January 2013

Table 6.61 shows the impact of higher secondary education on rice production. In Table 6.61, the calculated value of χ^2 is 8.744 and the critical value of χ^2 for 1 degrees of freedom at 1% level of significance is 6.635. Since 8.744 > 6.635, the null hypothesis can be rejected, and the alternative hypothesis is accepted at the 1% level of significance. That is to say, there is a relationship between the two variables. So, there is evidence of a relationship between rice production and higher secondary level of education.

Table 6.62: Impact of tertiary level of education on rice production

Yield(kg)	Tertiary level	Others	Total
18-23	11	222	233
23-28	23	102	125
Total	34	324	358
$\chi^2 = 17.711$ df=1 p value =0.000			

Source: Field survey, December 2012 and January 2013

Table 6.62 shows the impact of tertiary level education on rice production in the study area. In Table 6.62, the calculated value of χ^2 is 17.711 and the critical value of χ^2 for 1 degrees of freedom at 0.1 % level of significance is 10.827. Since 17.711>10.827, the null hypothesis can be rejected, and the alternative hypothesis is accepted at the 0.1% level of significance. That is to say, there is a relationship between the two variables. So, there is evidence of a relationship between rice production and tertiary level of education.

6.5.3 Analysis of variance

One way ANOVA is used to examine mean differences between two or more groups. It is a bivariate test with one independent variable and one dependent variable. The independent variable must be categorical and the dependent variable must be continuous. This analysis is based on 358 randomly selected variables with no missing observations.

The output includes in Table 6.63 that provides descriptive statistics for each of the factors for the independent variable. As the report indicates, the means of standardized reading scores for the Illiterate, Primary Education, Secondary Education, Higher Secondary Education and Tertiary Education are different from other groups. The results suggest that these groups may differ with regard to their mean. These differences, especially among the independent variables, warrant an ANOVA. The null hypothesis is that the means of various level of education group are equal.

Table 6.63: Descriptive of the independent variable

Yield Level of education	N	Mean	Standard Deviation	Standard Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Illiterate	52	19.5758	0.83053	0.11517	19.3446	19.8070	18.26	21.21
Primary	118	22.5397	0.85025	0.07827	22.3847	22.6947	20.00	24.44
Secondary	128	23.0942	1.37185	0.12126	22.8542	23.3341	19.70	26.42
Higher Secondary	26	23.5892	1.21591	0.23846	23.0981	24.0803	21.21	25.51
Tertiary	34	24.1279	1.94464	0.33350	23.4494	24.8064	20.72	26.77
Total	358	22.5345	1.77997	0.09407	22.3495	22.7195	18.26	26.77

Source: Field survey, December 2012 and January 2013

The results of the overall *F* test in the ANOVA summary Table 6.64 can be examined to determine whether group means are different or not. As indicated, the overall *F* test is significant (i.e., p value < 0.05), indicating that means between groups are not equal. Because, they reject the null hypothesis.

Table 6.64: ANOVA of yield and level of education

Level of education \ Yield	Sum of Squares	df	Mean Square	*F*	*P* value
Between Groups	610.551	4	152.638	103.513	0.000
Within Groups	520.525	353	1.475		
Total	1131.076	357			

Source: Field survey, December 2012 and January 2013

6.5.4 Empirical results

The empirical results of the production function in equation (3) presented in Table 6.65. In Table 6.65, the findings show that the input cost of production is insignificant and the coefficient of input cost of production is 0.00671. The results indicate that as input cost of production increases by Tk.1 with output increases by 0.00671 kilogram. The labour cost of production is statistically insignificant .The coefficient of labour cost of production is -0.0126. The results indicate that if the labour cost of production increases by Tk.1, then the total output decreases by -0.0126kilogram. The cultivable land is statistically highly significant. The coefficient of cultivable land is 0.995549.The results indicate that the cultivable land increases by 1 decimal, total production increases by 0.995549 kilogram per decimal.

The coefficient of illiterate farmer is 3.063489, which is highly significant. This is because, if the farmers experience increases, their total output increases. In this study, the level of experience is the highest of illiterate rice cultivators. The coefficient of primary education is (3.063489+0.056992) = 3.120481, which is highly significant. It indicates that if the primary education of farmer increases, their total output increases by 3.120481 kilogram. The coefficient of secondary education is (3.120481+0.080962) =3.201443, which is highly significant. If the secondary education of farmer increases, their total output increases by 3.201443 kilogram.

Table 6.65: Empirical results of multiple regressions

	$\hat{\beta}$	St. Error	t	P value	Eigenvalue	Tolerance	VIF
1	2	3	4	5	6	7	8
Intercept	3.063489*	0.091764	33.384	0.000	5.539388	-	-
Input cost(k)	0.00671	0.039505	0.169	0.865	1.033643	0.01099	90.991
Labor cost(L)	-0.0126	0.039364	-0.320	0.749	1.000343	0.010237	97.682
Cultivable Land(T)	0.995549*	0.021458	46.395	0.000	0.926455	0.036072	27.722
Primary(D_1)	0.056992*	0.006879	8.284	0.000	0.372796	0.457093	2.187
Secondary(D_2)	0.080962*	0.00703	11.517	0.000	0.12186	0.421079	2.374
Higher Secondary(D_3)	0.07161*	0.009472	7.559	0.000	0.005364	0.790974	1.264
Tertiary(D_4)	0.118321*	0.009398	12.589	0.000	0.000131	0.629602	1.588
Extension Service(S)	0.096339*	0.005167	18.645	0.000	0.000020	0.738342	1.354
R^2		0.99408					
Adjusted R^2		0.99394					
Mean square error		97.11					

Source: Field survey, December 2012 and January 2013

* Highly significant

The coefficient of higher secondary education is (3.201443+0.07161) =3.273053, which is highly significant. If the higher secondary education of farmer increases, their total output increases by 3.273053 kilogram. The coefficient of tertiary education is (3.273053+0.118321) =3.391374, which is highly significant. If the tertiary education of farmer increases, their total output increases by 3.391374 kilogram. The coefficient of extension service is 0.096339 and it is statistically significant. It indicates that if the extension service increases, their total output increases by 0.096339 kilogram.

In Table 6.65, two variables of this model provide insignificant results and one is opposite sign. So, this model may suffer from multicollinearity problem.

6.5.5 Reliability and validity

To ensure the reliability of the questionnaire Cronbach's alpha test has been used in this study. The result of Cronbach's alpha test is given in Table 6.66.

Table 6.66: Cronbach`s alpha reliability test

Number of observation	Number of items	Cronbach`s Alpha
358	9	0.7870

Source: Field survey, December 2012 and January 2013

In Table 6.66, it is observed that Cronbach`s alpha is 0.7870 which indicates a high level of internal consistency for our scale with this specific sample.

In this study, variables and questions are drawn from literature, which ensured the validity of the questionnaire (Ali and Noman, 2013).

6.5.6 Diagnostic of the model

To check the reliability of the above results, the diagnosis of normality, multicollinearity, heteroscedasticity and autocorrelation are essential. For our postulated model, following the rule of thumb multicollinearity is not a troublesome problem. Again, to judge the validity of the above-mentioned results, though not predictable for cross-section data, the test for presence or absence of autocorrelation or serial correlation and heteroscedasticity have been conducted.

6.5.7 Normality test

Normality is achieved by Jarque-Bera test. The null hypothesis of the JB test is that the residuals are normally distributed and alternative hypothesis is not normally distributed.

Table 6.67: Jarque-Bera normality test

Skewness	Kurtosis	Jarque-Bera statistic	*P* value
0.099	3.004	0.5954	0.7425

n Table 6.67, the Jarque-Bera statistics is 0.5954 and the corresponding *p*- value is 0.7425. Since the *p* value is more than 5 per cent, we accept null hypothesis meaning that the population residual is normally distributed which fulfills the assumption of a good regression line.

6.5.8 Fit of the model

Table 6.68: Analysis of variance (ANOVA)

	Sums of Squares	df	Mean Squares	*F*	*P* value
Regression	100.3533	8	12.54417	7330.727	0.000
Residual	0.597201	349	0.001711		
Total	100.9505	357			

Source: Field survey, December 2012 and January 2013

Table 6.68, ANOVA summarizes how much of the variance in the data (total sum of squares) are accounted for by the factor effect (factor sum of squares) and how much is random error (residual sum of squares). From Table 6.68, *F* value is 7330.727 and the *p* value is 0.000. This indicates that the results obtained from regression output are highly significant. Therefore, it is clear that the model is a better fit statistically.

6.5.9 Heteroscedasticity

Heteroscedasticity is obtained by white heteroscedasticity test. The null hypothesis of white test is that there is no heteroscedasticity and alternative hypothesis is that there is heteroscedasticity in the error term.

Table 6.69: White heteroscedasticity test

		P value
F Statistic	2.414949	0.006 661
Obser* R^2	25.52597	0.007630

Source: Field survey, December 2012 and January 2013

It can be observed from Table 6.69 that there is no heteroscedasticity in the error term of the model. The result is confirmed by White heteroscedaticity test. Obser*R^2=25.52597 which has, asymptotically, a chi-square distribution with 44 df. The 5% critical chi-square value for 44 df is 60.481. Since the calculated value of chi-square is less than the critical value at 5% level of significance, it can be said that there is no heteroscedasticity in the error term of the model.

6.5.10 Autocorrelation

The Breusch-Godfrey serial correlation and Durbin-Watson statistic have been used to test for presence of serial correlation among the residuals. The null hypothesis of the

Breusch-Godfrey serial correlation LM test is that there is no serial correlation in the residuals and alternative hypothesis is that there is serial correlation in the residuals. From Table 6.70, the*p*-value (0.483746) of Obs*R-squared is more than 5 percent ($p > 0.05$), we cannot reject null hypothesis meaning that residuals (u) are not serially correlated which is desirable. So, it can be said that the model is free from autocorrelation. The value of the Durbin-Watson statistic ranges from 0 to 4(Gujarati, 2003). As a general rule of thumb, the residuals are not correlated if the Durbin-Watson statistic is approximately 2 and an acceptable range is 1.50 to 2.50 (Alam et.al, 2013).

Table 6.70: Breusch-Godfrey serial correlation LM and Durbin-Watson tests

		P value
F Statistics	0.706749	0.493954
Obser**R-squared*	1.452391	0.483746
d Statistic (DW)	1.960	

Source: Field survey, December 2012 and January 2013

In Table 6.70, the value of *d* statistic is 1.960, which is about to 2. It indicates that there is no serial correlation.

6.5.11 Multicollinearity

Table 6.65 shows that there are three of eigenvalues close to zero and three VIF`s values more than 5. These results indicate that the model suffers from multicollinearity. It can also be found that the value of R^2 and adjusted R^2 are very high.

6.5.12 Results of ridge regression

Ridge regression has been applied to overcome the problem of multicollinearity. The ridge regression results have been shown in Table 6.68.

All VIF values are less than 5 which are shown in Table 6.71. These results indicate that this model is free from multicollinearity problems. It also shows the different results between Table 6.65 and Table 6.71. All variables are statistically significant in Table 6.71.

Table 6.71: Empirical results of ridge regression

	$\hat{\beta}$	St. Error	t	P value	Tolerance	VIF
1	2	3	4	5	6	7
Intercept	1.188777*	0.151448	7.849407	0.0000	-	-
Input cost(K)	0.262656*	0.029686	8.847782	0.0000	0.20677	4.836281
Labor cost(L)	0.259342*	0.028691	9.039091	0.0000	0.204726	4.88457
Cultivable Land(T)	0.416597*	0.028041	14.85656	0.0000	0.224411	4.456109
Primary (D1)	0.047597**	0.01819	2.616726	0.009264	0.69459	1.439698
Secondary (D2)	0.072074*	0.018429	3.910972	0.00011	0.650939	1.536243
Higher Secondary (D3)	0.053247***	0.02761	1.928534	0.0546	0.989059	1.011062
Tertiary (D4)	0.098399*	0.02613	3.765689	0.000195	0.865267	1.155712
Extension Service	0.061499*	0.014775	4.162458	0.000039	0.959303	1.042424
R^2	0.9371					
Adjusted R^2	0.9357					
Mean square error	1.66					
K(Ridge parameter)	0.13000					

Source: Field survey, December 2012 and January 2013

* Highly significant **1% level of significant***5% level of significant

The coefficient of input cost of production is 0.262656 and it is statistically highly significant. The results indicate that as input cost of production increases by Tk.1, output increases by 0.262656 kilogram.The coefficient of labour cost of production is 0.259342 it is statistically highly significant. The results indicate that if the labour cost of production increases by Tk.1, then output increases by 0.259342 kilogram. The coefficient of cultivable land is 0.416597 it is statistically highly significant. The results indicate that the cultivable land increases by 1 decimal, total production increases by 0.416597 kilogram per decimal.

The coefficient of illiterate farmer is 1.188777, which is highly significant. This is because, if the farmers experience increases, their total output increases. In this study, the level of experience is the high of illiterate rice farmers. The coefficient of primary education is (1.188777+0.047597) = 1.236374, which is significant. It indicates that if the primary education of farmer increases, their total output increases by 1.236374 kilogram. The coefficient of secondary education is (1.236374+0.072074) =1.308448, which is highly significant. If the secondary education of farmer increases, their total output increases by 1.308448, kilogram. The coefficient of higher secondary education is (1.308448, + 0.053247) =1.361695, which is significant. If the higher secondary education of farmer increases, their total output increases by 1.361695 kilogram. The coefficient of tertiary education is (1.361695+ 0.098399) =1.460094, which is highly significant. If the tertiary education of farmer increases, their total output increases by 1.460094 kilogram. The coefficient of extension service is 0.061499 and it is statistically highly significant. The results indicate that the extension service increases, total production increases.

6.5.13 Fit of the overall model

Table 6.72: Analysis of variance of ridge regression

	Sums of Squares	df	Mean Squares	*F*	*P* value
Regression	94.60583	8	11.82573	650.4913	0.000
Residual	6.344711	349	0.01818		
Total	100.9505	357			

Source: Field survey, December 2012 and January 2013

Table 6.72, analysis of variance of ridge regression summarizes how much of the variance in the data (total sum of squares) are accounted for by the factor effect (factor sum of squares) and how much is random error (residual sum of squares). In Table 6.72, *F* value is 650.49 and the overall results are highly significant.

6.5.14 Comparison between the OLS method and ridge regression method

It has been justified that ridge regression is better than OLS method. Table 6.65 shows the results of OLS. In Table 6.65, coefficient of determination (*R*-squared) is 0.99408, adjusted R^2 is 0.99394, 2nd column shows the OLS estimator of $\hat{\beta}$, Eigen values found column 6th, VIF is in column 8th. Here maximum VIF is 97.682, which indicates greater multicollinearity. It is observed that *R*-squared and adjusted *R*-squared is very high, and least squares estimates are unstable. The predictor variables are correlated so ridge regression techniques can be applied to stable set of correlation.

Table 6.71 shows the results of Ridge regression. In Table 6.71, the coefficient of determination is 0.9371, adjusted R^2 is 0.9357, 2nd column shows the ridge estimator of $\hat{\beta}_R$, VIF is in column 7th,italso observed that *R*-squared and adjusted *R*-squared are less than OLS, and ridge estimates are stable than OLS estimates.

It is found from Table 6.65 and Table 6.71, the tolerance of OLS estimates is less than the tolerance of ridge estimates. As a result, the VIF for ridge estimates is less than the VIF for OLS estimates. These results indicate that the ridge regression method is better than OLS as it is clear from the Table 6.65 and Table 6.71.

It is estimated $\hat{\beta}$ using OLS estimator and estimate $\hat{\beta}_R$ using ridge estimator with different choices of *k* from a grid (0.01, 0.02..., and 0.13). It is computed mean square error for OLS estimator and mean square error for ridge regression estimator. From Table 6.65 and Table 6.71 MSE for OLS is greater than the MSE for ridge regression. This result indicates that ridge estimator performs better than the OLS estimator does.

6.5.15 Discussion

In Table 6.71, the input cost of production has positive and significant (*p* value = 0.0000) effect on rice production. If the input cost of production increases, output increases. The same results in line with Appleton & Balihuta (1996) and Weir (1999). In Table 6.71 shows that the labour cost of production has positive and significant (*p* value = 0.0000) effect on rice production. The results indicate that if the labour cost of production increases, and then output increases. The findings were consistent with studies by Cotlear (1986), Appleton & Balihuta (1996), Yang (1997) and Weir (1999). In Table 6.71 reveals that the cultivable land has positive and significant (*p* value = 0.0000) effect on rice production. The results indicate that the cultivable land increases, then total production increases. The same results were found by Cotlear (1986), Appleton & Balihuta (1996), Yang (1997), Weir (1999) and Rehman et al. (2012).

Table 6.56 the *Chi-Square* results show that there is a significant (*p* value = 0.000) relationship between education and rice production in the study area. In Table 6.64 the analysis of variance (ANOVA) also shows that there are significant (*p* value = 0.000) relationships among the means of the different level of education. ANOVA results suggest that the different level of educational group may differ with regard their mean scores in Table 6.63. Years of schooling are found to be a significant impact on rice production, with one additional year of education yielding an increase in rice yields.

Table 6.59 the *Chi-Square* result shows that there is a significant (*p* value = 0.000) relationship between primary education and rice yield. Similar results are found in ANOVA. In Table 6.71 ridge regression result also shows that the primary education has positive and significant (*p* value = 0.009264) effect on rice production. So, the level of primary education of the rice cultivators increases, their total output increases. The similar results were found by Singh (1974), Dominique van de Walle (2003), Onphanhdala (2009) and Haq (2012).

Table 6.60 the *Chi-Square* results shows that there is a significant (*p* value = 0.000) relationship between secondary education and rice yield. ANOVA also supports this result. Ridge regression results show that the secondary education has positive and significant (*p* value = 0.00011) impact on rice production. So, the level of secondary education of the rice cultivators increases, their total output increases. The similar results were found by Singh (1974) and Asadullah & Rahman (2006).

Table 6.61 the *Chi-Square* results shows that there is a significant (*p* value = 0.003) relationship between higher secondary education and rice yield. ANOVA also confirm this result. Ridge regression results reveal that higher secondary education has positive and (*p* value = 0.0546) significant impact on rice production. So, the level of higher secondary education of the rice cultivators increases, their total output increases also. The similar result was found by Pudasaini (1983).

Table 6.62 the *Chi-Square* results shows that there **i**s a significant (*p* value = 0.000) relationship between tertiary education and rice yield. Similar results are found in ANOVA. Ridge regression results show that tertiary education has positive and significant (*p* value = 0.000195) effect on rice production. So, the level of tertiary education of the rice cultivators increases, their total output increases. The similar results were found by Pudasaini (1983) and Gemmell (1996).

Table 6.71reveals that the coefficient of primary, secondary, higher secondary and tertiary levels of education. Tertiary level of education remains statistically highly significant (*p* value = 0.000195) and the coefficient of this variable value is higher than other levels of education. This indicates that higher educated farmer can easily introduce the modern technology and technique namely HYV seed, soil quality (texture), fertilizer, insecticides, fungicides, herbicides etc. to their field to increase rice production than other levels of farmer.

As shown in Table 6.71, the results indicate that farmer education has a strong and positive effect on rice production. It is also found that the rice production increases with the coefficients of all education variables increase. In the study area majority of farmer, do not complete a primary and secondary education. If they completed primary and secondary education, the rice production would be increased significantly.

In Table 6.71, shows that the extension service has positive and significant (*p* value = 0.000039) effects in improving rice productivity. It is clarified that the more the extension contacts between extension agents and farmers leads to the higher the productivity. The similar results were found by Huffman (1974), Haq (2011) and Nargis & Lee (2013). According to Huffman (1974), extension activities help farmers who did not acquire enough school education to improve their ability to adjust. Similarly, the literate farmers may not be excluded from extension service unless the quality of education is life oriented and this situation prevails in Bangladesh and other developing countries. The difference in the educational background of the farmers keeps influence the effectiveness of extension services and thus magnifies the economic gap between farmers in a vicious cycle (Haq *et al.*, 2003). Therefore, extension service is necessary to provide among all farmers regarding their educational background since extension is a type of education for all.

Empirical studies have been shown that the rice production heavily depends on formal education because education affects individuals' attitudes, economic ambitions, effective working ability and receptiveness to new ideas, indicating a positive relationship between education and labour productivity. Moreover, statistical findings support the fact that the rice production is significantly related to the level of education of the farmers. If farmers are well educated formally or non-formally, they are able to handle their crop management and production, which leads to increase in overall rice production at national level. The main responsibility of the educated farmers in rice production is laid on shoulders of agricultural extension. It is through extension service that new findings could reach the farmer speedily.

Chapter Seven

Summary, Conclusion and Policy Suggestion

7.1 Conclusions

Impact of education on rice production is very important for policy formulation and strategies for the development of agricultural sector. In this study, *Chi Square* test has been used to find out the association between yield of rice and levels of education. The results show that there is a significant association between yield of rice and level of education. Analysis of Variance (ANOVA) is employed to find out the significant difference between mean of the independent groups. The result shows that there are significant differences between mean of the independent groups. A multiple regression model and ridge regression model have been used to estimate the impact of education on rice production. Findings also show that most of the variables are highly with the problem of multicolinearity. In order to overcome this problem ridge regression has been used. The results of ridge regression reveal that the various levels of education have positive and statistically significant effect on rice production. Therefore, the rice production increases with the increase of the level of education of farmer. This result suggests that the level of education of farmer has positive effect on rice production. The input cost of production has positive effect on rice production. The labour cost, cultivable land and extension service have also positive effect on rice production. Education, input cost, labour cost, cultivable land and extension service all these have positive effect on rice production. It is to be notable that the ridge regression models turn out better results than ordinary least squares method during the multicollinearity problem is prevailed in the model in the sense of smaller MSE of estimators, smaller variation for estimates and smaller coefficient of determination.

7.2 Summary of the main findings

Findings show that about 33 percent (Table 5.1) of the respondents who obtained primary education engaged in rice production in the study area.

Findings also show that about 35.8 percent of the respondents who obtained secondary education engaged in rice production in the study area and only 7.3 percent respondents who obtained higher secondary education engaged in rice production in the study area.

The findings of the *Chi Square* of the study show that there is a significant relationship between rice production and education. The study also shows that there is a relationship between rice production and primary education, rice production and secondary education, rice production and higher secondary education, rice production and tertiary education. The findings of the chi-square test show that the impact of primary, secondary as well as tertiary level of education is more significant than the higher secondary level of education on rice production in the study area.

The findings of the ANOVA reveal that there is a significant difference among the mean of the group of independent variables.

The findings of the ridge regression show that there is a positive relationship between rice production and different levels of education. The study also shows that primary education, secondary education, higher secondary education and tertiary education has a positive and significant effect on rice production. It indicates that if the farmers' education of all levels increases, their total output increases. It means that education is relevant with rice production in the study area.

The findings of ridge regression show that the impact of higher secondary level of education is significant and secondary level of education is relatively more significant.

The results also show that input cost has a positive and significant effect on rice production. The results indicate that as input cost of production increases, output increases. It means that input cost is important with rice production in the study area.

The result also shows that labour cost has a positive and significant effect on rice production. The results indicate that if the labour cost of production increases, and then output increases. It means that labour cost is relevant with rice production in the country.

The result also shows that cultivable land has a positive and significant effect on rice production. The results indicate that the cultivable land increases, total production increases. It means that cultivable land is the basic factor with rice production in the country.

The result also shows that the extension service has a positive and significant effect on rice production. It means that extension services are important on rice production for the study area.

7.3 Policy suggestions

The positive impact of education on rice production supports that education is of paramount importance which increases rice production. To boost up rice production of Bangladesh the government should put emphasis on education so that farmers can easily adopt modern agricultural practice. The literacy campaign, training, IPM and adult education programs should be undertaken for the farmers.

Findings confirm that most of primary educated people in the study area are involved in rice production. But there is no agro-oriented course or curricula in the Primary level schools or institutions. There are a few agricultural training institutions in our country. Agro-based courses must be included in the primary level schools or institutions. In addition, number of agricultural institutes must be increased throughout the country, which in turn will increase the number of people with agricultural knowledge. It certainly would have a positive impact on the agricultural productivity.

Findings also confirm that people having the secondary level of education are large in number than that of primary level of education in the study area who are involved in rice production. At present in the secondary level an optional agricultural science subject or course is offered which in our view is very inadequate. Therefore, compulsory courses should be introduced in the secondary and higher secondary levels. It is believed that it will help to increase the number of people with agricultural knowledge. For agricultural technological development, emphasis should be given on research and development activities. For this, agricultural universities, research institutes should be given enough funds or resource.

Thus, policy makers need to consider the influence of higher education and research on the efficiency level of the farmers in formulating policies. The training programmes must include the production of crops, storage of crops, pest control and management, livestock rearing etc. Development of agriculture always requires local research as the development of agriculture depends on many local factors like soil, weather and climate, water, construction of land, behavior of rivers, science and education, local technology cannels, history, indigenous skills, genetic food habit and many other issues.

There are no agriculture-trained teachers as well as institutions for agricultural training in the primary and secondary level. For sustainable growth, environment friendly agriculture must be encouraged. Policy makers would address this issue.

In Bangladesh, the government should take the responsibility of the expansion of agricultural education, research and development. The illiterate farmers do not know about the modern cultivation technique. At first, it needs to expand the general education on the emphasis of agricultural education in the rural area. Now, it is very important to expand the agricultural education and research because educated and trained farmers' show their eagerness to use the modern technology and input.

The lack of education is the barrier of development. Therefore, the government should try to expand education among the farmers. Besides, the government should take proper steps to teach the people practically who are educated in modern agricultural system, HYV seeds, the use of insecticide, fertilizer, irrigation etc., so that their productivity may increase.

To avoid excessive use of seeds, irrigation, fertilizer, pesticides, herbicides and fungicides, farmers' should have a minimum level of education and training. Schooling plays a pivotal role in raising the awareness of farmers' to respond in changing the production pattern.

Day by day land fragmentation is increasing. So, policies should be formulated in such a way that the existing land tenure and land fragmentation system cannot reduce cultivable land. Therefore, farmers' could be educated and proper trained so that they become capable to operate the latest technologies and inputs should be adapted by rice farmers' for reducing the land and environmental degradation, increasing productivity and welfare of rice cultivators in Bangladesh.

Labour cost is a very essential factor, which affects the rice production. Government should be emphasizing the need of education to improve the ability of rice cultivators to receive and understand information regarding modern technology, so that their productivity may increase. Level of education of the rice cultivators' is not high in Bangladesh. However, to increase technical efficiency of rice farmers' education, agricultural credit and training are important factors.

Extension service is important for rice farmers and it has positive effect on rice production. Government should take necessary steps to increase extension service to the farmers.

7.4 Further research

A sample size 358 has been taken in this study. Various data have been collected from individual farmers who are not habituated to keep necessary information in systematic written form. Moreover, some of them are reluctant to provide information. A research with motivated written information can be a further research. There is a scope for further research by using wider data set that provides more accurate outcome.

This study is based on five inputs, that is, input cost, labour cost, cultivable land, education and extension service. If researcher wants to employ more inputs like HYV (High Yielding Variety) seeds, pesticides, farmer experience, then it will be possible to make a wider comparison.

This study incorporates the Cobb-Douglas type production function for analyzing the contribution of impact of education on *aman* rice which is considered as output and for further research *aus* and *boro* rice can be considered. But this production function has some limitations which have not been considered in this research but may be considered in further research. There is a scope for application of the joint output model where rice and another crop, rice and livestock, rice and fisheries, rice and forest can be treated as separate outputs.

To estimate the impact of education on rice production Cobb-Douglas type production function has been used. In addition, constant elasticity of substitution or translog production function can be used instead of Cobb-Douglas type production function. Further research can be carried out by using time series data for rice production and various levels of education which might generate better results.

The study confines only three villages of Shibganj *Upazila* of Bogra district of Bangladesh for explaining the impact of education on rice production. There is a scope for further work by using the same procedure for different regions of the country to make the findings holistic.

References

Adnan, N., Ahmad,H.M., & Adnan,R.(2006). A Comparative Study on Some Methods for Handling Multicollinearity Problems, *MATHEMATIKA*, Department of Mathematics, Universiti Teknologi Malaysia (UTM) Volume 2,109-119.

Alam,G.M.,Hoque,K.E.,Khalifa,M.T.B.,Siraj,S.B.,& Ghani,M.F.B.A.(2009). The Role of Agriculture Education and Training on Agriculture Economics and National Development of Bangladesh. *African Journal of Agricultural Research* Vol.4 (2), pp1334-1350

Alam,S.M, Saeed,S.A., Sahabuddin,M., & Akter,S.(2013). Relationship between Employees Recognition and Employee Contribution in Service Industry, *International Journal of Business and Marketing Managemen*t, vol.1 (1), 1-8. Retrieved from http://www.resjournals.org/IJBMM.

Alene, A.D., & Manyong, V.M. (2006). Endogenous Technology Adoption and Household Food Security: The Case of Improved Cowpea Varieties in Northern Nigeria. *Quarterly Journal of International Agriculture* 45(3):211–30.

Ali & Noman. (2013). Determinants of Demand for Health care in Bangladesh: An Econometric Analysis, *World Journal of Social Sciences*,3(6),153-163.

Ali, M., &Flinn,J.C.(1989). Profit Efficiency among Basmati Rice Producers in Pakistan Punjab. *American Journal of Agricultural Economics*, 71(2), 303–10.

Appleton, S. & Balihuta,A.(1996). Education and Agricultural Productivity: Evidence from Uganda, *Journal of International Development*: Vol.8, No.3,415-444.

Asadullah, M .N, & Rahman, S.(2009).Farm Productivity and Efficiency in Rural Bangladesh: The Role of Education Revisited. *Applied Econometrics* 41(1):17-33.

Asfaw, A., &Admassie,A.(2004). The Role of Education on the Adoption of Chemical Fertiliser under Different Socioeconomic Environments in Ethiopia. *Agricultural Economics* 30(3):215–28.

Begum,R.(1998). An Economic Study of Wheat Production and Marketing in Bangladesh. Diss.Ehime U.

Bhuiyan, N. I., Paul, D. N. R.,&Jabber, M. A. (2002). Feeding the extra millions by 2025– Challenges for rice research and extension in Bangladesh, National Workshop on Rice Research and Extension in Bangladesh, *Bangladesh Rice Research Institute*, Gazipur,29-31.

Birdsall,N., Ibrahim,A. J., & Gupta,G.R.(2004).Task Force 3 Interim Report on Primary Education.

Cambell, D.G. (1993). Production Frontiers, Technical Efficiency and Productivity Measurement in a Panel of US Manufacturing Plants, Maryland college Park, PA.

Cambell, T.A. (1993). KLEMS Translog Cost Estimates and Energy Elasticities, Florida State University, USA.

Card,D.(2001). Estimating the Return to Schooling: Progress on Some Persistent EconometricProblems. *Econometrica* 69 (5), 1127-1160.

Chatterjee,S.,& Hadi, A. (2006).Regression Analysis by Example, 4th ed, A John Wiley & Sons, Inc, Publication.

Cotlear, D. (1986). Farmer Education and Farm Efficiency in Peru: The Role of Schooling, Extension Services & Migration. Education Policy Division, *Education and Training* Department, December, the World Bank.

Dempster, A. P., Schatzoff, M.,& Wermuth, N. (1977). A simulation study of alternatives to ordinary least squares. *Journal of the American Statistical Association*, 72, pp.77-91.

Dev, U.K.,&Hossain,M.(1995).Farmers' Education, Modern Technology and Technical Efficiency of Rice Growers. *Bangladesh Journal of Agricultural Economics*.Vol.18 (2).

Diewert,W.E.(1973). Separability and a Generalization of the Cobb-Douglas Cost, Production, and Indirect Utility Functions, *Institute for Mathematical Studies in the Social Science, Technical* Report: 86, Stanford University.

Dunn,O.J.,&Clark,V.A.(1974). Applied Statistics: Analysis of Variance and Regression. Toronto: John Wiley & Sons.

Duraisamy, P. (1989). Human Capital and Adoption of Innovation in Agricultural Production: Indian Evidence, Center Discussion Paper No.577, July, *Economic Growth Center*, Yale University.

Durbin, J., &Watson, G. S. (1950).Testing for Serial Correlation in Least Squares Regression, I. *Biometrika* 37, 409–428.

Durbin, J.,& Watson, G. S. (1951).Testing for Serial Correlation in Least Squares Regression, II.*Biometrika* 38, 159–179.

Fafchamps, M.,& Quisumbing,A.R.(199). Human Capital, Productivity and Labor Allocation in Rural Pakistan. *Journal of Human Resources*, 34(2), 369-406.

Feder, G., Just, R., & Zilberman, D. (1985).Adoption of agricultural innovations in developing countries: a survey, *Economic Development and Cultural Change,* 33, 255-298.

Foster, A.D., &. Rosenzweig,M. R.(1996a). Return and Investments from the Green Revolution. *American Economic Review* 86: 931-953.

Foster, A.D., & Rosenzweig, M. R. (1996b). Technical Change and Human-Capital Returns and Investments: Evidence from the Green Revolution. *The American Economic Review* 86(4): 931-953.

Gallacher, M.(2001). Education as an Input in Agricultural Production: Argentina, JEL Classification: D24, Q12.Universidad del CEMA, Buenos Aires.

George, D.,& Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference. 11.0 update* (4th ed.). Boston: Allyn & Bacon.

Ghasemi, A., & Zahediasl, S.(2012). Normality Tests for Statistical Analysis: A Guide for Non-Statisticians, *International Journal of Endocrinology & Metabolism,* 10(2),486-489.

Gisser, M. (1965).Schooling and the Farm Problem. *Econometrica*, 33:582-92.

GOB (1989). Handbook of Agricultural Statistics, *Bangladesh Bureau of Statistics* (BBS).

GOB (1990).*Bangladesh Economic Review* 1990, Ministry of Finance, Dhaka

GOB (1998). *The Fifth Five Year Plan 1997-2002,* Dhaka: Planning Commission, Ministry of Planning, Government of Bangladesh.

GOB (2001). Bangladesh Population Census Report, 2001. National Volume, Dhaka, *Bangladesh Bureau of Statistics* (BBS).

GOB (2005).Statistical Pocket Book of Bangladesh, 2005, Published November 2006, Dhaka; *Bangladesh Bureau of Statistics* (BBS), Planning Division, Ministry of Planning; Government of the People's Republic of Bangladesh

GOB (2007).Handbook of Agricultural Statistics, December 2007 *Bangladesh Bureau of Statistics* (BBS).

GOB (2010) Labour Force Survey 2010, *Bangladesh Bureau of Statistics* (BBS), Planning Division, Ministry of Planning; Government of the People's Republic of Bangladesh.

GOB (2012) Population and Housing Census 2011.National Volume, *Bangladesh Bureau of Statistics* (BBS), Planning Division, Ministry of Planning; Government of the People's Republic of Bangladesh.

GOB (2012).*Bangladesh Economic Review* 2012, Ministry of Finance, Dhaka.

GOB (2013).*Bangladesh Economic Review* 2013, Ministry of Finance, Dhaka.

GOB(2012) *Bangladesh Bureau of Statistics* (BBS), Bogra Branch.

Golstein, M.,& Smith, A. F. M. (1974).Ridge-type estimators for regression analysis.

Golstein, M.,& Smith, A.F.M.(1974). Ridge-Type Estimators for Regression Analysis, *Journal of the Royal Statistical Society* B-36, 284-291.

Greene, W. (2003).*Econometric Analysis* (5th Edition). Upper Saddle River: Prentice-Hall.

Grigin, A. (2011). The role of Education in Agricultural Productivity: The Case of Village

Griliches, Z.(1964).Research Expenditures, Education and the Aggregate Agricultural Production Function. American Economic Review, 54:961-74.

Gross & Tales. (1952).Characteristics associated with acceleration of recommended farm practices. *Rural Sociology,* 17(4):321-327.

Gujarati, D. (2003). *Basic Econometrics*. 4th edition, New York: McGraw-Hill.

Halim, A. (1976). The Economic Contribution of Schooling and Extension to Rice Production in the Province of Laguna, Republic of the Philippines. Ph.D. thesis, University of the Philippines at Los Banos.

Hansen, M N. (2001). Education and Economic Rewards-Variations by Social Class Origin and Income Measures. *European Sociological Review.*17 (3):209-231.

Haq, A. Z. M., Ishida, A., Yokoyama, S.,& Taniguchi, K. (2003). Outcomes and issues of agricultural extension services in Bangladesh. *Journal of Agricultural Extension and Research Japan.* 8(17): 17-22.

Haq, A.Z.M. (2012).Farmers Education and Farm Productivity in Bangladesh. *Mediterranean Journal of Social Sciences,* Vol 3 (15).

Haq,A.Z.M.(2011).Effect of Extension Contact on Rice Productivity in Some Selected Sites of Gazipur District, *Bangladesh Journal of Agricultural Research,* 36(4),723-732.

Hassan,Y.Z.M., Khan, S.A., & Malik, H.N.(2003). Effect of Education of the Respondents on the Adoption of Recommended Chemical Control Measures Among the Rice of Thesile Kamoke, District Gujran Wala–Pakistan, *International Journal of Agriculture &Biology,* Vol.4,564-565.

Hoerl, A. E., & Kennard, R. W. (1970).Ridge regression: biased estimation for nonorthogonal problems. *Technometrics,* 12, 55-67.

Hoerl, A. E., Kennard, R. W., & Baldwin, K. F. (1975).Ridge regression: some simulation, Communications in statistics, Vol. 4, 105-123.

Hossain, M. (1991).Agriculture in Bangladesh: Performance Problems and Prospects, Dhaka university Press Ltd.

Hossain, M. (1988). Nature and Impact of Green revolution in Bangladesh: Research Report 67, Washington, D.C: *International Food Policy Research Institute/Bangladesh Institute of Development Studies,*36.

Hossain, M.(1995). Fertilizer Consumption, Pricing and Food Grain Consumption in Bangladesh, Dhaka and Washington DC, *Bangladesh Institute of Development Studies /International food Policy Research Institute.*

Hossain, M., Quasem, M.A., Akash, M.M, & Jabber, M.A.(1990).Differential Impact of Modern Rice Technology: The Bangladesh Case. Dhaka: *Bangladesh institute of Development Studies.*

http://irri.org/index.php?option=com_k2&view=item&id=10824:rice-in-bangladesh (03/07/2013).

Huffman, W. E.(1974).Decision Making: The Role of Education. *American Journal of Agricultural Economics*, 56,85-97. Institutes in Turkey, 1940-1966, *School of Economics and Management*, Lund University.

Ishida, A., & Hassan, M. D. S.(2000).Why do Migrant Workers Intend to Extend their Stay in Host Countries?38(5),101-115.

Jamison, D.T., & Lau,L.J.(1982). *Farmer Education and Farm Efficiency*. Baltimore: John Hopkins University Press.

Jolliffe, D. (1998).Skills, Schooling and Household Income in Ghana. *World Bank Economic Review.*Vol.12. No. 1 .

Khaldi, N.(1975).Education and Allocative Efficiency in U.S. Agriculture. *American Journal of Agricultural Economics*, 57: 650-57.

Khalil, M.I. (1991).The Agriculture Sector in Bangladesh: A Data Base, Dhaka, US Agency for International Development.

Knight, J.,Weir,S.,& Woldehanna,T.(2003). The Role of Education in Facilitating Risk-Taking and Innovation in Agriculture. *Journal of Development Studies* 39 (6),1-22.

Krejcie & Morgan. (1970). Determining Sample Size for Research Activities, *Educational and Psychological Measurement, Vol. 30, 607-610.*

Lin, Y, J. (1991). Education and Innovation Adoption in Agriculture: Evidence from Hybrid Rice in China, *American Journal of Agricultural Economics*, Vol73 (3), 713-723.

Lockheed, M. E., Jamison. T.,& Lau, L. J. (1980).Farmer Education and Farm Efficiency: A Survey, *Economic Development and Cultural Change*, 29(1), 37-76.

Lopez,M.(2007). Estimation of Cronbach's alpha for sparse datasets,*20th Annual Conference of the National Advisory Committee on Computing Qualifications* (NACCQ), Nelson, New Zealand. Samuel Mann and Noel Bridgeman (Eds).

Mahmudul,H.A., Ishda,Z., & Taniguchi,K.(2002). The Role of Farmers' Education on Income in Bangladesh. Retrieved from *http:/www.zoukou.life.shamane-u.ac.ip/ruralm an a/isida/list.htm- 10k.*

Marquardt, D. W. (1970). Generalized inverses, ridge regression, biased linear estimation and nonlinear estimation. *Technometrics,* 15, 591-612.

McDonald, G. C.,& Galarneau, D. I. (1975).A Monte Carlo evaluation of some ridge-type estimators. *Journal of the American Statistical Association,* Vol. 70, 407-416.

Minh-Phuong, T. N. (2006).Education and Agricultural Growth in Vietnam, *Department of Agricultural and Applied Economics*, University of Wisconsin-Madision.

Murphy, L., Bilsborrow, R., & Pichon,F.(1997).Poverty and Prosperity among Migrant Settlersin the Amazon Rainforest Frontier of Ecuador. *Journal of Development Studies.Vol.34*(2).

Nargis, F.,& Lee, S, H.(2013). Efficiency Analysis of Boro Rice Production in North-Central Region of Bangladesh, *The Journal of Animal & Plant Science*, 23(2), 527-533

Nelson, R. R.,& Phelps, E. S. (1966). Investment in Humans, Technological Diffusion, and Economic Growth, *American Economic Review*. 56 (May), 69-75.

Neumann, J.V. (1941).Distribution of the ratio of the mean square successive difference to the variance, *Annals of Mathematical Statistics*, 12, 367-395.

Onphanhdala.P.(2009).Farmer Education and Agricultural Efficiency: Evidence from Lao PDR. No. 20, January, *Graduate School of International Cooperation Studies*, Kobe University.

Phillips, M, J.(1987). A Comment on Farmer Education and Farm Efficiency: A Survey, *Economic Development and Cultural Cahnge,* Vol.35 (3), 637-641.

Pritchett, L. (2001). Where Has All the Education Gone? *The World Bank Economic Review vol.15 (3)367–91.*

Psacharopoulos, G., & Patrinos, H.A.(2002). Returns to Investment in Education: A Further Update. *World Bank Policy Research Working Paper* 2881.

Pudasaini, S.P. (1983).The Effects of Education in Agriculture: Evidence from Nepal, *American Journal of Agricultural Economics*, Vol. 65(3), 509-515.

Quasem, A.(1994).Irrigation and Household Income : A Case Study of Bangladesh. *Bangladesh Journal of Agricultural Economics.*Vol.17, No.1 and 2.

Rahman, M.A., & Karim.(1999).Water Use, Yield and Yield Component of wheal Under Irrigation and Nitrogen Application in Terrace Soil of Bangladesh, Annals of Bangladesh Agriculture, Vol.9.

Rahman, S.(1999).Impact of Technological Change on Income Distribution and Poverty in Bangladesh Agriculture: An Empirical Analysis, *Journal of International Development.*Vol.11 (7).

Rahman, S,M.,& Haque, Z. (2013).Adoption of Selected Wheat Production Technologies in Two Northern Districts of Bangladesh. International *Journal of Agricultural Research Innovation & Technology*,3(1), 5-11.

Rehaman, U.N., Khan. J.,& Tariq. M.(2012). The impact of education on agricultural productionin Khyber Pakhtunkhwa province of Pakistan (1975-2008). *Sarhad Journal of Agriculture*, 28(2), 345-352.

Rogerson, P.A. (2003). Statistical Methods for Geography, SAGE Publication Ltd.,6 Bonhill Street, London.

Salehin, M, M., Kabir, S.M., Morshed.M.K.,& Farid.S.K. (2009). Socioeconomic Changes of Farmers Due to Adoption of Rice Production Technologies in Selected Areas of Sherpur District, *Journal of Bangladesh Agricultural University*,7(2),335-341.

Sargan, J.D.,&Bhargava, A. (1983).Testing residuals from least squares regression for being generated by the Gaussian random walk, *Econometrica,* 51, 153-174.

Schultz, T. P. (2003). Human Capital, Schooling and Health Returns. *Economic Growth Center – Yale University Discussion Paper* 853.

Schultz, T.W. (1975). The Value of the Ability to Deal with Disequilibria. *Journal of Economic Literature* 13(3), 827-46.

Sharif, N.R., & Dar, A. A.(1996). An empirical study of the patterns and sources of technical in efficiency in traditional and HYV rice cultivation in Bangladesh. *Journal of Development Studies,* 32, 612-629.

Singh,B.(1974). Impact of Education on Farm Production, *Economic and Political Weekly* (September), A92-A96.

Tavakol, M., & Dennick, R.(2011).Making sense of Cronbach's alpha, *International Journal of Medical Education*,2,53-55.

Taylor, J. E., &Yunez-Naude, A.(2000). The Returns from Schooling in a Diversified Rural Economy. *American Journal of Agricultural Economics* 82, 287-297.

Tetlay, K.A., Heisey, P.W., Ahmed, Z., & Munir, A. (1991). Farmers' sources of wheat seed and wheat seed management in three irrigated regions of Pakistan. *Seed Science Technoogyl.*19, 123-138.

Van de Walle, D.(2003). Are Returns to Investment Lower for the Poor? Human and Physical Capital Interactions in Rural Vietnam. *Review of Development Economics*, 7(4), 636.

Weir, S.(1999). The Effects of Education on Farmer Productivity in Rural Ethiopia. CSAE Working Paper 99 (7).

Weir, S., & Knight, J. (2004). Adoption and Diffusion of Agricultural Innovations in Ethiopia: The Role of Education. *Economic Development and Cultural Change* in press.

Welch,F.(1970).Education in Production, *Journal of Political Economy*,78,35-59.

William, R.D.& Goldstein,M. (1984). Multivariate analysis. John Wiley and Sons. New York, 53-99.

World Bank.(1990). World Development Report 1990, Washington DC: Oxford University Press.

Wu, C.C.(1977).Education in Farm Production: The Case of Taiwan, *American Journal of Agricultural Economics*, 59,699-709.

www.knowledgebank-brri.orgriceinban.ph (03/07/2013).

Yang, T,D. (1997).Education in Production : Measuring Labour Quality and Management, *American Journal of Agricultural Economics*,79,764-772.

Yasmeen, K.A.E., & Hussain, T. (2011). Impact of educated farmer on Agricultural Product, *Journal of Public Administration and Governce*, Vol.1, No.2.

Conference paper form this research

Islam, M.S., Ali, K.J. and Khan, M.M.H. (2014).Impact of Education on Rice Production in the Northern Districts of Bangladesh: A Ridge Regression Analysis, Presented at regional seminar titled `Regional Issues in Economic Development of Bangladesh' jointly organized by the Department of Economics, University of Rajshahi and Bangladesh Economic Association, Dhaka, held on 10 May 2014 at the Senate Bhaban, University of Rajshahi.

Appendix I

Impact of Education on Rice Production in Bangladesh: A Case Study of Three Villages of Shibgonj Upazilla of Bogra

(This information only used for research)

ID No.	

Part ''A'' Personal and Socio-economic Information

1.1. Name: .... Age ....Years

1.2. Experience of the farmer:

1.3. (a) Gender: .... (Male; Female) (b) Marital Status: ....

1.4. Village: Union:

1.5. Yearly Income (Tk.)...

1.6. Source of household income :(a) Agriculture: ....,(b) Non-Agriculture (Other):.........

1.7. Educational level (Years of schooling- last class passed):............ Years

1.8. Total number of family members...

1.9. What type of latrine do you use?

a) Pucca, B) Semi pucca and c) Katcha.

1.10. What is the size of your total cultivable land? ....bigha/acres/decimal

1.11. How much bigha/acres of land have you cultivated for rice during this season?.......Bigha/acres/decimal.

1.12.(a) Owned land: ....decimal (b) Rented-in: .... decimal

(c) Mortgaged-in: .... decimal (d) Leased-in (agreement): decimal

Part''B" Cost and Production in this season

2.1. **Amount of production during this year**

Rice	Items	Aman
	Total Production of Rice (kg)	
	Production of Rice per bigha (kg)	

2.2. **Cost of factors of production in this year**

SL.No	Items	Unit/ Quantity	Price of Factors (Tk.)	Human Labour(Hired/familylabour) Cost for Production(Tk.)
1	Source of Seed:			
	a) Farmer own seed	Kg		
	b)Quality seed procured	Kg		
	c) Variety of seed	Name		
2	Soil Quality(Texture)	Value [1/]		
3	Seedbed preperation and management			
4	Pulling and Transplanting of seedling			

5	Land Preperatipon			
6	No. of fillage before planting(Draft animal/Power tiller)	Number		
7	Irrigation on planted field	Number		
	Source of irrigation	Value [2/]		
8	Weeding before planting	Number		
9	Fertilizer(Total)	Kg		
	No. of application of fertilizer and cost	Number		
10	Amount of cowdung used	Kg		
11	No. of application of Insecticides	Number		
12	No. of application of Fungicides	Number		
13	No. of application of Herbicides	Number		
14	Hervesting			
15	Threshing			

1/. Soil Quality (Texture): 1=Clay, 2=clay loam, 3=loam, 4=Sandy, 5=sandy loam

2/.Irrigation: 1=Not irrigated, 2=Low lift pump, 3=Shallow tube-well, 4=Deep tube-well, 5= Gravity flow (canal irrigation), 6= Traditional method (Swing basket, done, etc)

Part‘‘C” Extension Service

3. Do you have any suggestion from SAAO in rice production? Yes/No
If yes, then what type of role-plays? Explain.

Appendix-II

$$MSE(\hat{\beta}) = \hat{\sigma}^2 \sum_{i=1}^{9} \frac{1}{\lambda_i}$$ (Where λ is the eigenvalues)

$$= \hat{\sigma}^2 [\frac{1}{\lambda_1} + \frac{1}{\lambda_2} + \frac{1}{\lambda_3} + \frac{1}{\lambda_4} + \frac{1}{\lambda_5} + \frac{1}{\lambda_6} + \frac{1}{\lambda_7} + \frac{1}{\lambda_8} + \frac{1}{\lambda_9}]$$

$$= (0.03375)^2 [\frac{1}{5.462153} + \frac{1}{1.017213} + \frac{1}{1.00009} + \frac{1}{1.000001} + \frac{1}{0.413662} + \frac{1}{0.101371} + \frac{1}{0.005349} + \frac{1}{0.000136} + \frac{1}{0.000024}]$$

= (0.001139) × (0.183078+0.983078+0.99991+0.999999+2.417431+9.864777+186.9474+7360.138+40173.21)

= (0.001139) × (47735.74)

= 54.374

[$\hat{\sigma}$ =Std. Error of estimate (Regression) =0.03375]

$$X'X = \begin{bmatrix} n & \sum x_1 & \sum x_2 & \sum x_3 & \sum x_4 & \sum x_5 & \sum x_6 & \sum x_7 & \sum x_8 \\ \sum x_1 & \sum x_1^2 & \sum x_1x_2 & \sum x_1x_3 & \sum x_1x_4 & \sum x_1x_5 & \sum x_1x_6 & \sum x_1x_7 & \sum x_1x_8 \\ \sum x_2 & \sum x_1x_2 & \sum x_2^2 & \sum x_2x_3 & \sum x_2x_4 & \sum x_2x_5 & \sum x_2x_6 & \sum x_2x_7 & \sum x_2x_8 \\ \sum x_3 & \sum x_1x_3 & \sum x_2x_3 & \sum x_3^2 & \sum x_3x_4 & \sum x_3x_5 & \sum x_3x_6 & \sum x_3x_7 & \sum x_3x_8 \\ \sum x_4 & \sum x_1x_4 & \sum x_2x_4 & \sum x_3x_4 & \sum x_4^2 & \sum x_4x_5 & \sum x_4x_6 & \sum x_4x_7 & \sum x_4x_8 \\ \sum x_5 & \sum x_1x_5 & \sum x_2x_5 & \sum x_3x_5 & \sum x_4x_5 & \sum x_5^2 & \sum x_5x_6 & \sum x_5x_7 & \sum x_5x_8 \\ \sum x_6 & \sum x_1x_6 & \sum x_2x_6 & \sum x_3x_6 & \sum x_4x_6 & \sum x_5x_6 & \sum x_6^2 & \sum x_6x_7 & \sum x_6x_8 \\ \sum x_7 & \sum x_1x_7 & \sum x_2x_7 & \sum x_3x_7 & \sum x_4x_7 & \sum x_5x_7 & \sum x_6x_7 & \sum x_7^2 & \sum x_7x_8 \\ \sum x_8 & \sum x_1x_8 & \sum x_2x_8 & \sum x_3x_8 & \sum x_4x_8 & \sum x_5x_8 & \sum x_6x_8 & \sum x_7x_8 & \sum x_8^2 \end{bmatrix}$$

$$
= \begin{bmatrix}
171 & 1548.87 & 1506.25 & 832.40 & 52 & 64 & 11 & 16 & 90 \\
1548.87 & 14077.67 & 13693.42 & 7586.11 & 467.08 & 585.54 & 99.44 & 146.02 & 809.16 \\
1506.25 & 13693.42 & 13320.6 & 7380.72 & 454.34 & 570.14 & 96.25 & 142.50 & 786.40 \\
832.40 & 7586.11 & 7380.72 & 4098.51 & 250.02 & 317.42 & 52.52 & 79.08 & 431.20 \\
52 & 467.08 & 454.34 & 250.02 & 52 & 0 & 0 & 0 & 31 \\
64 & 585.54 & 570.14 & 317.42 & 0 & 64 & 0 & 0 & 38 \\
11 & 99.44 & 96.25 & 52.52 & 0 & 0 & 11 & 0 & 10 \\
16 & 146.02 & 142.50 & 79.08 & 0 & 0 & 0 & 16 & 11 \\
90 & 809.16 & 786.40 & 431.20 & 31 & 38 & 10 & 11 & 90
\end{bmatrix}
$$

$$
I = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}
$$

$$
KI = 0.13 \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix} = \begin{bmatrix}
.13 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & .13 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & .13 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & .13 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & .13 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & .13 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & .13 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & .13 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .13
\end{bmatrix}
$$

Where K (=0.13) is Ridge Parameter and I is identity matrix.

$X'X + KI$

$$
= \begin{bmatrix}
171 & 1548.87 & 1506.25 & 832.40 & 52 & 64 & 11 & 16 & 90 \\
1548.87 & 14077.67 & 13693.42 & 7586.11 & 467.08 & 585.54 & 99.44 & 146.02 & 809.16 \\
1506.25 & 13693.42 & 13320.6 & 7380.72 & 454.34 & 570.14 & 96.25 & 142.50 & 786.40 \\
832.40 & 7586.11 & 7380.72 & 4098.51 & 250.02 & 317.42 & 52.52 & 79.08 & 431.20 \\
52 & 467.08 & 454.34 & 250.02 & 52 & 0 & 0 & 0 & 31 \\
64 & 585.54 & 570.14 & 317.42 & 0 & 64 & 0 & 0 & 38 \\
11 & 99.44 & 96.25 & 52.52 & 0 & 0 & 11 & 0 & 10 \\
16 & 146.02 & 142.50 & 79.08 & 0 & 0 & 0 & 16 & 11 \\
90 & 809.16 & 786.40 & 431.20 & 31 & 38 & 10 & 11 & 90
\end{bmatrix} +
$$

$$
\begin{bmatrix}
.13 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & .13 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & .13 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & .13 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & .13 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & .13 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & .13 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & .13 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .13
\end{bmatrix}
$$

$$
= \begin{bmatrix}
171.13 & 1548.87 & 1506.25 & 832.40 & 52 & 64 & 11 & 16 & 90 \\
1548.87 & 14077.8 & 13693.42 & 7586.11 & 467.08 & 585.54 & 99.44 & 146.02 & 809.16 \\
1506.25 & 13693.42 & 13320.73 & 7380.72 & 454.34 & 570.14 & 96.25 & 142.50 & 786.40 \\
832.40 & 7586.11 & 7380.72 & 4098.64 & 250.02 & 317.42 & 52.52 & 79.08 & 431.20 \\
52 & 467.08 & 454.34 & 250.02 & 52.13 & 0 & 0 & 0 & 31 \\
64 & 585.54 & 570.14 & 317.42 & 0 & 64.13 & 0 & 0 & 38 \\
11 & 99.44 & 96.25 & 52.52 & 0 & 0 & 11.13 & 0 & 10 \\
16 & 146.02 & 142.50 & 79.08 & 0 & 0 & 0 & 16.13 & 11 \\
90 & 809.16 & 786.40 & 431.20 & 31 & 38 & 10 & 11 & 90.13
\end{bmatrix}
$$

$(X'X + KI)^{-1}$

$$
= \begin{pmatrix}
4.0560 & -0.7977 & -0.1205 & 0.8718 & -0.0316 & -0.0102 & 0.0430 & -0.0140 & 0.0047 \\
-0.7977 & 1.0851 & -0.9907 & -0.0669 & 0.0379 & 0.0366 & -0.0091 & 0.0591 & -0.0159 \\
-0.1205 & -0.9907 & 1.1489 & -0.2053 & -0.0372 & -0.0374 & -0.0043 & -0.0615 & 0.0088 \\
0.8718 & -0.0669 & -0.2053 & 0.3162 & -0.0041 & -0.0057 & 0.0087 & -0.0035 & 0.0127 \\
-0.0316 & 0.0379 & -0.0372 & -0.0041 & 0.0666 & 0.0480 & 0.0512 & 0.0503 & -0.0194 \\
-0.0102 & 0.0366 & -0.0374 & -0.0057 & 0.0480 & 0.0648 & 0.0520 & 0.0514 & -0.0202 \\
0.0430 & -0.0091 & -0.0043 & 0.0087 & 0.0512 & 0.0520 & 0.1501 & 0.0543 & -0.0281 \\
-0.0140 & 0.0591 & -0.0615 & -0.0035 & 0.0503 & 0.0514 & 0.0543 & 0.1164 & -0.0232 \\
0.0047 & -0.0159 & 0.0088 & 0.0127 & -0.0194 & -0.0202 & -0.0281 & -0.0232 & 0.0323
\end{pmatrix}
$$

$$(X'X+KI)^{-2}$$

$$= \begin{pmatrix} 17.8651 & -4.0428 & -0.0136 & 3.8904 & -0.1586 & -0.0717 & 0.1931 & -0.1010 & 0.0418 \\ -4.0428 & 2.8063 & -2.1098 & -0.5866 & 0.1105 & 0.0923 & -0.0343 & 0.1470 & -0.0336 \\ -0.0136 & -2.1098 & 2.3647 & -0.3389 & -0.0833 & -0.0846 & -0.0109 & -0.1381 & 0.0259 \\ 3.8904 & -0.5866 & -0.3389 & 0.9069 & -0.0242 & -0.0060 & 0.0419 & -0.0053 & 0.0078 \\ -0.1586 & 0.1105 & -0.0833 & -0.0242 & 0.0161 & 0.0150 & 0.0152 & 0.0198 & -0.0066 \\ -0.0717 & 0.0923 & -0.0846 & -0.0060 & 0.0150 & 0.0151 & 0.0163 & 0.0196 & -0.0065 \\ 0.1931 & -0.0343 & -0.0109 & 0.0419 & 0.0152 & 0.01633 & 0.0336 & 0.0194 & -0.0080 \\ -0.1010 & 0.1470 & -0.1381 & -0.0053 & 0.0198 & 0.0196 & 0.0194 & 0.0296 & -0.0085 \\ 0.0418 & -0.0336 & 0.0259 & 0.0078 & -0.0066 & -0.0065 & -0.0080 & -0.0085 & 0.0036 \end{pmatrix}$$

$$\hat{\beta} = \begin{bmatrix} 3.1093 \\ -0.0111 \\ -0.0139 \\ 1.0186 \\ 0.0814 \\ 0.1009 \\ 0.1298 \\ 0.1272 \\ 0.0767 \end{bmatrix}$$

$$\hat{\beta}' =$$

$$\begin{bmatrix} 3.1093 & -0.0111 & -0.0139 & 1.0186 & 0.0814 & 0.1009 & 0.1298 & 0.1272 & 0.0767 \end{bmatrix}$$

$$\left(X'X+KI\right)^{-2}\hat{\beta}$$

$$= \begin{pmatrix} 17.8651 & -4.0428 & -0.0136 & 3.8904 & -0.1586 & -0.0717 & 0.1931 & -0.1010 & 0.0418 \\ -4.0428 & 2.8063 & -2.1098 & -0.5866 & 0.1105 & 0.0923 & -0.0343 & 0.1470 & -0.0336 \\ -0.0136 & -2.1098 & 2.3647 & -0.3389 & -0.0833 & -0.0846 & -0.0109 & -0.1381 & 0.0259 \\ 3.8904 & -0.5866 & -0.3389 & 0.9069 & -0.0242 & -0.0060 & 0.0419 & -0.0053 & 0.0078 \\ -0.1586 & 0.1105 & -0.0833 & -0.0242 & 0.0161 & 0.0150 & 0.0152 & 0.0198 & -0.0066 \\ -0.0717 & 0.0923 & -0.0846 & -0.0060 & 0.0150 & 0.0151 & 0.0163 & 0.0196 & -0.0065 \\ 0.1931 & -0.0343 & -0.0109 & 0.0419 & 0.0152 & 0.01633 & 0.0336 & 0.0194 & -0.0080 \\ -0.1010 & 0.1470 & -0.1381 & -0.0053 & 0.0198 & 0.0196 & 0.0194 & 0.0296 & -0.0085 \\ 0.0418 & -0.0336 & 0.0259 & 0.0078 & -0.0066 & -0.0065 & -0.0080 & -0.0085 & 0.0036 \end{pmatrix}$$

$$\times \begin{bmatrix} 3.1093 \\ -0.0111 \\ -0.0139 \\ 1.0186 \\ 0.0814 \\ 0.1009 \\ 0.1298 \\ 0.1272 \\ 0.0767 \end{bmatrix} = \begin{bmatrix} 59.5510 \\ -13.1396 \\ -0.4292 \\ 13.0341 \\ -0.5110 \\ -0.2220 \\ 0.6527 \\ -0.3099 \\ 0.1348 \end{bmatrix}$$

$\hat{\beta}'(X'X + KI)^{-2}\hat{\beta}$ =

$$\begin{bmatrix} 3.1093 & -0.0111 & -0.0139 & 1.0186 & 0.0814 & 0.1009 & 0.1298 & 0.1272 & 0.0767 \end{bmatrix} \times$$

$$\begin{bmatrix} 59.5510 \\ -13.1396 \\ -0.4292 \\ 13.0341 \\ -0.5110 \\ -0.2220 \\ 0.6527 \\ -0.3099 \\ 0.1348 \end{bmatrix} = 198.5819$$

$K^2 \hat{\beta}'(X'X + KI)^{-2}\hat{\beta}$ = $(0.13)^2$ × (198.5819) = (0.0169) × (198.5819) = 3.3560

$$\hat{\sigma}^2 \sum_{j=1}^{9} \frac{\lambda_j}{(\lambda_j + K)^2} =$$

$$\hat{\sigma}^2 \left[\frac{\lambda_1}{(\lambda_1+K)^2} + \frac{\lambda_2}{(\lambda_2+K)^2} + \frac{\lambda_3}{(\lambda_3+K)^2} + \frac{\lambda_4}{(\lambda_4+K)^2} + \frac{\lambda_5}{(\lambda_5+K)^2} + \frac{\lambda_6}{(\lambda_6+K)^2} + \frac{\lambda_7}{(\lambda_7+K)^2} + \frac{\lambda_8}{(\lambda_8+K)^2} + \frac{\lambda_9}{(\lambda_9+K)^2}\right]$$

$$= (0.03375)^2 \left[\frac{5.462153}{(5.462153+0.13)^2} + \frac{1.017213}{(1.017213+0.13)^2} + \frac{1.00009}{(1.00009+0.13)^2} + \frac{1.000001}{(1.000001+0.13)^2} + \frac{0.413662}{(0.413662+0.13)^2} + \frac{0.101371}{(0.101371+0.13)^2} + \frac{0.005349}{(0.005349+0.13)^2} + \frac{0.000136}{(0.000136+0.13)^2} + \frac{0.000024}{(0.000024+0.13)^2}\right]$$

= (0.001139063) × (0.174665 + 0.772901+0.783092+0.783146+1.399547+1.893633+0.291991+0.008023+0.001472)

= (0.001139063) × (6.108471)

= 0.006958

[$\hat{\sigma}$ =Std. Error of estimate (Regression) =0.03375]

$$\therefore MSE(\hat{\beta}_R) = \hat{\sigma}^2 \sum_{j=1}^{9} \frac{\lambda_j}{(\lambda_j + K)^2} + K^2 \hat{\beta}'(X'X + KI)^{-2}\hat{\beta}$$

= 0.006958 + 3.356034

= 3.362992

So $MSE(\hat{\beta}_R) < MSE(\hat{\beta})$

Since, $MSE(\hat{\beta}_R)$ =3.362992and $MSE(\hat{\beta})$ = 54.374

So $MSE_{\hat{\beta}_R}(= 3.36) < MSE_{\hat{\beta}}(= 54.37)$

Appendix-III

$$MSE(\hat{\beta}) = \hat{\sigma}^2 \sum_{i=1}^{9} \frac{1}{\lambda_i}$$ (Where λ is the eigenvalues)

$$= \hat{\sigma}^2 [\frac{1}{\lambda_1} + \frac{1}{\lambda_2} + \frac{1}{\lambda_3} + \frac{1}{\lambda_4} + \frac{1}{\lambda_5} + \frac{1}{\lambda_6} + \frac{1}{\lambda_7} + \frac{1}{\lambda_8} + \frac{1}{\lambda_9}]$$

=

$$(0.03701)^2 [\frac{1}{5.580373} + \frac{1}{1.003255} + \frac{1}{1.000264} + \frac{1}{1.000003} + \frac{1}{0.331188} + \frac{1}{0.079328} + \frac{1}{0.005480} + \frac{1}{0.000088} + \frac{1}{0.00001\ldots}]$$

= (0.00137) ×
(0.179199+0.996755+0.999735+0.999997+3.019426+12.60587+182.4549+11355.69+59039.36)

= (0.00137) × (70596.3)

= 96.698

[$\hat{\sigma}$ =Std. Error of estimate (Regression) = 0.03701]

$$X'X = \begin{bmatrix} n & \sum x_1 & \sum x_2 & \sum x_3 & \sum x_4 & \sum x_5 & \sum x_6 & \sum x_7 & \sum x_8 \\ \sum x_1 & \sum x_1^2 & \sum x_1 x_2 & \sum x_1 x_3 & \sum x_1 x_4 & \sum x_1 x_5 & \sum x_1 x_6 & \sum x_1 x_7 & \sum x_1 x_8 \\ \sum x_2 & \sum x_1 x_2 & \sum x_2^2 & \sum x_2 x_3 & \sum x_2 x_4 & \sum x_2 x_5 & \sum x_2 x_6 & \sum x_2 x_7 & \sum x_2 x_8 \\ \sum x_3 & \sum x_1 x_3 & \sum x_2 x_3 & \sum x_3^2 & \sum x_3 x_4 & \sum x_3 x_5 & \sum x_3 x_6 & \sum x_3 x_7 & \sum x_3 x_8 \\ \sum x_4 & \sum x_1 x_4 & \sum x_2 x_4 & \sum x_3 x_4 & \sum x_4^2 & \sum x_4 x_5 & \sum x_4 x_6 & \sum x_4 x_7 & \sum x_4 x_8 \\ \sum x_5 & \sum x_1 x_5 & \sum x_2 x_5 & \sum x_3 x_5 & \sum x_4 x_5 & \sum x_5^2 & \sum x_5 x_6 & \sum x_5 x_7 & \sum x_5 x_8 \\ \sum x_6 & \sum x_1 x_6 & \sum x_2 x_6 & \sum x_3 x_6 & \sum x_4 x_6 & \sum x_5 x_6 & \sum x_6^2 & \sum x_6 x_7 & \sum x_6 x_8 \\ \sum x_7 & \sum x_1 x_7 & \sum x_2 x_7 & \sum x_3 x_7 & \sum x_4 x_7 & \sum x_5 x_7 & \sum x_6 x_7 & \sum x_7^2 & \sum x_7 x_8 \\ \sum x_8 & \sum x_1 x_8 & \sum x_2 x_8 & \sum x_3 x_8 & \sum x_4 x_8 & \sum x_5 x_8 & \sum x_6 x_8 & \sum x_7 x_8 & \sum x_8^2 \end{bmatrix}$$

$$= \begin{pmatrix} 96 & 871.25 & 848.47 & 466.05 & 38 & 32 & 6 & 5 & 61 \\ 871.25 & 7932.67 & 7726.72 & 4256.42 & 341.76 & 294.86 & 56.54 & 45.91 & 551.76 \\ 848.47 & 7726.72 & 7526.48 & 4146.86 & 332.82 & 287.63 & 54.95 & 44.63 & 537.30 \\ 466.05 & 4256.42 & 4146.86 & 2291.52 & 181.40 & 160.14 & 30.70 & 24.72 & 293.01 \\ 38 & 341.76 & 332.82 & 181.40 & 38 & 0 & 0 & 0 & 29 \\ 32 & 294.86 & 287.63 & 160.14 & 0 & 32 & 0 & 0 & 24 \\ 6 & 56.54 & 54.95 & 30.70 & 0 & 0 & 6 & 0 & 5 \\ 5 & 45.91 & 44.63 & 24.72 & 0 & 0 & 0 & 5 & 3 \\ 61 & 551.76 & 537.30 & 293.01 & 29 & 24 & 5 & 3 & 61 \end{pmatrix}$$

$$I= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$KI=0.13 \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} .13 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & .13 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & .13 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & .13 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & .13 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & .13 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & .13 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & .13 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .13 \end{bmatrix}$$

Where *K* (=0.13) is Ridge Parameter and *I* is identity matrix.

$$X'X + KI$$

$$= \begin{bmatrix} 96 & 871.25 & 848.47 & 466.05 & 38 & 32 & 6 & 5 & 61 \\ 871.25 & 7932.67 & 7726.72 & 4256.42 & 341.76 & 294.86 & 56.54 & 45.91 & 551.76 \\ 848.47 & 7726.72 & 7526.48 & 4146.86 & 332.82 & 287.63 & 54.95 & 44.63 & 537.30 \\ 466.05 & 4256.42 & 4146.86 & 2291.52 & 181.40 & 160.14 & 30.70 & 24.72 & 293.01 \\ 38 & 341.76 & 332.82 & 181.40 & 38 & 0 & 0 & 0 & 29 \\ 32 & 294.86 & 287.63 & 160.14 & 0 & 32 & 0 & 0 & 24 \\ 6 & 56.54 & 54.95 & 30.70 & 0 & 0 & 6 & 0 & 5 \\ 5 & 45.91 & 44.63 & 24.72 & 0 & 0 & 0 & 5 & 3 \\ 61 & 551.76 & 537.30 & 293.01 & 29 & 24 & 5 & 3 & 61 \end{bmatrix}$$

$$
+\begin{bmatrix}
.13 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & .13 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & .13 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & .13 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & .13 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & .13 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & .13 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & .13 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .13
\end{bmatrix}
$$

$$
=\begin{pmatrix}
96.13 & 871.25 & 848.47 & 466.05 & 38 & 32 & 6 & 5 & 61 \\
871.25 & 7932.8 & 7726.72 & 4256.42 & 341.76 & 294.86 & 56.54 & 45.91 & 551.76 \\
848.47 & 7726.72 & 7526.61 & 4146.86 & 332.82 & 287.63 & 54.95 & 44.63 & 537.3 \\
466.05 & 4256.42 & 4146.86 & 2291.65 & 181.4 & 160.14 & 30.7 & 24.72 & 293.01 \\
38 & 341.76 & 332.82 & 181.4 & 38.13 & 0 & 0 & 0 & 29 \\
32 & 294.86 & 287.63 & 160.14 & 0 & 32.13 & 0 & 0 & 24 \\
6 & 56.54 & 54.95 & 30.7 & 0 & 0 & 6.13 & 0 & 5 \\
5 & 45.91 & 44.63 & 24.72 & 0 & 0 & 0 & 5.13 & 3 \\
61 & 551.76 & 537.3 & 293.01 & 29 & 24 & 5 & 3 & 61.13
\end{pmatrix}
$$

$(X'X+KI)^{-1}$

$$
=\begin{pmatrix}
5.7547 & -0.8120 & -0.4290 & 1.1057 & 0.0017 & 0.0248 & 0.1376 & 0.0426 & 0.0337 \\
-0.8120 & 1.8685 & -1.8411 & 0.0255 & 0.0076 & 0.0128 & -0.0664 & -0.0384 & 0.0035 \\
-0.4290 & -1.8411 & 2.1447 & -0.3706 & -0.0044 & -0.0062 & 0.0569 & 0.0400 & -0.0304 \\
1.1057 & 0.0255 & -0.3706 & 0.3978 & -0.0204 & -0.0319 & -0.0221 & -0.0234 & 0.0420 \\
0.0017 & 0.0076 & -0.0044 & -0.0204 & 0.1333 & 0.1103 & 0.1143 & 0.0987 & -0.0549 \\
0.0248 & 0.0128 & -0.0062 & -0.0319 & 0.1103 & 0.1466 & 0.1201 & 0.1028 & -0.0573 \\
0.1376 & -0.0664 & 0.0569 & -0.0221 & 0.1143 & 0.1201 & 0.2928 & 0.1084 & -0.0626 \\
0.0426 & -0.0384 & 0.0400 & -0.0234 & 0.0987 & 0.1028 & 0.1084 & 0.2888 & -0.0447 \\
0.0337 & 0.0035 & -0.0304 & 0.0420 & -0.0549 & -0.0573 & -0.0626 & -0.0447 & 0.0727
\end{pmatrix}
$$

$(X'X+KI)^{-2}$

$$
=\begin{pmatrix}
35.2050 & -5.3823 & -2.2952 & 6.9376 & 0.0039 & 0.1225 & 0.8428 & 0.2617 & 0.2409 \\
-5.3823 & 7.5470 & -7.0553 & -0.1557 & 0.0112 & 0.0049 & -0.3625 & -0.1969 & 0.0411 \\
-2.2952 & -7.0553 & 8.3165 & -1.4667 & -0.0057 & -0.0243 & 0.2150 & 0.1648 & -0.1086 \\
6.9376 & -0.1557 & -1.4667 & 1.5230 & -0.0177 & 0.0029 & 0.1027 & 0.0056 & 0.0737 \\
0.0039 & 0.0112 & -0.0057 & -0.0177 & 0.0562 & 0.0587 & 0.0760 & 0.0679 & -0.0298 \\
0.1225 & 0.0049 & -0.0243 & 0.0029 & 0.0587 & 0.0637 & 0.0830 & 0.0722 & -0.0310 \\
0.8428 & -0.3625 & 0.2150 & 0.1027 & 0.0760 & 0.0830 & 0.1559 & 0.1006 & -0.0391 \\
0.2617 & -0.1969 & 0.1648 & 0.0056 & 0.0679 & 0.0722 & 0.1006 & 0.1229 & -0.0351 \\
0.2409 & 0.0411 & -0.1086 & 0.0737 & -0.0298 & -0.0310 & -0.0391 & -0.0351 & 0.0213
\end{pmatrix}
$$

$$\hat{\beta} = \begin{bmatrix} 3.0475 \\ 0.0698 \\ -0.0780 \\ 0.9926 \\ 0.1142 \\ 0.1423 \\ 0.1366 \\ 0.1609 \\ 0.0742 \end{bmatrix}$$

$$\hat{\beta}' = \begin{bmatrix} 3.0475 & 0.0698 & -0.0780 & 0.9926 & 0.1142 & 0.1423 & 0.1366 & 0.1609 & 0.0742 \end{bmatrix}$$

$$\left(X'X + KI\right)^{-2} \hat{\beta}$$

$$= \begin{pmatrix} 35.2050 & -5.3823 & -2.2952 & 6.9376 & 0.0039 & 0.1225 & 0.8428 & 0.2617 & 0.2409 \\ -5.3823 & 7.5470 & -7.0553 & -0.1557 & 0.0112 & 0.0049 & -0.3625 & -0.1969 & 0.0411 \\ -2.2952 & -7.0553 & 8.3165 & -1.4667 & -0.0057 & -0.0243 & 0.2150 & 0.1648 & -0.1086 \\ 6.9376 & -0.1557 & -1.4667 & 1.5230 & -0.0177 & 0.0029 & 0.1027 & 0.0056 & 0.0737 \\ 0.0039 & 0.0112 & -0.0057 & -0.0177 & 0.0562 & 0.0587 & 0.0760 & 0.0679 & -0.0298 \\ 0.1225 & 0.0049 & -0.0243 & 0.0029 & 0.0587 & 0.0637 & 0.0830 & 0.0722 & -0.0310 \\ 0.8428 & -0.3625 & 0.2150 & 0.1027 & 0.0760 & 0.0830 & 0.1559 & 0.1006 & -0.0391 \\ 0.2617 & -0.1969 & 0.1648 & 0.0056 & 0.0679 & 0.0722 & 0.1006 & 0.1229 & -0.0351 \\ 0.2409 & 0.0411 & -0.1086 & 0.0737 & -0.0298 & -0.0310 & -0.0391 & -0.0351 & 0.0213 \end{pmatrix}$$

$$\times \begin{bmatrix} 3.0475 \\ 0.0698 \\ -0.0780 \\ 0.9926 \\ 0.1142 \\ 0.1423 \\ 0.1366 \\ 0.1609 \\ 0.0742 \end{bmatrix} = \begin{bmatrix} 114.1698 \\ -15.5561 \\ -9.5478 \\ 22.7763 \\ 0.0294 \\ 0.4148 \\ 2.6833 \\ 0.8254 \\ 0.8014 \end{bmatrix}$$

$\hat{\beta}'(X'X + KI)^{-2}\hat{\beta}$ =

$[3.0475 \quad 0.0698 \quad -0.0780 \quad 0.9926 \quad 0.1142 \quad 0.1423 \quad 0.1366 \quad 0.1609 \quad 0.0742] \times$

$$\begin{bmatrix} 114.1698 \\ -15.5561 \\ -9.5478 \\ 22.7763 \\ 0.0294 \\ 0.4148 \\ 2.6833 \\ 0.8254 \\ 0.8014 \end{bmatrix} = 370.82$$

$K^2 \hat{\beta}'(X'X + KI)^{-2}\hat{\beta}$ = $(0.13)^2$ × (370.82) = (0.0169) × (370.82) = 6.267

$$\hat{\sigma}^2 \sum_{j=1}^{9} \frac{\lambda_j}{(\lambda_j + K)^2} =$$

$$\hat{\sigma}^2[\frac{\lambda_1}{(\lambda_1+K)^2} + \frac{\lambda_2}{(\lambda_2+K)^2} + \frac{\lambda_3}{(\lambda_3+K)^2} + \frac{\lambda_4}{(\lambda_4+K)^2} + \frac{\lambda_5}{(\lambda_5+K)^2} + \frac{\lambda_6}{(\lambda_6+K)^2} + \frac{\lambda_7}{(\lambda_7+K)^2} + \frac{\lambda_8}{(\lambda_8+K)^2} + \frac{\lambda_9}{(\lambda_9+K)^2}]$$

$$= (0.03701)^2[\frac{5.580373}{(5.580373+0.13)^2} + \frac{1.003255}{(1.003255+0.13)^2} + \frac{1.000264}{(1.000264+0.13)^2} + \frac{1.000003}{(1.000003+0.13)^2} + \frac{0.331188}{(0.331188+0.13)^2}$$

$$+ \frac{0.079328}{(0.079328+0.13)^2} + \frac{0.005480}{(0.005480+0.13)^2} + \frac{0.000088}{(0.000088+0.13)^2} + \frac{0.000016}{(0.000016+0.13)^2}]$$

= (0.00137) ×
(0.171133+0.781188+0.782987+0.783145+1.557106+1.81039+0.298599+0.005204+0.001002)

= (0.00137) × (6.190754)

= 0.00848

[$\hat{\sigma}$ =Std. Error of estimate (Regression) = 0.03701]

$$\therefore MSE(\hat{\beta}_R) = \hat{\sigma}^2 \sum_{j=1}^{9} \frac{\lambda_j}{(\lambda_j + K)^2} + K^2 \hat{\beta}'(X'X + KI)^{-2}\hat{\beta}$$

= 0.00848 + 6.267

= 6.275

So $MSE(\hat{\beta}_R) < MSE(\hat{\beta})$

Since, $MSE(\hat{\beta}_R)$ = 6.275 and $MSE(\hat{\beta})$ = 96.698

So $MSE_{\hat{\beta}_R} (= 6.275) < MSE_{\hat{\beta}} (= 96.698)$

Appendix-IV

$$MSE(\hat{\beta}) = \hat{\sigma}^2 \sum_{i=1}^{9} \frac{1}{\lambda_i}$$ (Where λ is the eigenvalues)

$$= \hat{\sigma}^2 [\frac{1}{\lambda_1} + \frac{1}{\lambda_2} + \frac{1}{\lambda_3} + \frac{1}{\lambda_4} + \frac{1}{\lambda_5} + \frac{1}{\lambda_6} + \frac{1}{\lambda_7} + \frac{1}{\lambda_8} + \frac{1}{\lambda_9}]$$

$$= (0.04486)^2 [\frac{1}{5.635941} + \frac{1}{1.165418} + \frac{1}{1.000053} + \frac{1}{0.736965} + \frac{1}{0.313616} + \frac{1}{0.143514} + \frac{1}{0.004353} + \frac{1}{0.000127} + \frac{1}{0.000012}]$$

= (0.002012) × (0.177433+0.858061+0.999947+1.356916+3.188611+6.967964+229.7137+7858.834+81671.3)

= (0.002012) × (89773.39)

= 180.66

[$\hat{\sigma}$ =Std. Error of estimate (Regression) =0.04486]

$$X'X = \begin{bmatrix} n & \sum x_1 & \sum x_2 & \sum x_3 & \sum x_4 & \sum x_5 & \sum x_6 & \sum x_7 & \sum x_8 \\ \sum x_1 & \sum x_1^2 & \sum x_1x_2 & \sum x_1x_3 & \sum x_1x_4 & \sum x_1x_5 & \sum x_1x_6 & \sum x_1x_7 & \sum x_1x_8 \\ \sum x_2 & \sum x_1x_2 & \sum x_2^2 & \sum x_2x_3 & \sum x_2x_4 & \sum x_2x_5 & \sum x_2x_6 & \sum x_2x_7 & \sum x_2x_8 \\ \sum x_3 & \sum x_1x_3 & \sum x_2x_3 & \sum x_3^2 & \sum x_3x_4 & \sum x_3x_5 & \sum x_3x_6 & \sum x_3x_7 & \sum x_3x_8 \\ \sum x_4 & \sum x_1x_4 & \sum x_2x_4 & \sum x_3x_4 & \sum x_4^2 & \sum x_4x_5 & \sum x_4x_6 & \sum x_4x_7 & \sum x_4x_8 \\ \sum x_5 & \sum x_1x_5 & \sum x_2x_5 & \sum x_3x_5 & \sum x_4x_5 & \sum x_5^2 & \sum x_5x_6 & \sum x_5x_7 & \sum x_5x_8 \\ \sum x_6 & \sum x_1x_6 & \sum x_2x_6 & \sum x_3x_6 & \sum x_4x_6 & \sum x_5x_6 & \sum x_6^2 & \sum x_6x_7 & \sum x_6x_8 \\ \sum x_7 & \sum x_1x_7 & \sum x_2x_7 & \sum x_3x_7 & \sum x_4x_7 & \sum x_5x_7 & \sum x_6x_7 & \sum x_7^2 & \sum x_7x_8 \\ \sum x_8 & \sum x_1x_8 & \sum x_2x_8 & \sum x_3x_8 & \sum x_4x_8 & \sum x_5x_8 & \sum x_6x_8 & \sum x_7x_8 & \sum x_8^2 \end{bmatrix}$$

$$= \begin{pmatrix} 91 & 836.60 & 815.59 & 452.67 & 28 & 32 & 9 & 13 & 59 \\ 836.60 & 7716.00 & 7523.51 & 4186.78 & 257.94 & 294.24 & 84.87 & 117.86 & 533.37 \\ 815.59 & 7523.51 & 7336.11 & 4083.21 & 251.65 & 286.86 & 82.90 & 114.74 & 519.11 \\ 452.67 & 4186.78 & 4083.21 & 2278.57 & 140.18 & 158.97 & 47.17 & 62.86 & 283.29 \\ 28 & 257.94 & 251.65 & 140.18 & 28 & 0 & 5 & 0 & 17 \\ 32 & 294.24 & 286.86 & 158.97 & 0 & 32 & 2 & 0 & 25 \\ 9 & 84.87 & 82.90 & 47.17 & 5 & 2 & 9 & 0 & 5 \\ 13 & 117.86 & 114.74 & 62.86 & 0 & 0 & 0 & 13 & 11 \\ 59 & 533.37 & 519.11 & 283.29 & 17 & 25 & 5 & 11 & 59 \end{pmatrix}$$

$$I=\begin{bmatrix}1&0&0&0&0&0&0&0&0\\0&1&0&0&0&0&0&0&0\\0&0&1&0&0&0&0&0&0\\0&0&0&1&0&0&0&0&0\\0&0&0&0&1&0&0&0&0\\0&0&0&0&0&1&0&0&0\\0&0&0&0&0&0&1&0&0\\0&0&0&0&0&0&0&1&0\\0&0&0&0&0&0&0&0&1\end{bmatrix}$$

$$KI=0.13\begin{bmatrix}1&0&0&0&0&0&0&0&0\\0&1&0&0&0&0&0&0&0\\0&0&1&0&0&0&0&0&0\\0&0&0&1&0&0&0&0&0\\0&0&0&0&1&0&0&0&0\\0&0&0&0&0&1&0&0&0\\0&0&0&0&0&0&1&0&0\\0&0&0&0&0&0&0&1&0\\0&0&0&0&0&0&0&0&1\end{bmatrix}=\begin{bmatrix}.13&0&0&0&0&0&0&0&0\\0&.13&0&0&0&0&0&0&0\\0&0&.13&0&0&0&0&0&0\\0&0&0&.13&0&0&0&0&0\\0&0&0&0&.13&0&0&0&0\\0&0&0&0&0&.13&0&0&0\\0&0&0&0&0&0&.13&0&0\\0&0&0&0&0&0&0&.13&0\\0&0&0&0&0&0&0&0&.13\end{bmatrix}$$

Where *K* (=0.13) is Ridge Parameter and *I* is identity matrix.

$$X'X+KI$$

$$=\begin{pmatrix}\mathbf{91}&\mathbf{836.60}&\mathbf{815.59}&\mathbf{452.67}&\mathbf{28}&\mathbf{32}&\mathbf{9}&\mathbf{13}&\mathbf{59}\\\mathbf{836.60}&\mathbf{7716.00}&\mathbf{7523.51}&\mathbf{4186.78}&\mathbf{257.94}&\mathbf{294.24}&\mathbf{84.87}&\mathbf{117.86}&\mathbf{533.37}\\\mathbf{815.59}&\mathbf{7523.51}&\mathbf{7336.11}&\mathbf{4083.21}&\mathbf{251.65}&\mathbf{286.86}&\mathbf{82.90}&\mathbf{114.74}&\mathbf{519.11}\\\mathbf{452.67}&\mathbf{4186.78}&\mathbf{4083.21}&\mathbf{2278.57}&\mathbf{140.18}&\mathbf{158.97}&\mathbf{47.17}&\mathbf{62.86}&\mathbf{283.29}\\\mathbf{28}&\mathbf{257.94}&\mathbf{251.65}&\mathbf{140.18}&\mathbf{28}&\mathbf{0}&\mathbf{5}&\mathbf{0}&\mathbf{17}\\\mathbf{32}&\mathbf{294.24}&\mathbf{286.86}&\mathbf{158.97}&\mathbf{0}&\mathbf{32}&\mathbf{2}&\mathbf{0}&\mathbf{25}\\\mathbf{9}&\mathbf{84.87}&\mathbf{82.90}&\mathbf{47.17}&\mathbf{5}&\mathbf{2}&\mathbf{9}&\mathbf{0}&\mathbf{5}\\\mathbf{13}&\mathbf{117.86}&\mathbf{114.74}&\mathbf{62.86}&\mathbf{0}&\mathbf{0}&\mathbf{0}&\mathbf{13}&\mathbf{11}\\\mathbf{59}&\mathbf{533.37}&\mathbf{519.11}&\mathbf{283.29}&\mathbf{17}&\mathbf{25}&\mathbf{5}&\mathbf{11}&\mathbf{59}\end{pmatrix}$$

$$+\begin{bmatrix}.13&0&0&0&0&0&0&0&0\\0&.13&0&0&0&0&0&0&0\\0&0&.13&0&0&0&0&0&0\\0&0&0&.13&0&0&0&0&0\\0&0&0&0&.13&0&0&0&0\\0&0&0&0&0&.13&0&0&0\\0&0&0&0&0&0&.13&0&0\\0&0&0&0&0&0&0&.13&0\\0&0&0&0&0&0&0&0&.13\end{bmatrix}$$

$$
= \begin{pmatrix}
91.13 & 836.6 & 815.59 & 452.67 & 28 & 32 & 9 & 13 & 59 \\
836.6 & 7716.13 & 7523.51 & 4186.78 & 257.94 & 294.24 & 84.87 & 117.86 & 533.37 \\
815.59 & 7523.51 & 7336.24 & 4083.21 & 251.65 & 286.86 & 82.9 & 114.74 & 519.11 \\
452.67 & 4186.78 & 4083.21 & 2278.70 & 140.18 & 158.97 & 47.17 & 62.86 & 283.29 \\
28 & 257.94 & 251.65 & 140.18 & 28.13 & 0 & 5 & 0 & 17 \\
32 & 294.24 & 286.86 & 158.97 & 0 & 32.13 & 2 & 0 & 25 \\
9 & 84.87 & 82.9 & 47.17 & 5 & 2 & 9.13 & 0 & 5 \\
13 & 117.86 & 114.74 & 62.86 & 0 & 0 & 0 & 13.13 & 11 \\
59 & 533.37 & 519.11 & 283.29 & 17 & 25 & 5 & 11 & 59.13
\end{pmatrix}
$$

$(X'X + KI)^{-1}$

$$
= \begin{pmatrix}
5.3787 & -0.5498 & -0.6145 & 1.0471 & -0.0213 & -0.0010 & 0.0013 & -0.0166 & -0.0196 \\
-0.5498 & 2.1157 & -2.1466 & 0.0702 & 0.0313 & 0.0309 & 0.0083 & 0.0201 & -0.0535 \\
-0.6145 & -2.1466 & 2.5132 & -0.4357 & -0.0342 & -0.0344 & 0.0020 & -0.0229 & 0.0282 \\
1.0471 & 0.0702 & -0.4357 & 0.4390 & -0.0013 & -0.0039 & -0.0217 & -0.0014 & 0.0472 \\
-0.0213 & 0.0312 & -0.0341 & -0.0013 & 0.0947 & 0.0611 & -0.0078 & 0.0606 & -0.0180 \\
-0.0010 & 0.0309 & -0.0344 & -0.0039 & 0.0611 & 0.0977 & 0.0073 & 0.0673 & -0.0290 \\
0.0013 & 0.0083 & 0.0020 & -0.0217 & -0.0078 & 0.0073 & 0.1298 & 0.0141 & -0.0045 \\
-0.0166 & 0.0201 & -0.0229 & -0.0014 & 0.0606 & 0.0673 & 0.0141 & 0.1451 & -0.0303 \\
-0.0196 & -0.0535 & 0.0282 & 0.0472 & -0.0180 & -0.0290 & -0.0045 & -0.0303 & 0.0692
\end{pmatrix}
$$

$(X'X + KI)^{-2}$

$$
= \begin{pmatrix}
30.7078 & -2.7277 & -4.1250 & 6.3199 & -0.1148 & -0.0072 & -0.0213 & -0.0908 & -0.0443 \\
-2.7278 & 9.3965 & -9.6332 & 0.5360 & 0.1582 & 0.1474 & 0.0126 & 0.1093 & -0.1654 \\
-4.1250 & -9.6332 & 11.4952 & -2.0789 & -0.1467 & -0.1582 & -0.0043 & -0.0984 & 0.1814 \\
6.3199 & 0.5360 & -2.0789 & 1.4866 & -0.0069 & 0.0123 & -0.0115 & -0.0088 & -0.0123 \\
-0.1148 & 0.1577 & -0.1462 & -0.0069 & 0.0193 & 0.0184 & -0.0002 & 0.0208 & -0.0087 \\
-0.0072 & 0.1473 & -0.1582 & 0.0123 & 0.0184 & 0.0208 & 0.0025 & 0.0224 & -0.0107 \\
-0.0213 & 0.0126 & -0.0043 & -0.0115 & -0.0002 & 0.0025 & 0.0177 & 0.0041 & -0.0028 \\
-0.0908 & 0.1093 & -0.0984 & -0.0088 & 0.0208 & 0.0224 & 0.0041 & 0.0315 & -0.0110 \\
-0.0443 & -0.1654 & 0.1814 & -0.0123 & -0.0088 & -0.0107 & -0.0028 & -0.0110 & 0.0131
\end{pmatrix}
$$

$$
\hat{\beta} = \begin{bmatrix}
3.3981 \\ -0.0417 \\ -0.0188 \\ 1.0363 \\ 0.0192 \\ 0.0439 \\ 0.0237 \\ 0.0963 \\ 0.1088
\end{bmatrix}
$$

$\hat{\beta}' =$

$$
\begin{bmatrix} 3.3981 & -0.0417 & -0.0188 & 1.0363 & 0.0192 & 0.0439 & 0.0237 & 0.0963 & 0.1088 \end{bmatrix}
$$

$\left(X'X + KI\right)^{-2}\hat{\beta}$

$$= \begin{pmatrix} 30.7078 & -2.7277 & -4.1250 & 6.3199 & -0.1148 & -0.0072 & -0.0213 & -0.0908 & -0.0443 \\ -2.7278 & 9.3965 & -9.6332 & 0.5360 & 0.1582 & 0.1474 & 0.0126 & 0.1093 & -0.1654 \\ -4.1250 & -9.6332 & 11.4952 & -2.0789 & -0.1467 & -0.1582 & -0.0043 & -0.0984 & 0.1814 \\ 6.3199 & 0.5360 & -2.0789 & 1.4866 & -0.0069 & 0.0123 & -0.0115 & -0.0088 & -0.0123 \\ -0.1148 & 0.1577 & -0.1462 & -0.0069 & 0.0193 & 0.0184 & -0.0002 & 0.0208 & -0.0087 \\ -0.0072 & 0.1473 & -0.1582 & 0.0123 & 0.0184 & 0.0208 & 0.0025 & 0.0224 & -0.0107 \\ -0.0213 & 0.0126 & -0.0043 & -0.0115 & -0.0002 & 0.0025 & 0.0177 & 0.0041 & -0.0028 \\ -0.0908 & 0.1093 & -0.0984 & -0.0088 & 0.0208 & 0.0224 & 0.0041 & 0.0315 & -0.0110 \\ -0.0443 & -0.1654 & 0.1814 & -0.0123 & -0.0088 & -0.0107 & -0.0028 & -0.0110 & 0.0131 \end{pmatrix}$$

$$\times \begin{bmatrix} 3.3981 \\ -0.0417 \\ -0.0188 \\ 1.0363 \\ 0.0192 \\ 0.0439 \\ 0.0237 \\ 0.0963 \\ 0.1088 \end{bmatrix} = \begin{bmatrix} 111.0722 \\ -8.9222 \\ -15.9855 \\ 23.0308 \\ -0.3988 \\ -0.0125 \\ -0.0841 \\ -0.3170 \\ -0.1601 \end{bmatrix}$$

$\hat{\beta}'(X'X + KI)^{-2}\hat{\beta} =$

$$\begin{bmatrix} 3.3981 & -0.0417 & -0.0188 & 1.0363 & 0.0192 & 0.0439 & 0.0237 & 0.0963 & 0.1088 \end{bmatrix} \times$$

$$\begin{bmatrix} 111.0722 \\ -8.9222 \\ -15.9855 \\ 23.0308 \\ -0.3988 \\ -0.0125 \\ -0.0841 \\ -0.3170 \\ -0.1601 \end{bmatrix} = 401.916$$

$K^2\hat{\beta}'(X'X + KI)^{-2}\hat{\beta} = (0.13)^2 \times (401.916) = (0.0169) \times (401.916) = 6.79238$

$$\hat{\sigma}^2 \sum_{j=1}^{9} \frac{\lambda_j}{(\lambda_j + K)^2} =$$

$$\hat{\sigma}^2[\frac{\lambda_1}{(\lambda_1+K)^2} + \frac{\lambda_2}{(\lambda_2+K)^2} + \frac{\lambda_3}{(\lambda_3+K)^2} + \frac{\lambda_4}{(\lambda_4+K)^2} + \frac{\lambda_5}{(\lambda_5+K)^2} + \frac{\lambda_6}{(\lambda_6+K)^2} + \frac{\lambda_7}{(\lambda_7+K)^2} + \frac{\lambda_8}{(\lambda_8+K)^2} + \frac{\lambda_9}{(\lambda_9+K)^2}]$$

$$= (0.04486)^2 \Big[\frac{5.635941}{(5.635941+0.13)^2} + \frac{1.165418}{(1.165418+0.13)^2} + \frac{1.000053}{(1.000053+0.13)^2} + \frac{0.736965}{(0.736965+0.13)^2} + \frac{0.313616}{(0.313616+0.13)^2} + \frac{0.143514}{(0.143514+0.13)^2} + \frac{0.004353}{(0.004353+0.13)^2} + \frac{0.000127}{(0.000127+0.13)^2} + \frac{0.000012}{(0.000012+0.13)^2} \Big]$$

= (0.002012) × (0.169522+0.694483+0.783115+0.980491+1.593616+1.918382+0.241166+0.007515+0.000724)

= (0.002012) × (6.389015)

= 0.012857

[$\hat{\sigma}$ =Std. Error of estimate (Regression) =0.04486]

$$\therefore MSE(\hat{\beta}_R) = \hat{\sigma}^2 \sum_{j=1}^{9} \frac{\lambda_j}{(\lambda_j + K)^2} + K^2 \hat{\beta}'(X'X + KI)^{-2} \hat{\beta}$$

= 0.012857 + $(0.13)^2$ (401.916) = 0.012857+ 6.79238

= 6.805238

So $MSE(\hat{\beta}_R) < MSE(\hat{\beta})$

Since, $MSE(\hat{\beta}_R)$ = 6.80 and $MSE(\hat{\beta})$ = 180.66

So $MSE_{\hat{\beta}_R} (= 6.80) < MSE_{\hat{\beta}} (= 180.66)$

Appendix-V

$$MSE(\hat{\beta}) = \hat{\sigma}^2 \sum_{i=1}^{9} \frac{1}{\lambda_i}$$ (Where λ is the eigenvalues)

$$= \hat{\sigma}^2 [\frac{1}{\lambda_1} + \frac{1}{\lambda_2} + \frac{1}{\lambda_3} + \frac{1}{\lambda_4} + \frac{1}{\lambda_5} + \frac{1}{\lambda_6} + \frac{1}{\lambda_7} + \frac{1}{\lambda_8} + \frac{1}{\lambda_9}]$$

$$= (0.04137)^2 [\frac{1}{5.539388} + \frac{1}{1.033643} + \frac{1}{1.000343} + \frac{1}{0.926455} + \frac{1}{0.372796} + \frac{1}{0.12186} + \frac{1}{0.005364} + \frac{1}{0.000131} + \frac{1}{0.00002}]$$

= (0.001711) × (0.180525+0.967452+0.999657+1.079383+2.682434+8.20613+186.4266+7634.43+48903.2)

= (0.001711) × (56738.17)

= 97.11

[$\hat{\sigma}$ =Std. Error of estimate (Regression) =0.04137]

$$X'X = \begin{bmatrix} n & \sum x_1 & \sum x_2 & \sum x_3 & \sum x_4 & \sum x_5 & \sum x_6 & \sum x_7 & \sum x_8 \\ \sum x_1 & \sum x_1^2 & \sum x_1x_2 & \sum x_1x_3 & \sum x_1x_4 & \sum x_1x_5 & \sum x_1x_6 & \sum x_1x_7 & \sum x_1x_8 \\ \sum x_2 & \sum x_1x_2 & \sum x_2^2 & \sum x_2x_3 & \sum x_2x_4 & \sum x_2x_5 & \sum x_2x_6 & \sum x_2x_7 & \sum x_2x_8 \\ \sum x_3 & \sum x_1x_3 & \sum x_2x_3 & \sum x_3^2 & \sum x_3x_4 & \sum x_3x_5 & \sum x_3x_6 & \sum x_3x_7 & \sum x_3x_8 \\ \sum x_4 & \sum x_1x_4 & \sum x_2x_4 & \sum x_3x_4 & \sum x_4^2 & \sum x_4x_5 & \sum x_4x_6 & \sum x_4x_7 & \sum x_4x_8 \\ \sum x_5 & \sum x_1x_5 & \sum x_2x_5 & \sum x_3x_5 & \sum x_4x_5 & \sum x_5^2 & \sum x_5x_6 & \sum x_5x_7 & \sum x_5x_8 \\ \sum x_6 & \sum x_1x_6 & \sum x_2x_6 & \sum x_3x_6 & \sum x_4x_6 & \sum x_5x_6 & \sum x_6^2 & \sum x_6x_7 & \sum x_6x_8 \\ \sum x_7 & \sum x_1x_7 & \sum x_2x_7 & \sum x_3x_7 & \sum x_4x_7 & \sum x_5x_7 & \sum x_6x_7 & \sum x_7^2 & \sum x_7x_8 \\ \sum x_8 & \sum x_1x_8 & \sum x_2x_8 & \sum x_3x_8 & \sum x_4x_8 & \sum x_5x_8 & \sum x_6x_8 & \sum x_7x_8 & \sum x_8^2 \end{bmatrix}$$

$$= \begin{pmatrix} 358 & 3256.73 & 3170.32 & 1751.13 & 118 & 128 & 26 & 34 & 210 \\ 3256.73 & 29726.36 & 28943.65 & 16029.32 & 1066.79 & 1174.66 & 240.86 & 309.81 & 1894.30 \\ 3170.32 & 28943.65 & 28183.2 & 15610.8 & 1038.82 & 1144.64 & 234.11 & 301.88 & 1842.82 \\ 1751.13 & 16029.32 & 15610.8 & 8668.61 & 571.60 & 636.54 & 130.40 & 166.67 & 1007.51 \\ 118 & 1066.79 & 1038.82 & 571.60 & 118 & 0 & 5 & 0 & 77 \\ 128 & 1174.66 & 1144.64 & 636.54 & 0 & 128 & 2 & 0 & 87 \\ 26 & 240.86 & 234.11 & 130.40 & 5 & 2 & 26 & 0 & 20 \\ 34 & 309.81 & 301.88 & 166.67 & 0 & 0 & 0 & 34 & 25 \\ 210 & 1894.30 & 1842.82 & 1007.51 & 77 & 87 & 20 & 25 & 210 \end{pmatrix}$$

$$I=\begin{bmatrix}1&0&0&0&0&0&0&0&0\\0&1&0&0&0&0&0&0&0\\0&0&1&0&0&0&0&0&0\\0&0&0&1&0&0&0&0&0\\0&0&0&0&1&0&0&0&0\\0&0&0&0&0&1&0&0&0\\0&0&0&0&0&0&1&0&0\\0&0&0&0&0&0&0&1&0\\0&0&0&0&0&0&0&0&1\end{bmatrix}$$

$$KI=0.13\begin{bmatrix}1&0&0&0&0&0&0&0&0\\0&1&0&0&0&0&0&0&0\\0&0&1&0&0&0&0&0&0\\0&0&0&1&0&0&0&0&0\\0&0&0&0&1&0&0&0&0\\0&0&0&0&0&1&0&0&0\\0&0&0&0&0&0&1&0&0\\0&0&0&0&0&0&0&1&0\\0&0&0&0&0&0&0&0&1\end{bmatrix}=\begin{bmatrix}.13&0&0&0&0&0&0&0&0\\0&.13&0&0&0&0&0&0&0\\0&0&.13&0&0&0&0&0&0\\0&0&0&.13&0&0&0&0&0\\0&0&0&0&.13&0&0&0&0\\0&0&0&0&0&.13&0&0&0\\0&0&0&0&0&0&.13&0&0\\0&0&0&0&0&0&0&.13&0\\0&0&0&0&0&0&0&0&.13\end{bmatrix}$$

Where *K* (=0.13) is Ridge Parameter and *I* is identity matrix.

$$X'X+KI$$

$$=\begin{pmatrix}358&3256.73&3170.32&1751.13&118&128&26&34&210\\3256.73&29726.36&28943.65&16029.32&1066.79&1174.66&240.86&309.81&1894.30\\3170.32&28943.65&28183.2&15610.8&1038.82&1144.64&234.11&301.88&1842.82\\1751.13&16029.32&15610.8&8668.61&571.60&636.54&130.40&166.67&1007.51\\118&1066.79&1038.82&571.60&118&0&5&0&77\\128&1174.66&1144.64&636.54&0&128&2&0&87\\26&240.86&234.11&130.40&5&2&26&0&20\\34&309.81&301.88&166.67&0&0&0&34&25\\210&1894.30&1842.82&1007.51&77&87&20&25&210\end{pmatrix}$$

$$
+\begin{bmatrix}
.13 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & .13 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & .13 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & .13 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & .13 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & .13 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & .13 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & .13 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .13
\end{bmatrix}
$$

$$
=\begin{pmatrix}
358.13 & 3256.73 & 3170.32 & 1751.13 & 118 & 128 & 26 & 34 & 210 \\
3256.73 & 29726.49 & 28943.65 & 16029.32 & 1066.79 & 1174.66 & 240.86 & 309.81 & 1894.3 \\
3170.32 & 28943.65 & 28183.33 & 15610.8 & 1038.82 & 1144.64 & 234.11 & 301.88 & 1842.82 \\
1751.13 & 16029.32 & 15610.8 & 8668.74 & 571.6 & 636.54 & 130.4 & 166.67 & 1007.51 \\
118 & 1066.79 & 1038.82 & 571.6 & 118.13 & 0 & 5 & 0 & 77 \\
128 & 1174.66 & 1144.64 & 636.54 & 0 & 128.13 & 2 & 0 & 87 \\
26 & 240.86 & 234.11 & 130.4 & 5 & 2 & 26.13 & 0 & 20 \\
34 & 309.81 & 301.88 & 166.67 & 0 & 0 & 0 & 34.13 & 25 \\
210 & 1894.3 & 1842.82 & 1007.51 & 77 & 87 & 20 & 25 & 210.13
\end{pmatrix}
$$

$(X'X+KI)^{-1}$

$$
=\begin{pmatrix}
2.9174 & -0.5226 & -0.1219 & 0.5968 & -0.0189 & -0.0073 & 0.0236 & -0.0070 & 0.0125 \\
-0.5226 & 0.6976 & -0.6383 & -0.0367 & 0.0179 & 0.0162 & -0.0074 & 0.0186 & -0.0081 \\
-0.1219 & -0.6383 & 0.7427 & -0.1303 & -0.0156 & -0.0148 & 0.0044 & -0.0181 & 0.0007 \\
0.5968 & -0.0367 & -0.1303 & 0.1817 & -0.0042 & -0.0048 & -0.0011 & -0.0032 & 0.0108 \\
-0.0189 & 0.0179 & -0.0156 & -0.0042 & 0.0272 & 0.0196 & 0.0133 & 0.0200 & -0.0076 \\
-0.0073 & 0.0162 & -0.0148 & -0.0048 & 0.0196 & 0.0285 & 0.0154 & 0.0211 & -0.0087 \\
0.0238 & -0.0074 & 0.0044 & -0.0011 & 0.0133 & 0.0154 & 0.0518 & 0.0163 & -0.0078 \\
-0.0070 & 0.0186 & -0.0181 & -0.0032 & 0.0200 & 0.0211 & 0.0163 & 0.0509 & -0.0094 \\
0.0125 & -0.0081 & 0.0007 & 0.0108 & -0.0077 & -0.0087 & -0.0078 & -0.0094 & 0.0154
\end{pmatrix}
$$

$(X'X+KI)^{-2}$

$$
=\begin{pmatrix}
9.1565 & -1.8341 & -0.1897 & 1.8848 & -0.0656 & -0.0313 & 0.0722 & -0.0304 & 0.0473 \\
-1.8341 & 1.1695 & -0.8517 & -0.2613 & 0.0336 & 0.0259 & -0.0197 & 0.0299 & -0.0135 \\
-0.1897 & -0.8517 & 0.9916 & -0.1695 & -0.0212 & -0.0208 & 0.0047 & -0.0255 & 0.0031 \\
1.8848 & -0.2613 & -0.1695 & 0.4076 & -0.0110 & -0.0043 & 0.0132 & -0.0035 & 0.0099 \\
-0.0656 & 0.0336 & -0.0212 & -0.0110 & 0.0027 & 0.0024 & 0.0011 & 0.0030 & -0.0012 \\
-0.0313 & 0.0259 & -0.0208 & -0.0043 & 0.0025 & 0.0025 & 0.0015 & 0.0030 & -0.0011 \\
0.0727 & -0.0199 & 0.0047 & 0.0134 & 0.0011 & 0.0015 & 0.0041 & 0.0019 & -0.0005 \\
-0.0304 & 0.0299 & -0.0255 & -0.0035 & 0.0030 & 0.0030 & 0.0019 & 0.0045 & -0.0014 \\
0.0473 & -0.0135 & 0.0031 & 0.0099 & -0.0012 & -0.0011 & -0.0005 & -0.0014 & 0.0008
\end{pmatrix}
$$

$$\hat{\beta} = \begin{bmatrix} 3.0635 \\ 0.0067 \\ -0.0126 \\ 0.9955 \\ 0.0569 \\ 0.0809 \\ 0.0716 \\ 0.1183 \\ 0.0963 \end{bmatrix}$$

$$\hat{\beta}' = \begin{bmatrix} 3.0635 & 0.0067 & -0.0126 & 0.9955 & 0.0569 & 0.0809 & 0.0716 & 0.1183 & 0.0963 \end{bmatrix}$$

$$\left(X'X + KI\right)^{-2} \hat{\beta}$$

$$= \begin{pmatrix} 9.1565 & -1.8341 & -0.1897 & 1.8848 & -0.0656 & -0.0313 & 0.0722 & -0.0304 & 0.0473 \\ -1.8341 & 1.1695 & -0.8517 & -0.2613 & 0.0336 & 0.0259 & -0.0197 & 0.0299 & -0.0135 \\ -0.1897 & -0.8517 & 0.9916 & -0.1695 & -0.0212 & -0.0208 & 0.0047 & -0.0255 & 0.0031 \\ 1.8848 & -0.2613 & -0.1695 & 0.4076 & -0.0110 & -0.0043 & 0.0132 & -0.0035 & 0.0099 \\ -0.0656 & 0.0336 & -0.0212 & -0.0110 & 0.0027 & 0.0024 & 0.0011 & 0.0030 & -0.0012 \\ -0.0313 & 0.0259 & -0.0208 & -0.0043 & 0.0025 & 0.0025 & 0.0015 & 0.0030 & -0.0011 \\ 0.0727 & -0.0199 & 0.0047 & 0.0134 & 0.0011 & 0.0015 & 0.0041 & 0.0019 & -0.0005 \\ -0.0304 & 0.0299 & -0.0255 & -0.0035 & 0.0030 & 0.0030 & 0.0019 & 0.0045 & -0.0014 \\ 0.0473 & -0.0135 & 0.0031 & 0.0099 & -0.0012 & -0.0011 & -0.0005 & -0.0014 & 0.0008 \end{pmatrix}$$

$$\times \begin{bmatrix} 3.0635 \\ 0.0067 \\ -0.0126 \\ 0.9955 \\ 0.0569 \\ 0.0809 \\ 0.0716 \\ 0.1183 \\ 0.0963 \end{bmatrix} = \begin{bmatrix} 29.9172 \\ -5.8555 \\ -0.7733 \\ 6.1807 \\ -0.2107 \\ -0.0990 \\ 0.2365 \\ -0.0951 \\ 0.1543 \end{bmatrix}$$

$\hat{\beta}'(X'X + KI)^{-2}\hat{\beta}$ =

$[3.0635 \quad 0.0067 \quad -0.0126 \quad 0.9955 \quad 0.0569 \quad 0.0809 \quad 0.0716 \quad 0.1183 \quad 0.0963] \times$

$$\begin{bmatrix} 29.9172 \\ -5.8555 \\ -0.7733 \\ 6.1807 \\ -0.2107 \\ -0.0990 \\ 0.2365 \\ -0.0951 \\ 0.1543 \end{bmatrix} = 97.7752$$

$K^2 \hat{\beta}'(X'X + KI)^{-2}\hat{\beta}$ = $(0.13)^2$ × (198.5819) = (0.0169) × (198.5819) = 3.3560

$$\hat{\sigma}^2 \sum_{j=1}^{9} \frac{\lambda_j}{(\lambda_j + K)^2} =$$

$$\hat{\sigma}^2[\frac{\lambda_1}{(\lambda_1+K)^2} + \frac{\lambda_2}{(\lambda_2+K)^2} + \frac{\lambda_3}{(\lambda_3+K)^2} + \frac{\lambda_4}{(\lambda_4+K)^2} + \frac{\lambda_5}{(\lambda_5+K)^2} + \frac{\lambda_6}{(\lambda_6+K)^2} + \frac{\lambda_7}{(\lambda_7+K)^2} + \frac{\lambda_8}{(\lambda_8+K)^2} + \frac{\lambda_9}{(\lambda_9+K)^2}]$$

=

$$(0.04137)^2[\frac{5.539388}{(5.539388+0.13)^2} + \frac{1.033643}{(1.033643+0.13)^2} + \frac{1.000343}{(1.000343+0.13)^2} + \frac{0.926455}{(0.926455+0.13)^2} +$$

$$\frac{0.372796}{(0.372796+0.13)^2} + \frac{0.12186}{(0.12186+0.13)^2} + \frac{0.005364}{(0.005364+0.13)^2} + \frac{0.000131}{(0.000131+0.13)^2} + \frac{0.000020}{(0.000020+0.13)^2}]$$

= (0.001711) ×
(0.172341+0.763363+0.78294+0.830085+1.474646+1.921068+0.292742+0.007735+0.00121)

= (0.001711) × (6.24613)

= 0.01069

[$\hat{\sigma}$ =Std. Error of estimate (Regression) =0.04137]

$$\therefore MSE(\hat{\beta}_R) = \hat{\sigma}^2 \sum_{j=1}^{9} \frac{\lambda_j}{(\lambda_j + K)^2} + K^2 \hat{\beta}'(X'X + KI)^{-2}\hat{\beta}$$

= 0.01069 + $(0.13)^2$ (97.77)

= 0.01069+ (0.0169) (97.77)

= 0.01069+1.652313

= 1.663003

So $MSE(\hat{\beta}_R) < MSE(\hat{\beta})$

Since, $MSE(\hat{\beta}_R)$ =1.663003 and $MSE(\hat{\beta})$ = 97.11

So $MSE_{\hat{\beta}_R}(=1.66) < MSE_{\hat{\beta}}(=97.11)$