CHAPTER- V

Analysis on Impact of Health, Safety, and Welfare on the Productivity of the Tea Workers

5.1: Regression Analysis Considering all the Tea Estates

5.1.1: Descriptive Statistics Results

During the period of the study, the variables- health, safety, welfare, and productivity of the tea workers are found to be very stable and not much varying from their mean values. If we give a look at the table-5.1, the low value of standard deviation of all the three variables in this regard also confirms the stability.

Table: 5.1

Descriptive Statistics Results: Considering all the Tea Estates

	Log_ Productivity	**Log_Welfare_ Expenses**	**Log_Health_ Expenses**	**Log_Safety_ Expenses**
Mean	9.201378	17.64843	16.22087	15.16971
c Median	9.200733	17.65360	16.22706	15.16957
Maximum	9.241884	17.80104	16.36440	15.31646
Minimum	9.165331	17.49301	16.07209	15.03614
Std. Dev.	0.023381	0.096231	0.091400	0.092318
Skewness	0.136704	0.049067	-0.12064	0.089090
Kurtosis	1.899013	1.979847	1.814895	1.772737
Jarque-Bera	1.072436	0.875285	1.218907	1.281601
Probability	**0.584956**	**0.645557**	**0.543648**	**0.526870**
Sum	184.0276	352.9686	324.4173	303.3942
Sum Sq. Dev.	0.010386	0.175946	0.158726	0.161930
Observations	20	20	20	20

Source: *Computed by the author*

The above table exhibits that in the case of all the four variables, p values of Jarque-Bera statistics are greater than 0.05. Therefore, we can assert that all the variables approximately conform to the normality; and it is also observed that the results of median of various variables are more or less equal to the respective mean values.

5.1.2: Unit Root Test Results

Unit root test has been conducted to see whether the time series variables are non-stationary and possesses a unit root. The null hypothesis here is- the series are non-stationary and the alternative hypothesis is- the series are stationary.

Table: 5.2

Unit Root Test Results : Considering all the Tea Estates

Variables	Level			First Difference			Second Difference		
	Intercept	Trend & Intercept	None	Intercept	Trend & Intercept	None	Intercept	Trend & Intercept	None
Log_ Productivity	2.657682 0.9999	0.055670 0.9938	2.538315 0.9952	-2.504703 0.1307	-3.382750 0.0852	-0.174362 0.6094	**-6.108511 0.0001**	-5.885552 0.0011	-6.205288 0.0000
Log_Welfare _ Expenses	-0.473816 0.8763	-3.321310 0.0929	2.678828 0.9966	-6.026958 0.0001	-5.843379 0.0010	-4.155452 0.0003	**-5.416747 0.0006**	-5.083370 0.0049	-5.614677 0.0000
Log_Health_ Expenses	-0.365327 0.8960	-4.459825 0.0115	2.764969 0.9970	-7.956645 0.0000	-7.709586 0.0000	-6.323046 0.0000	**-5.785412 0.0004**	-5.562310 0.0026	-5.935153 0.0000
Log_Safety_ Expenses	-0.424328 0.8851	-4.388991 0.0132	2.527052 0.9951	-6.914468 0.0000	-6.697145 0.0002	-5.605216 0.0000	**-4.251704 0.0058**	-3.178027 0.1282	-4.43567 0.0002

Source: *Computed by the author*

We can see the ADF test result in detail in table-5.2. Here, at the 2nd difference with intercept, the t-statistics of all the variables- productivity, welfare, health, and safety are significant; meaning that all variables are stationary at the 2nd difference with intercept only.

5.1.3: Relation between Welfare, Health, Safety, and Labour Productivity: Multiple Regression Model

If we concentrate on the relation between welfare, health, safety & labour productivity, we get the following regression model, which has come out from log estimation of the above-mentioned variables, where productivity is the dependent variable and welfare, health, and safety expenses are explanatory variables.

Table: 5.3

Multiple Regression Test Result: Considering all the Tea Estates

Dependent Variable: LOG_PRODUCTIVITY				
Method: Least Squares				
Sample: 1998 - 2017				
Included observations: 20				
Variable	**Coefficient**	**Std. Error**	**t-Statistic**	**Prob.**
C	**5.021285**	0.144725	34.69543	**0.0000**
LOG_WELFARE_EXPENSES	**0.121847**	0.030738	3.964035	**0.0011**
LOG_HEALTH_EXPENSES	**0.058822**	0.025744	2.284874	**0.0363**
LOG_SAFETY_EXPENSES	**0.070901**	0.026038	2.722954	**0.015**
R-squared	**0.982356**	Mean dependent var		9.201378
Adjusted R-squared	0.979048	S.D. dependent var		0.023381
S.E. of regression	0.003384	Akaike info criterion		-8.36249
Sum squared resid	0.000183	Schwarz criterion		-8.16335
Log-likelihood	87.62493	Hannan-Quinn criter.		-8.32362
F-statistic	**296.9452**	Durbin-Watson stat		1.70217
Prob(F-statistic)	**0.0000**			

Source: *Computed by the author*

From table-5.3, the following regression equation can be formed:

$$Log(y) = 5.021285 + 0.121847 \log(x_1) + 0.058822 \log(x_2) + 0.070901 \log(x_3)$$

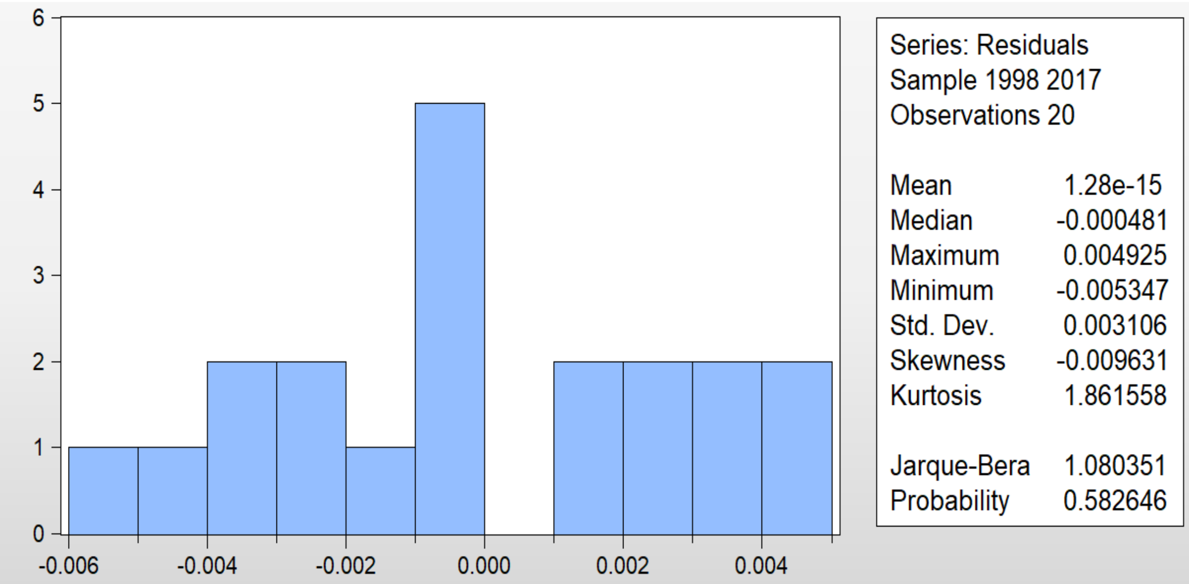
---- (Equation –5. 1)

Where, R^2 = 0.982356, F= 296.9452*, DW= 1.70217, y =labour productivity, x_1 = welfare expenses, x_2 = health expenses, x_3 = safety expenses *=significant at 5% level.

A quick glance at the results, displayed by table-5.1, reveals that the coefficients, in equation-5.1, are statistically significant and the fit is moderately tight. But, before forecasting and making an estimation, normality has been tested to know whether residuals are normally distributed or not.

Figure: 5.1

Jarque-Bera Statistics Result : Multiple Regression



***Source**: Computed by authors*

From the above figure-5.1, we can get the result of Jarque-Bera Statistics. Here, the null hypothesis is- residuals are normally distributed. In this case, looking at the probability of Jarque- Bera statistics, we can easily accept the null hypothesis because of the insignificance of its probability value. So, we can assert that the residuals are normally distributed. But before making an estimation, we have also looked into the matter of heteroskedasticity in the residuals of equation-5.1.

Table: 5.4

Breusch-Pagan-Godfrey Heteroskedasticity Test Result: Multiple Regression

Heteroskedasticity Test: Breusch-Pagan-Godfrey				
F-statistic	1.414599	Prob. F(3,16)		**0.2751**
Obs*R-squared	4.192689	Prob. Chi-Square(3)		**0.2414**
Scaled explained SS	1.155918	Prob. Chi-Square(3)		**0.7636**
Test Equation:				
Dependent Variable: RESID^2				
Method: Least Squares				
Sample: 1998 2017				
Included observations: 20				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.00034	0.000362	-0.9435	0.3595
LOG_WELFARE_EXPENSES	1.70E-05	7.68E-05	0.221335	0.8276
LOG_HEALTH_EXPENSES	9.60E-05	6.43E-05	1.493669	0.1547
LOG_SAFETY_EXPENSES	-9.94E-05	6.50E-05	-1.52812	0.146
R-squared	0.209634	Mean dependent var		9.16E-06
Adjusted R-squared	0.061441	S.D. dependent var		8.73E-06
S.E. of regression	8.45E-06	Akaike info criterion		-20.3471
Sum squared resid	1.14E-09	Schwarz criterion		-20.148
Log-likelihood	207.4714	Hannan-Quinn criter.		-20.3083
F-statistic	1.414599	Durbin-Watson stat		2.46017
Prob(F-statistic)	0.275111			

Source: *Computed by the author*

The above table-5.4 shows us the Breusch-Pagan-Godfrey Heteroskedasticity result. It is known to us that homoscedasticity is one of the prerequisites for an accurate regression model. Here the null hypothesis is- residuals are homoskedastic. We can easily accept the null hypothesis looking at the p- values, as p- values are greater than 0.05. So, it proves that the problem of heteroscedasticity does not exist in equation-5.1. But another problem which often restricts us from making estimation is the existence of serial correlation. So, before making an estimation, we will have to check whether there is any existence of serial correlation in the above-mentioned equation.

Table: 5.5

Breusch-Godfrey Serial Correlation LM Test Result: Multiple Regression

Breusch-Godfrey Serial Correlation LM Test:			
F-statistic	0.142957	Prob. F(2,14)	**0.868**
Obs*R-squared	0.400275	Prob. Chi-Square(2)	**0.8186**

Source: *Computed by the author*

The above table-5.5 gives us Breusch-Godfrey Serial Correlation LM Test result. Here the null hypothesis is- residuals are not serially correlated. If we look at probability values, we find that the values are much greater than 0.05. So, we will have to accept the null hypothesis; that means the equation is free from serial correlation. Since we have tested all the prerequisites of multiple regression, and all the conditions are satisfied, using equation 5.1, we can proceed further for making an estimation. It is mentioned below.

Log (labour productivity) = 5.021285+ 0.121847 log (welfare expenses) + 0.058822 log (health expenses) + 0.070901 log (safety expenses)

----- (Equation – 5.2)

5.1.4: <u>Relation between Welfare, Health, Safety, and Labour Productivity: Simple Regression Model</u>

We have also tested the impact of welfare, health, and safety, separately, on labour productivity to observe how the above-mentioned variables create impact, individually, on labour productivity.

5.1.4.1: <u>Relation between Welfare and Productivity: Simple Regression Model</u>

Firstly, simple regression has been conducted, considering labour productivity as a dependent variable and welfare as an independent variable.

Table: 5.6

Result of Simple Regression Model: Considering Welfare and Productivity

Dependent Variable: LOG_PRODUCTIVITY				
Method: Least Squares				
Sample: 1998 - 2017				
Included observations: 20				
Variable	**Coefficient**	**Std. Error**	**t-Statistic**	**Prob.**
C	**4.988762**	0.188582	26.45407	**0.0000**
LOG_WELFARE_EXPENSES	**0.238696**	0.010685	22.3387	**0.0000**
R-squared	**0.965185**	Mean dependent var		9.201378
Adjusted R-squared	0.963251	S.D. dependent var		0.023381
S.E. of regression	0.004482	Akaike info criterion		-7.88283
Sum squared resid	0.000362	Schwarz criterion		-7.78325
Log-likelihood	80.82825	Hannan-Quinn criter.		-7.86339
F-statistic	**499.0173**	Durbin-Watson stat		1.761112
Prob(F-statistic)	**0.0000**			

Source: *Computed by authors*

From table-5.6, the following regression equation can be formed:

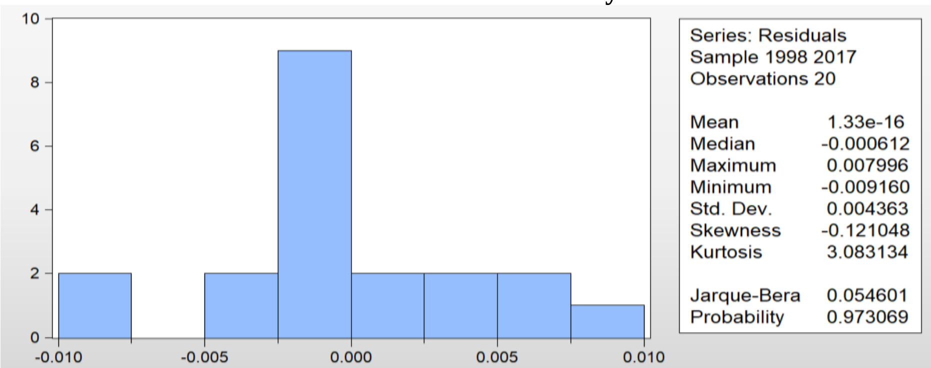
$$Log(y) = 4.988762 + 0.238696 \log (x_1)$$

------ (Equation – 5.3)

Where, R^2 = 0.965185, F= 499.0173*, DW = 1.761112, y = labour productivity, x_1 = welfare expenses, *=significant at 5% level.

A quick glance at the results reveals that the coefficients, in equation-5.3, are statistically significant and the fit is moderately tight. But before forecasting and making an estimation, normality has been tested, to check whether residuals are normally distributed or not.

Figure: 5.2
Jarque-Bera Statistics Result of Simple Regression : Considering Welfare and Productivity



Source: *Computed by the author*

From figure-5.2, we get the result of Jarque-Bera Statistics. The null hypothesis is residuals are normally distributed. Here from the probability of Jarque- Bera Statistics we can easily accept the null hypothesis because of the insignificance of its probability value. So, we can assert that the residuals are normally distributed. But before estimation, we have also looked into the matter of heteroskedasticity in the residuals of equation-5.3.

Table: 5.7
Breusch-Pagan-Godfrey Heteroskedasticity Test Result of Simple Regression : Considering Welfare and Productivity

Heteroskedasticity Test: Breusch-Pagan-Godfrey				
F-statistic	0.12574	Prob. F(1,18)		**0.727**
Obs*R-squared	0.138742	Prob. Chi-Square(1)		**0.7095**
Scaled explained SS	0.117052	Prob. Chi-Square(1)		**0.7323**
Test Equation:				
Dependent Variable: RESID^2				
Method: Least Squares				
Sample: 1998 - 2017				
Included observations: 20				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.00039	0.001153	-0.33892	0.7386
LOG_WELFARE_EXPENSES	2.32E-05	6.53E-05	0.354598	0.727
R-squared	0.006937	Mean dependent var		1.81E-05
Adjusted R-squared	-0.04823	S.D. dependent var		2.68E-05
S.E. of regression	2.74E-05	Akaike info criterion		-18.0766
Sum squared resid	1.35E-08	Schwarz criterion		-17.977
Log-likelihood	182.7661	Hannan-Quinn criter.		-18.0572
F-statistic	0.12574	Durbin-Watson stat		2.452367
Prob(F-statistic)	0.727011			

Source: *Computed by the author*

The above table shows us Breusch-Pagan-Godfrey Heteroskedasticity result. It is known to us that homoscedasticity is one of the prerequisites for an accurate regression model. Here the null hypothesis is residuals are homoskedastic. We can easily accept the null hypothesis looking at the p- values as p- values are greater than 0.05. So, it proves that the problem of heteroscedasticity is not there in this equation. But as we know that the existence of serial correlation restricts us from making an estimation, we have looked into the matter of the existence of serial correlation in this equation through Breusch-Godfrey Serial Correlation LM Test.

Table: 5.8

Breusch-Godfrey Serial Correlation LM Test Result of

Simple Regression : Considering Welfare and Productivity

Breusch-Godfrey Serial Correlation LM Test:			
F-statistic	0.109034	Prob. F(2,16)	**0.8974**
Obs*R-squared	0.26892	Prob. Chi-Square(2)	**0.8742**

Source: *Computed by the author*

The table-5.8 gives us Breusch-Godfrey Serial Correlation LM Test result. Here the null hypothesis is residuals are not serially correlated. If we look at probability values, the vales are much greater than 0.05. So, we will have to accept the null hypothesis; that means the equation is free from serial correlation. As we have tested all the prerequisites of regression we can proceed further for estimation using equation-5.3.

Log (labor productivity) = 4.988762 + 0.238696 log (welfare expenses)

---- (Equation –5.4)

From the equation-5.4, we can assert that 1% increase in welfare expenditure per year in the tea estates leads to 0.238696 % increase in labour productivity per year during the period of 1998 – 2017, which is significant at 5% level.

5.1.4.2: Relation between Health and Productivity: Simple Regression Model

Now, we have conducted simple regression, taking labour productivity as the dependent variable and health as an independent variable.

Table: 5.9

Result of Simple Regression Model: Considering Health and Productivity

Dependent Variable: LOG_PRODUCTIVITY				
Method: Least Squares				
Sample: 1998 - 2017				
Included observations: 20				
Variable	**Coefficient**	**Std. Error**	**t-Statistic**	**Prob.**
C	**5.229787**	0.283209	18.46618	**0.0000**
LOG_HEALTH_EXPENSES	**0.244845**	0.017459	14.02375	**0.0000**
R-squared	**0.916149**	Mean dependent var		9.201378
Adjusted R-squared	0.91149	S.D. dependent var		0.023381
S.E. of regression	0.006956	Akaike info criterion		-7.00383
Sum squared resid	0.000871	Schwarz criterion		-6.90426
Log likelihood	72.0383	Hannan-Quinn criter.		-6.98439
F-statistic	**196.6656**	Durbin-Watson stat		2.055627
Prob(F-statistic)	**0.0000**			

Source: *Computed by the author*

From table-5.9, the following regression equation can be formed:

$$Log(y) = 5.229787 + 0.244845 \log (x_2)$$

----- (Equation-5.5)

R^2=0.916149, F=196.6656*, DW= 2.055627, y =labour productivity, x_2 = Health expenses, *=significant at 5% level.

A quick glance at the results reveals- the coefficients, in equation-5, are statistically significant and the fit is moderately tight. But before estimation, we have also looked into the matter of heteroskedasticity in the residuals of our equation through Breusch-Pagan-Godfrey test.

Table: 5.10

Breusch-Pagan-Godfrey Heteroskedasticity Test Result of Simple Regression : Considering Health and Productivity

Heteroskedasticity Test: Breusch-Pagan-Godfrey				
F-statistic	0.007435	Prob. F(1,18)		**0.9322**
Obs*R-squared	0.008257	Prob. Chi-Square(1)		**0.9276**
Scaled explained SS	0.015939	Prob. Chi-Square(1)		**0.8995**
Test Equation:				
Dependent Variable: RESID^2				
Method: Least Squares				
Sample: 1998 2017				
Included observations: 20				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.00031	0.004079	-0.07555	0.9406
LOG_HEALTH_EXPENSES	2.17E-05	0.000251	0.086225	0.9322
R-squared	0.000413	Mean dependent var		4.35E-05
Adjusted R-squared	-0.05512	S.D. dependent var		9.75E-05
S.E. of regression	0.0001	Akaike info criterion		-15.4844
Sum squared resid	1.81E-07	Schwarz criterion		-15.3848
Log-likelihood	156.844	Hannan-Quinn criter.		-15.465
F-statistic	0.007435	Durbin-Watson stat		2.322208
Prob(F-statistic)	0.93224			

Source: *Computed by the author*

The table-5.10 shows us Breusch-Pagan-Godfrey Heteroskedasticity result. We know that homoscedasticity is one of the prerequisites for an accurate regression model. Here the null hypothesis is residuals are homoskedastic. We can easily accept the null hypothesis looking at the p- values as p- values are greater than 0.05. So, it proves that there is no existence of heteroscedasticity in equation-5.5. But another problem which often restricts us for making estimation is the existence of serial correlation. So before making estimation, we will have to check whether there is any existence of serial correlation in the above mentioned equation.

Table: 5.11

Breusch-Godfrey Serial Correlation LM Test Result of

Simple Regression : Considering Health and Productivity

Breusch-Godfrey Serial Correlation LM Test:			
F-statistic	0.334333	Prob. F(2,16)	**0.7207**
Obs*R-squared	0.802302	Prob. Chi-Square(2)	**0.6695**

Source: *Computed by the author*

The table-5.11gives us Breusch-Godfrey Serial Correlation LM Test result. Here the null hypothesis is residuals are not serially correlated. If we look at probability values, the values are much greater than 0.05. So, we will have to accept the null hypothesis; that means equation 5.5 is free from serial correlation. As we have tested all the prerequisites of regression we can proceed further for estimation using equation-5.5.

Log (labor productivity) = 5.229787+ 0.244845 log (health expenses)

---- (Equation-5.6)

From equation-5.6, we can assert that 1% increase in health expenditure per year in the tea estates lead to 0.244845 % increase in labour productivity per year during the period of 1998 – 2017, which is significant at 5% level.

5.1.4.3: <u>Relation between Safety and Productivity: Simple Regression Model</u>

This time we have conducted simple regression taking labour productivity as the dependent variable and safety as an independent variable.

Table: 5.12

Result of Simple Regression Model: Considering Safety and Productivity

Dependent Variable: LOG_PRODUCTIVITY				
Method: Least Squares				
Sample: 1998 2017				
Included observations: 20				
Variable	**Coefficient**	**Std. Error**	**t-Statistic**	**Prob.**
C	**5.502057**	0.244403	22.51221	**0.0000**
LOG_SAFETY_EXPENSES	**0.243862**	0.016111	15.1364	**0.0000**
R-squared	**0.927158**	Mean dependent var		9.201378
Adjusted R-squared	0.923111	S.D. dependent var		0.023381
S.E. of regression	0.006483	Akaike info criterion		-7.14459
Sum squared resid	0.000757	Schwarz criterion		-7.04501
Log likelihood	73.44585	Hannan-Quinn criter.		-7.12515
F-statistic	**229.1107**	Durbin-Watson stat		1.90786
Prob(F-statistic)	**0.0000**			

Source: *Computed by the author*

From table-5.12, the following regression equation can be formed:

$$\mathbf{Log(y) = 5.502057 + 0.243862 log\ (x_3)}$$

----- (Equation- 5.7)

y =labour productivity, x_3 = Safety expenses, *=significant at 5% level.

A quick glance at the results reveals- the coefficients, in equation-5.7, are statistically significant and the fit is moderately tight. But before estimation, we have also looked into the matter of heteroskedasticity in the residuals of equation-5.7.

Table: 5.13

Breusch-Pagan-Godfrey Heteroskedasticity Test Result of Simple Regression : Considering Safety and Productivity

Heteroskedasticity Test: Breusch-Pagan-Godfrey				
F-statistic	0.043543	Prob. F(1,18)		**0.837**
Obs*R-squared	0.048265	Prob. Chi-Square(1)		**0.8261**
Scaled explained SS	0.09836	Prob. Chi-Square(1)		**0.7538**
Test Equation:				
Dependent Variable: RESID^2				
Method: Least Squares				
Sample: 1998 - 2017				
Included observations: 20				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.000741	0.003368	0.219898	0.8284
LOG_SAFETY_EXPENSES	-4.63E-05	0.000222	-0.20867	0.837
R-squared	0.002413	Mean dependent var		3.78E-05
Adjusted R-squared	-0.05301	S.D. dependent var		8.71E-05
S.E. of regression	8.93E-05	Akaike info criterion		-15.7137
Sum squared resid	1.44E-07	Schwarz criterion		-15.6141
Log-likelihood	159.1366	Hannan-Quinn criter.		-15.6942
F-statistic	0.043543	Durbin-Watson stat		2.239943
Prob(F-statistic)	0.837049			

Source: *Computed by the author*

Table-5.13 shows us Breusch-Pagan-Godfrey Heteroskedasticity result. It is known to us that homoscedasticity is one of the prerequisites for an accurate regression model. Here the null hypothesis is residuals are homoskedastic. We can easily accept the null hypothesis looking at the p- values, as p- values are greater than 0.05. So, it proves that the problem of heteroscedasticity does not exist in equation 5.7. But it is known to us that another problem which often restricts us for making estimation is the existence of serial correlation. So before making an estimation, we will have to check whether there is any existence of serial correlation in the above-mentioned equation.

Table: 5.14

Breusch-Godfrey Serial Correlation LM Test Result of Simple Regression : Considering Safety and Productivity

Breusch-Godfrey Serial Correlation LM Test:			
F-statistic	0.018555	Prob. F(2,16)	**0.9816**
Obs*R-squared	0.046281	Prob. Chi-Square(2)	**0.9771**

Source: *Computed by the author*

Table-5.14 gives us Breusch-Godfrey Serial Correlation LM Test result. Here the null hypothesis is residuals are not serially correlated. If we look at probability values, the values are much greater than 0.05. So, we will have to accept the null hypothesis; that means the equation is free from serial correlation. As we have tested all the prerequisites of regression we can proceed further for estimation using equation-5.7.

Log (labor productivity) = 5.502057+ 0.243862 log (safety expenses)

---- (Equation-5.8)

From equation-5.8, we can assert that 1% increase in safety expenditure per year in the tea estates leads to 0.243862 % increase in labour productivity per year during the period of 1998 – 2017, which is significant at 5% level.