Advances and Applications in Mathematical Sciences
Volume 22, Issue 6, April 2023, Pages 1361-1369
ISSN 0974-6803

STATISTICAL EVALUATION OF PRODUCTIVITY IN ORGANISED MANUFACTURING SECTOR - INDIAN PERSPECTIVE

S. A. JYOTHI RANI[1], P. PREETHI[2] and V. V. HARAGOPAL[3]

[1]Associate Professor
[2]Research Scholar
[3]Professor (Retd.)
Statistics Department
Osmania University, Hyderabad, India
E-mail: jyothi.rani@osmania.ac.in
preethipanchakshari@gmail.com
haragopal_vajjha@yahoo.com

Abstract

This paper attempts to investigate the factors (i.e., labour, capital) that influence overall manufacturing output and identify the model that can best fit the data. The annual survey of industries data from 1981 to 2017 is used. Both multiple linear regression and non-linear regression models are fitted and tested for their reliability. The study identified that the non-linear model better fits the data. The study also shows that the Indian organized manufacturing sector is heavily driven by the capital intensive sector and experiencing an increased return of scale and also observed that there is a huge gap between labour and capital intensive sectors in India which can affect the employment growth in the country.

I. Introduction

The Indian organised manufacturing sector has seen an upward trend of growth rate from 1990's reforms onwards. The organised manufacturing sector plays an important role in the industrial development of the country. At present, the major share (i.e. around 74%) of the country's gross domestic product (GDP) is coming from the industry and service sectors. Rapid industrialization is the only way for faster development. There were many

2020 Mathematics Subject Classification: 62J05.
Keywords: organised manufacturing sector, linear, nonlinear models, labour, capital, output.
Received July 7, 2022; Accepted December 22, 2022

reforms like trade liberalization, privatization and globalization which took place in 1991s which provided many challenges and opportunities to Indian industries (Peter, [1]). The main objective of the paper is to investigate the factors that are affecting the overall output of the manufacturing sector. To analyze the change, the multiple linear and non-linear models are constructed for estimation.

II. Objectives of the Study

- The main objective of this paper is to investigate the factors (i.e. labour, capital) that influence overall manufacturing output and identify the model that can best fit the data.
- The multiple linear regression model is fitted and tested for its reliability.
- The non-linear regression model (Cobb-Douglas Model) is fitted and tested for its reliability.

III. Research Methodology

3.1. Data Source. The study considers the annual survey of industries data, published by the Ministry of Statistics and Programme Implementation (MOSPI) every year. This study examines the abovementioned objectives at all Indian levels. The period of study is from 1981 to 2017. The factors like output, capital, labour are used for this data.

Software used: Stata-14 package is used for finding the results.

3.2. Multiple Linear Regression Model. Linear regression analysis is a technique that is used for identifying and modelling the relationship between variables. If there is more than one independent variable then it is called the multiple linear regression model.

In general, the dependent variable 'y' may be related to k independent variables x such that

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + \varepsilon$$

To investigate the factors (i.e. labour, capital) that influences the overall manufacturing output.

The multiple linear regression

$$Q = \beta_0 + \beta_1 L + \beta_2 K \tag{1}$$

Where Q = Output, L = Labour, K = Capital

Least Squares Method. Least squares method is used to estimate the parameters β_0, β_1 such β_1 that the sum of the squares of the difference between the observed and expected values are minimum.

3.3. Non-Linear Regression Model: Cobb Douglas Model

The Cobb Douglas function is mathematically defined as,

$$Q = AL^{\alpha}K^{\beta} \tag{2}$$

Where Q, L, K represents output, labour and capital respectively. Since it is a non equation and can be converted into a linear equation by applying the logarithmic transformation.

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

Further, it can be generalised as

$$\log Q = \beta_0 + \beta_1 \log L + \beta_2 \log K \tag{3}$$

Where $\beta_0, \beta_1, \beta_2$ are coefficients of the above-said equation.

In general, the Cobb-substitution Douglas production function assumes constant and unitary elasticity of between labour and capital (Peter, Pesala, [1]).

The implications of the production function parameters α and β. There is an increasing (decreasing) return to scale if $\alpha + \beta$ constant returns to scale if $\alpha + \beta = 1$ (Peter, Pesala, [1]).

3.4. Test for the Reliability of the model. R-squared, test for significance of regression, variance inflation factor, diagnostics are used to test the reliability of the model.

IV. Multiple Linear Regression Model

Table 1. Multiple Linear Regression Model.

Source	SS	df	MS			
Model	2.2631e+18	2	1.1316e+18	Number of obs	=	36
Residual	2.0974e+16	33	6.3558e+14	F(2, 33)	=	1780.35
				Prob > F	=	0.0000
				R-squared	=	0.9908
				Adj R-squared	=	0.9903
Total	2.2841e+18	35	6.5260e+16	Root MSE	=	2.5e+07

Q	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
L	4.046839	6.699186	0.60	0.550	-9.582757	17.67643
K	1.802342	.1202739	14.99	0.000	1.557643	2.047041
_cons	-3.42e+07	5.31e+07	-0.64	0.524	-1.42e+08	7.38e+07

The fitted multiple regression model is

$$\text{Q-hat} = (-3.42\text{e} + 07) + 4.04L + 1.80K \tag{4}$$

By using the above model, if the value of Labour(L) and Capital(C) is known then the Output(Q) of the overall manufacturing sector can be predicted.

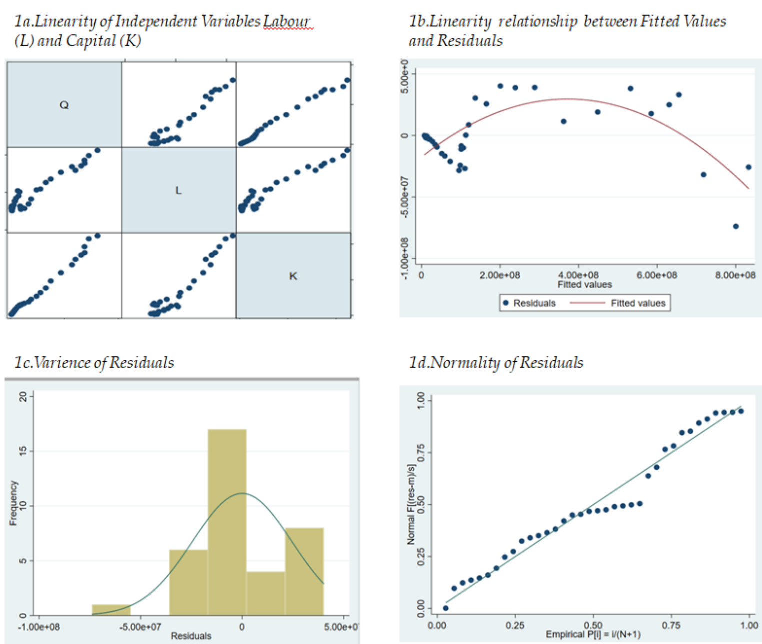

Figure 1. Plot Diagnostics (1/2).

1e.Multicolinearlity

estat vif

Variable	VIF	1/VIF
K	14.69	0.068077
L	14.69	0.068077
Mean VIF	14.69	

1f.Test for Heteroskedasticity

```
Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
        Ho: Constant variance
        Variables: fitted values of Q

        chi2(1)      =    16.58
        Prob > chi2  =   0.0000
```

4.2. Test for the Reliability of the Model. Once the model is fitted the reliability of the model is tested for its assumptions.

4.2.1. R-Square Test. "R-squared is a measure in statistics of how close the data are to the fitted regression line" (Sunthornjittanon, Supichaya, [5]).

As $R^2 = 0.9908$ from the above table.1, which indicates that the model explains 99% of variation present in the data. Since R-squared is high, the model best fits the data.

4.2.2 F-Test. "F-test is used to test the significance of the model Supichaya, [5]. As $\text{Prob} > F = 0.000$ which is less than α (α is the level of significance at 1% (0.01), 5% (0.05), 10% (0.10)). Hence, it can be concluded that the null hypothesis is rejected i.e., H_0 : R-square = 0. Thus, the regression model is statistically significant.

4.2.3 t-Test. "t-test is used to check the significance of individual regression".

(Sunthornjittanon, Supichaya [5]). It is observed at the p values for t in the above table here all the p values are greater than except K which implies that there is no linear relationship between the independent.

4.2.4 Linearity of Independent Variables. From Figure 1a it is observed that each independent variable in the model also follows a linear trend.

4.2.5 VIF. VIF is used to test the multicollinearity from Figure 1e observes that VIF >10 which implies there is a problem of multicollinearity.

4.2.6 Residuals Diagnostic. Linearity: From Figure 1b, it is seen that the residuals do not exhibit a linear trend in the data which suggests that the model is not appropriate for the data.

Normality: The normality plot from shows that the residuals are not much deviated from the normality. Therefore, the normality condition is satisfied.

Heteroscedasticity: It assumes that all the residuals are drawn from a population that has a constant variance. From figure 1f&1c, prob $> \chi^2(0.000) < \alpha(0.01, 0.05, 0.10)$. Hence, reject H_0 : Residual variance is constant and accept H_1 : Residual variance is not constant in this model.

4.3 Non-Linear Regression Model

Table 2. Non-Linear Regression Model.

Source	SS	df	MS			
				Number of obs	=	36
				F(2, 33)	=	3241.22
Model	78.9403864	2	39.4701932	Prob > F	=	0.0000
Residual	.401859933	33	.012177574	R-squared	=	0.9949
				Adj R-squared	=	0.9946
Total	79.3422463	35	2.26692132	Root MSE	=	.11035

logQ	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
logL	-.4048831	.1673264	-2.42	0.021	-.7453113	-.064455
logK	1.166706	.0281092	41.51	0.000	1.109518	1.223895
_cons	4.026888	2.267749	1.78	0.085	-.5868817	8.640658

The fitted multiple regression model is

$$\log Q\text{-hat} = 4.026\text{-}0.404 \log L + 1.166 \log K \tag{5}$$

By using the above model, if the value of Labour(L) and Capital(C) is known then predict the output of the overall manufacturing test.

Return to Scale. The sum of the coefficient indicates returns to scale and from the estimated results from the above table it is seen that $(\alpha + \beta = \text{-}0.404 + 1.166 = 0.762)$ which is less than 1 indicating that decrease in returns to scale in the overall manufacturing sector output in India.

Factor Share. Factor share means factor intensity of output. In this study, it is seen that the value of labour wages coefficient $(\alpha \text{ - } 0.404)$ far lesser than the value of capital coefficient $(\beta \text{ - } 1.166)$ which indicates that the output of the overall manufacturing sector in India is heavily driven by capital intensive sectors.

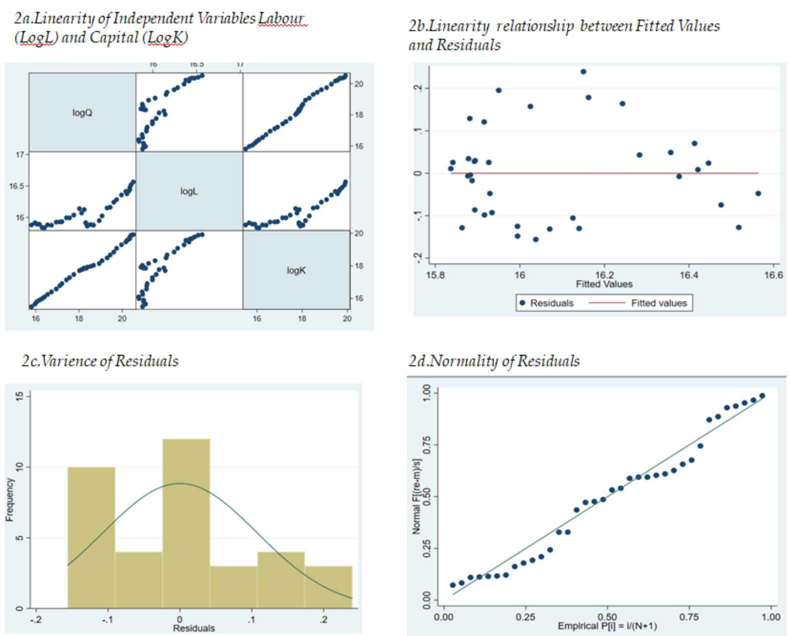

Figure 2. Plot Diagnostics (1/2).

2e.Multicolinearlity

```
estat vif

    Variable |       VIF       1/VIF
-------------+----------------------
        logK |      4.17    0.239642
        logL |      4.17    0.239642
-------------+----------------------
    Mean VIF |      4.17
```

2f.Test for Heteroskedasticity

```
Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
         Ho: Constant variance
         Variables: fitted values of logQ

         chi2(1)      =     0.30
         Prob > chi2  =   0.5819
```

4.4 Test for the Reliability of the Model. Once the model is fitted the reliability of the model is tested for its assumptions.

4.4.1. R-Square Test. As $R^2 = 0.9949$ from the above table which indicates that the model explains 99% of variation present in the data. Since R-square is high, the model best fits the data.

4.4.2. F-Test. As $\text{Prob} > F = 0.000$ which is less than α (α is the level of significance at 1% (0.01), 5% (0.05), 10% (0.10)). We can say that we reject the null hypothesis i.e. H_0 : R-square=0 and accept the alternative hypothesis i.e. R-Square $\neq 0$. We can say the regression model is statistically significant.

4.4.3. t-Test. From the p values for t in the above table it is observed that any variable having p less α is said to be statistically significant. Here, the 2 variables i.e. $\log L < 0.05$ and $\log K < 0.05$ and constant variables i.e. $\text{cons} < 0.10$ which implies there is a linear relationship between the independent variables (log L and log K) and dependent variable $(\log Q)$.

4.4.4. Linearity of Independent Variables. From Figure 2a, it shows that each independent variable in the model also follows a linear trend.

4.4.5. VIF. VIF is used to test the multicollinearity shown in Figure 2e. It is observed that VIF< 5 implies there is no presence of multicollinearity in the data.

4.4.6. Residuals Diagnostic Linearity: From figure 2b, it is observed that residuals are scattered and exhibit a linear trend in the data which suggests the model is appropriate for the data.

Normality: The normality plot from Figure 2d shows that the residuals are not much deviated from the normality. The normality condition is satisfied.

Heteroscedasticity: It assumes that all the residuals are drawn from a population that has a constant variance. From Figure 2f and 2c it is said to be the $\text{prob} > \chi^2(0.581) > \alpha(0.01, 0.05, 0.10)$.

Hence, accept H_0 : Residual variance is constant.

V. Conclusion

The multiple linear regression models is not a good fit for this data as some of the assumptions/validation required for the linear model are not satisfied here. As it is observed from VIF(Multicollinearity) (Figure 1e), the linearity condition of residuals (Figure 1b) and Heteroscedasticity (Figure 1f) does not satisfy the assumptions.

The non-linear regression model (Cobb Douglas Model) is a good fit for this data.

By applying log on the nonlinear model the non-linear regression model is converted into a linear regression model and tested for all the assumptions/validation checks required for any linear model and found that all the assumptions hold good for this model.

- The non-linear model can be interpreted as, if the capital is increased by 1% then there will be an increase of 1.16% in the output and if the labour is increased by 1% then there will be a decrease in the output by 0.40%. Which indicate the Indian organised manufacturing sector is heavily dependent on the capital intensity sector but not on the labour intensity sector and there is a huge gap between the labour and capital intensive sector in India.

- The constant value of 4.026 indicates that in the absence of labour and capital if the other factors like material inputs, no of industry etc is increased by 1% then the output will be increased by 4.026%.

VI. Limitation

- The study is restricted to only the organised manufacturing sector in India. Further, it can be extended to study for both organised and unorganized manufacturing sectors in India.

- The data from 1981-82 to 2017-18 is considered for the study. Further analysis can be done using the data from 1950 onwards.

- The study analyzes only two independent variables Labour(L) and Capital(K) on the dependent variable Output(Q). However, in future research, it can be done on more variables that will provide better results.

References

[1] Peter Pesala, Indian manufacturing industry in the Era of globalization: a Cobb-Douglas production function analysis, Indian Journal of Economics and Development 5(1) (2017), 1-11.

[2] Pathania Rajni and Saleem Anwar, Production coefficient analysis of Indian manufacturing industry, IOSR Journal of Business and Management (IOSRJBM) ISSN: 2278-487X 3(4) (2012), 48-53.

[3] M. A. Hossain and Neil Dias Karunaratne, Trade liberalisation and technical efficiency: evidence from Bangladesh manufacturing industries, Journal of Development Studies 40(3) (2004), 87-114.

[4] MoSPI: Ministry of Statistics and Programme Implementation, 2021.

[5] Sunthornjittanon Supichaya, Linear Regression Analysis on Net Income of an Agrochemical Company in Thailand, (2015).