BA/BBA

ECONOMETRICS

Semester – 07

Individual Assignment – I
ESTIMATION OF A REGRESSION MODEL

Submitted to

Dr. Dhyani Mehta

Date

(27/08/23)

Submitted by

Roll No.	Name
20BABBA025	Arni Parikh

INTRODUCTION

The objective is to test the hypothesis that the increase in the number of workers (employment) will lead to an increase in the Value of Output. The analysis is focused on a particular industry of Pharmaceuticals, Medicinal Chemical & Botanical Products. The data from this industry is tested in order to establish a relationship between output and employment.

In the classical theory, output and employment are determined by the production function and the demand for labour and the supply of labour in the economy. Given the capital stock, technical knowledge and other factors, a precise relation exists between total output and amount of employment, i.e., number of workers. This is shown in the form of the following production function: $Q=f(K, T, N)$ where total output (Q) is a function (f) of capital stock (K), technical knowledge (T), and the number of workers (N). Output is an increasing function of the number of workers, output increases as the employment of labour rises.

HYPOTHESIS

Null Hypothesis- H_0: there is no relationship between output and employment.

Alternate Hypothesis- H_1: Output is influenced by employment and they share a positive relationship.

LITERATURE REVIEW

- The relationship between output and employment tends to improve as we move from monthly to quarterly and annual data. Annual comparisons show a fairly high coefficient of correlation. (Wilson, G. W. (1960).

- An International Labour Organization Report (1996) concluded that the responsiveness of employment growth to GDP growth has not declined in industrialized countries as a whole.

- Pianta, Evangelista and Perani (1996) found that a positive and significant relationship between growth in value added and employment was found only in Germany and the US. (Sahin, Afsin and Tansel, Aysit and Berument, Hakan, 2013)

- Caporale and Skare (2011) found a positive effect of employment growth on output growth both in the short and long run using Granger causality. (Sahin, Afsin and Tansel, Aysit and Berument, Hakan, 2013)

METHODOLOGY AND DATA

Data: EPWRF Annual Survey of Industries

Time Period: 1980-2020

OLS Regression Analysis- EViews

Descriptive Statistics- Excel

Pharmaceuticals, Medicinal Chemical & Botanical Products

Year	Value Of Gross Output (Rs Crores)	Number of Workers
1980 - 1981	1289	104888
1981 - 1982	1555	111025
1982 - 1983	1849	112199
1983 - 1984	2155	119446
1984 - 1985	2391	122509
1985 - 1986	2618	117340
1986 - 1987	2854	119091
1987 - 1988	3509	133773
1988 - 1989	4182	134252
1989 - 1990	5279	138220
1990 - 1991	5989	133267
1991 - 1992	7534	146980
1992 - 1993	9624	166032
1993 - 1994	12136	174407
1994 - 1995	14181	181497
1995 - 1996	16831	204589
1996 - 1997	19537	216127
1997 - 1998	21833	184819
1998 - 1999	28468	213999
1999 - 2000	31305	243410
2000 - 2001	31706	233111
2001 - 2002	32375	225916
2002 - 2003	36198	223014
2003 - 2004	41533	240334
2004 - 2005	46459	264845
2005 - 2006	55079	289453
2006 - 2007	70991	335623
2007 - 2008	75098	353117
2008 - 2009	102000	378413
2009 - 2010	111629	414025
2010 - 2011	135728	469479
2011 - 2012	181536	548874
2012 - 2013	190282	531550
2013 - 2014	214564	617977
2014 - 2015	231834	609754
2015 - 2016	265742	650066
2016 - 2017	271959	679373
2017 - 2018	285014	739728
2018 - 2019	342227	757689
2019 - 2020	348658	806935

DESCRIPTIVE STATISTICS

Value Of Gross Output (Rs Crores)	
Mean	81643.29
Standard Error	16652.96
Median	31505.93
Mode	#N/A
Standard Deviation	105322.5
Sample Variance	1.11E+10
Kurtosis	0.530987
Skewness	1.345643
Range	347369
Minimum	1288.6
Maximum	348657.6
Sum	3265732
Count	40

Mean (Average): The average value of the gross output is approximately Rs 81643.29 Crores.

Standard Error: This value (16652.95512) is an estimate of the standard deviation of the sample mean. It indicates how much the sample mean might vary from the true population mean.

Median (Middle Value): The middle value of the gross output when the data is arranged in ascending order is approximately Rs 31505.93 Crores.

Mode: The mode is the value that appears most frequently in the dataset. In this case, it's not available.

Standard Deviation: The standard deviation (105322.5359) measures the amount of variation or dispersion in the dataset. A higher standard deviation indicates greater variability in the values.

Sample Variance: The sample variance (11092836571) quantifies how much the values in the dataset deviate from the mean. It's the average of the squared differences between each data point and the mean.

Kurtosis: Kurtosis (0.530987329) measures the tailedness of the distribution. A positive value indicates heavier tails compared to a normal distribution.

Skewness: Skewness (1.345643343) measures the asymmetry of the distribution. A positive skewness indicates that the distribution is skewed to the right (tail on the right side), and a negative skewness indicates skewness to the left.

Range: The range (347369.01) is the difference between the maximum and minimum values in the dataset.

Minimum: The minimum value in the dataset is Rs 1288.6 Crores.

Maximum: The maximum value in the dataset is Rs 348657.61 Crores.

Sum: The sum of all the values in the dataset is approximately Rs 3265731.64 Crores.

Count: The number of data points in the dataset is 40.

Number of Workers	
Mean	311178.7
Standard Error	33647.58
Median	224465
Mode	#N/A
Standard Deviation	212806
Sample Variance	4.53E+10
Kurtosis	-0.25033
Skewness	1.033313
Range	702047
Minimum	104888
Maximum	806935
Sum	12447146
Count	40

Mean (Average): The average number of workers is approximately 311178.65.

Standard Error: This value (33647.58346) is an estimate of the standard deviation of the sample mean for the number of workers.

Median (Middle Value): The middle value of the number of workers when the data is arranged in ascending order is 224465.

Mode: The mode is the value that appears most frequently in the dataset. In this case, it's not available ("N/A").

Standard Deviation: The standard deviation (212806.003) measures the amount of variation or dispersion in the number of workers. A higher standard deviation indicates greater variability in the values.

Sample Variance: The sample variance (45286394907) quantifies how much the values in the dataset deviate from the mean for the number of workers. It's the average of the squared differences between each data point and the mean.

Kurtosis: Kurtosis (-0.250327192) measures the tailedness of the distribution. A negative value indicates lighter tails compared to a normal distribution.

Skewness: Skewness (1.033313492) measures the asymmetry of the distribution. A positive skewness indicates that the distribution is skewed to the right (tail on the right side).

Range: The range (702047) is the difference between the maximum and minimum values of the number of workers.

Minimum: The minimum number of workers in the dataset is 104888.

Maximum: The maximum number of workers in the dataset is 806935.

Sum: The sum of all the values in the dataset is 12447146.

Count: The number of data points in the dataset is 40.

THE MODEL

Dependent Variable: VALUE_OF_GROSS OUTPUT   RS CRORES
Method: Least Squares
Date: 08/26/23   Time: 21:40
Sample: 1 41
Included observations: 41

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1674.269	8286.927	-0.202037	0.8409
NUMBER OF WORKERS	0.265125	0.004187	63.31393	0.0000

R-squared	0.990365	Mean dependent var	159304.0
Adjusted R-squared	0.990118	S.D. dependent var	508029.5
S.E. of regression	50503.02	Akaike info criterion	24.54500
Sum squared resid	9.95E+10	Schwarz criterion	24.62859
Log likelihood	-501.1726	Hannan-Quinn criter.	24.57544
F-statistic	4008.653	Durbin-Watson stat	0.340118
Prob(F-statistic)	0.000000		

INTERPRETATION

- Value of Output: y (Dependent Variable)
- Number of Workers: x (Independent Variable)

Therefore, the function y = f(x) is y = -1674.26 + 0.2651x

Where $\beta_0 =$ -1674.26 (constant) and $\beta_1 =$ 0.2651

β_1 represents that 1% change in 'x' will lead to a 26.51% change in 'y'.

- The data above is the regression function drawn out of the Value of Output of Pharmaceuticals, Medicinal Chemical & Botanical Products being the dependent variable and Number of workers in the same sector being the independent variable from the year 1980-2020.
- Probability of $\beta_1 =$ 0.0000 < 0.01 which shows that the coefficient is significant and that the hypothesis is accepted at 90%, 95% and 99% confidence intervals. (p value is less than alpha value at all levels of significance and hence are significant to the model)
- R^2 at 99.03% depicts the goodness of fit (strength) of the model derived. It expresses the proportion of the variation in 'y' which is explained by variation in 'x'.

- Adjusted R^2 at 99.01% shows the accuracy of the model.
- F-statistic's significance shows that the dependant variable is made up of the model and the error both.
- The number of workers have a positive coefficient depicting a positive relation with the dependent variable meaning that an increase in the number of workers would lead to an increase in the value of output of Pharmaceuticals, Medicinal Chemical & Botanical Products.
- Standard Error of the estimate captures the vagueness of the model.

Coefficient Confidence Intervals
Date: 08/26/23   Time: 22:40
Sample: 1 41
Included observations: 41

Variable	Coefficient	90% CI		95% CI		99% CI	
		Low	High	Low	High	Low	High
C	-1674.269	-15636.71	12288.17	-18436.16	15087.62	-24114.55	20766.01
NUMBER_OF_WO...	0.265125	0.258070	0.272181	0.256655	0.273595	0.253786	0.276465

- Coefficient should lie between the upper and lower limits at every confidence level intervals. For example, $0.2537 \leq \beta_1 \leq 0.2764$ at 99% confidence. This determines whether the hypothesis is accepted or not. Since β_1 lies between the lower and the upper limits at all confidence levels depicts that the hypothesis is accepted.

CONCLUSION

There is significant evidence to show that the output is positively related to employment (number of workers) and hence the hypothesis is accepted at all levels of significance. The value of output increases as the number of workers associated with a particular industry (Pharmaceuticals, Medicinal Chemical & Botanical Products) increases according to the model above. Output and Employment through theoretical and statistical proof also show a bidirectional relationship and therefore are positively correlated to each other.

REFERENCES

- The Relationship between Output and Employment. *The MIT Press*. (https://www.jstor.org/stable/1926093)

- Şahin, A., Tansel, A., & Berument, M. H. (2013). Output-Employment Relationship across Sectors: A Long- versus Short-Run Perspective. *Forschungsinstitut zur Zukunft der Arbeit Institute for the Study of Labor*. (https://doi.org/https://dx.doi.org/10.2139/ssrn.2325857)