Nour Mohamed Eletreby

Task 1

A research about a study of the impact of 4 different variables on profitability in Egypt

The statistical studies:

Introduction:

The primary goal of this study is to investigate how the capital structure decisions made by Egyptian companies in the Industrial Goods, Automobile, and Services sectors are linked to their profitability. Furthermore, this research aims to address the existing gap in the literature regarding the determination of the ideal capital structure for firms operating in these specific sectors in Egypt.

To examine the impact of capital structure on firm profitability, the researcher will test some variables that utilized in previous empirical studies to clarify the correlation between capital structure independent variables and profitability dependent variable in a linear model.

Descriptive Statistics

This descriptive statistics is for the 4 independent variables which are X1: total debt to total asset (TD/TA), X2: total debt to total equity (TD/TE), X (TD/TA), X3: short-term debt (STD/TA), X4: long-term debt (LTD/TA) and the dependent variable (Y1): Profitability (ROA)

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Debt-to-assets ratio	150	.000000000000000	4.1184441065546060	.449334810366752	.480843818115892
Debt-to-equity ratio	150	-13.0788659918609720	31.7649740226311700	.881370505711322	3.027466511684184
Short-term Debt-to-assets ratio	150	.000000000000000	1.2050236801854075	.308540209629759	.198027959186843
Long-term Debt-to-assets ratio	150	-.011263786815 3783	2.9134204263691985	.140794600736992	.347227558821381
Profitabillity (ROA)	150	-.500371507627 2441	1.2549301691109305	.071038502580910	.162976992574771
Valid N (listwise)	150				

Figure 1 : Descriptive statistics for all variables

The dependent variable Y : Profitability (RAO)

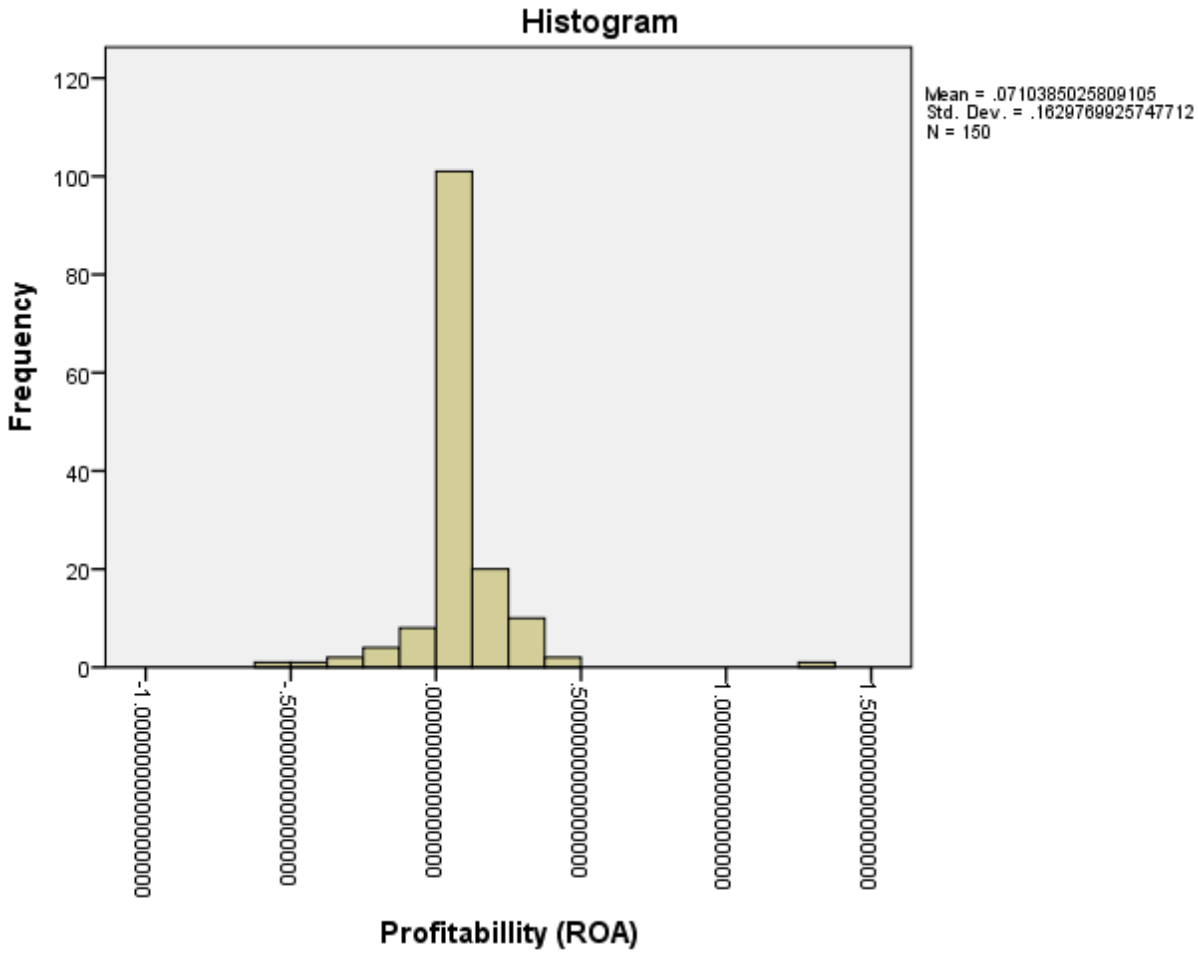

Figure 2: histogram for Y (RAO)

Comment : It is seen that the average of RAO is 0.071038502580910 with minimum -0.5003715076272441 and maximum 1.2549301691109305 while the standard deviation is 0.162976992574771 we can see the graph skewed to the right with outliers (exceeded 1)

The independent variables

First X1: Total Debt to Total Assets (TDTA)

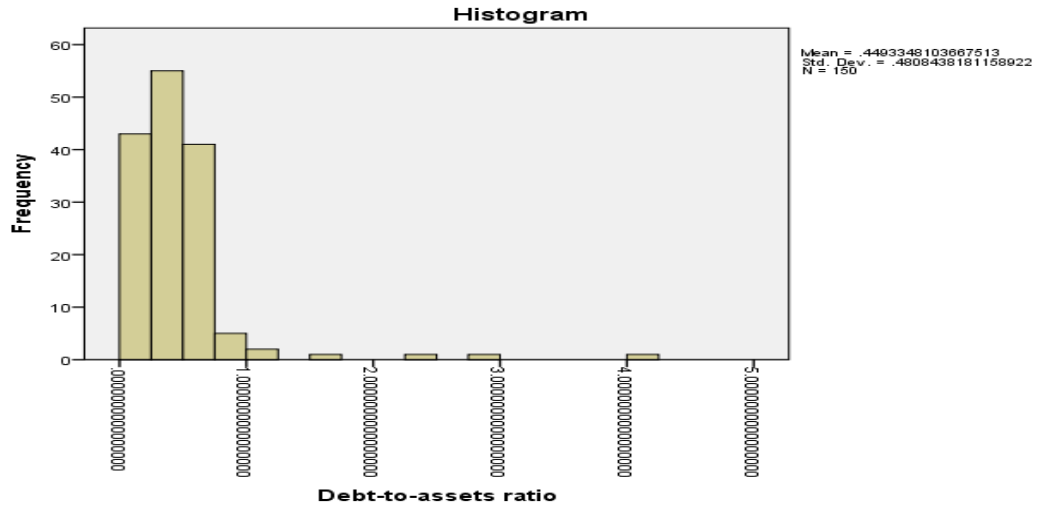

Figure 3: histogram for debt-to-assets ratio

Comment: it is seen the average is 0.449334810366752 with minimum 0 and maximum 4.1184441065546060 while the standard deviation is 0.480843818115892 we can see from the graph that it is skewed to the right with outliers greater than 1

Second X2: total debt to total equity (TD/TE)

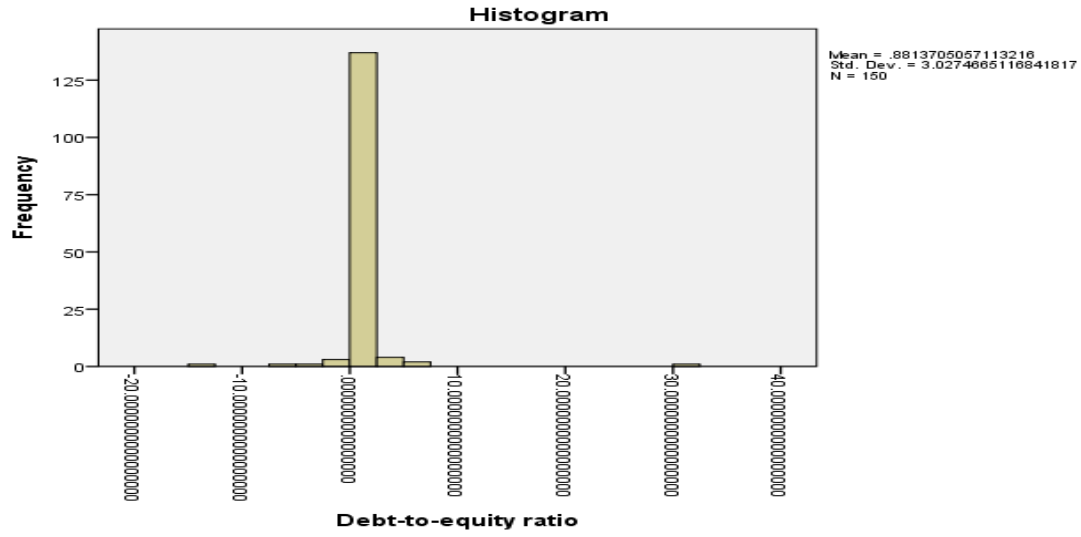

Figure 4 : histogram for debt to equity

Comment: it is seen that the average is 0.881370505711322 with minimum -13.0788659918609720 and maximum 31.7649740226311700 while the standard deviation is 3.027466511684184 we can see from the graph that it is skewed to the right with outliers greater than 30

Third X3:short-term debt (STD/TA)

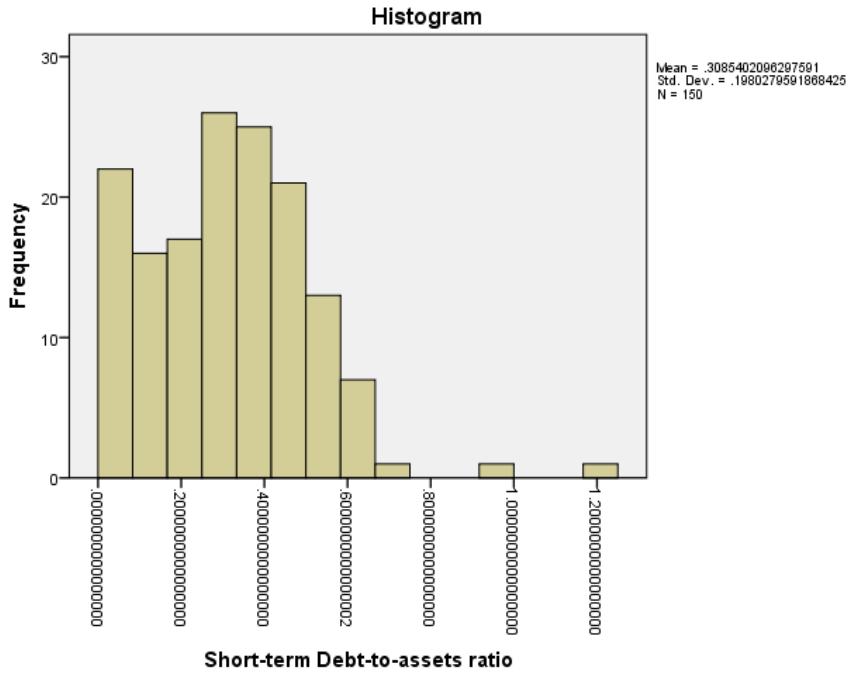

Figure 5: histogram for short term debt to assets

Comment: it is seen that the average is 0.308540209629759 with minimum 0 and maximum 1.2050236801854075 while the standard deviation is 0.198027959186843 we can see from the graph that it is skewed to the right with outliers greater than 1.3

Fourth X4: long-term debt (LTD/TA)

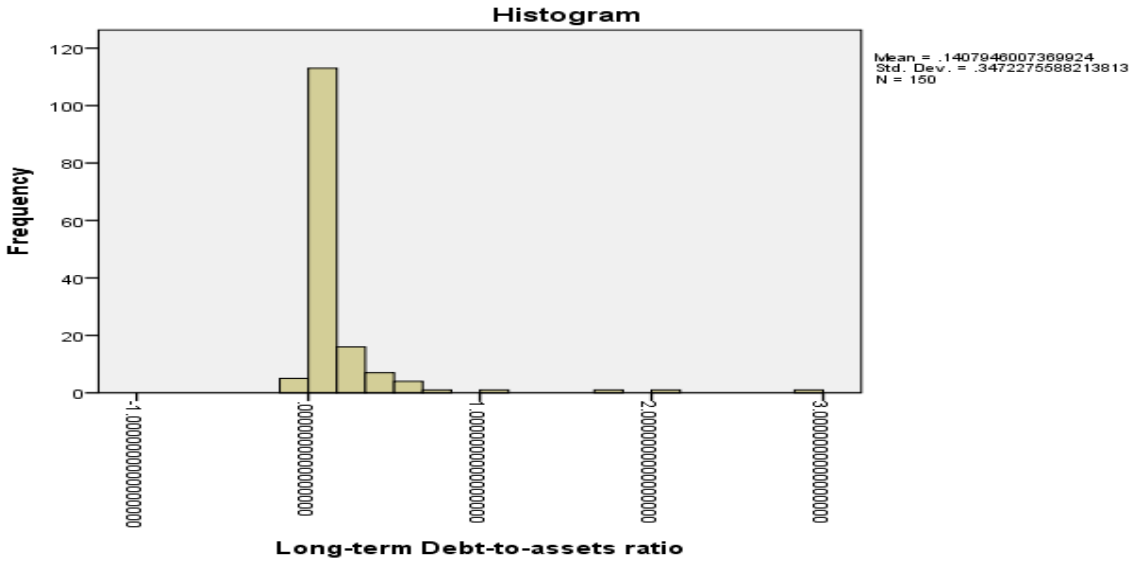

Figure 6: histogram for long term debt

Comment: it is seen that the average is 0.140794600736992 with minimum -0.0112637868153783 and maximum 2.913420426 while the standard deviation is 0.3472275588 we can see from the graph that it is skewed to the right with outliers greater than 1.5

Statistical Analysis

We are going to make regression analysis for our model but first we should remove the outliers from the model all of this using spss software

To remove the outliers, we used "casewise diagnostic" by enter method in spss and we removed 5 outliers

Hence, our sample size became 145 now

Now we will try regression analysis using different models to try to reach the best one

The first model using ENTER method without removing constant term

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.735[a]	.541	.531	.081308308658907

Figure 7

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1.098	3	.366	55.373	.000[b]
	Residual	.932	141	.007		
	Total	2.030	144			

Figure 8

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.077	.013		5.916	.000

	Debt-to-equity ratio	.000	.002	.005	.079	.937
	Short-term Debt-to-assets ratio	.042	.040	.070	1.049	.296
	Long-term Debt-to-assets ratio	-.260	.023	-.770	-11.490	.000

Figure 9

Excluded Variables[a]

Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics Tolerance
1	Debt-to-assets ratio	.[b]	.	.	.	.000

Figure 10

Comment : from figures 7,8,9,10 we can find this as a weak model as only one variable which is the long term asset is significant and there is high multicollinearity as debt to assets is removed (highlighted numbers)

After trying other model (the 4 other models) we also found the same problem of only the long term asset is significant

To solve this problem we will remove the constant term and now we will check what will happen

<u>The second model using ENTER method after removing the constant term</u>

After trying we removed an extra outlier now our sample size is 144

Our model:

Model Summary[c,d]

Model	R	R Square[b]	Adjusted R Square	Std. Error of the Estimate
1	.734[a]	.539	.529	.087675842947545

Figure 11

ANOVA[a,b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1.265	3	.422	54.869	.000[c]
	Residual	1.084	141	.008		
	Total	2.349[d]	144			

Figure 12

Coefficients[a,b]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	Debt-to-equity ratio	.001	.002	.029	.482	.631
	Short-term Debt-to-assets ratio	.233	.025	.678	9.302	.000
	Long-term Debt-to-assets ratio	-.291	.024	-.866	-12.316	.000

Figure 13

Comment : from figures 11 , 12 , 13 we can see that now 2 variables are significant not only one which are short term debt and long term debt (the highlighted numbers in figure 13)

And the Rsquared is 0.539 that means 54 percent of the dependent variable is explained by the independent variables (highlighted in figure 11)

And the highlighted part in figure 12 means the model is significant and good fit.

I tried also the 4 other models after removing the constant but they do not give any better.

By comparing model 1 and model 2 : the choice is the second model as 2 variables are significant not only 1

After choosing model 2 I will check the assumptions of the linear regression

First homodasticity

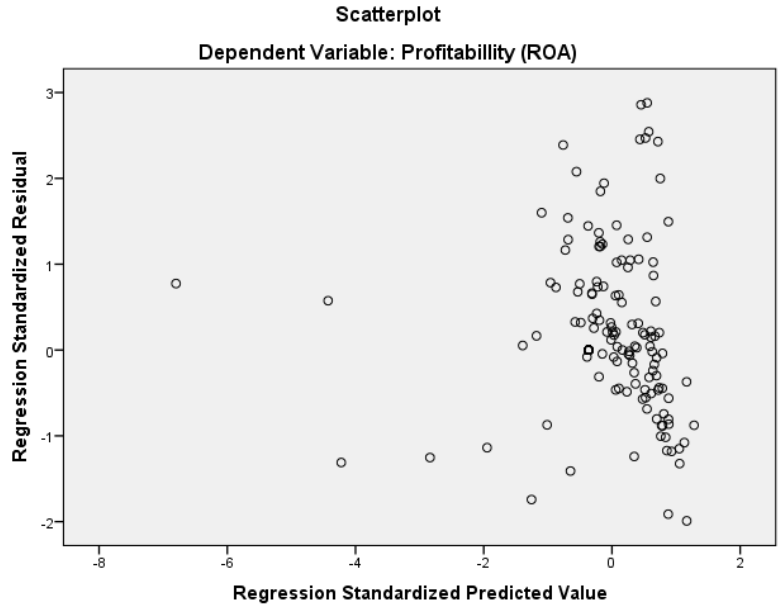

Figure 14: scatterplot for standardized residual vs standardized predicted

Comment: based on this pattern we can see that the scatterplot do not form any specific form so we can see homoscedasicity so assumption is valid

Second multicollinearity

Variable	Tolerance	VIF
Debt to equity	0.917	1.090
Short term debt	0.617	1.621
Long term debt	0.662	1.511

Comment: A variance inflation factor (VIF) is a measure of the amount of multicollinearity in regression analysis. Multicollinearity exists when there is a correlation between multiple independent variables in a multiple regression model. This can adversely affect the regression results. As the VIF in the 3 variables is close to 1 it means that there is no multicollinearity thus this assumption is valid

Third normality

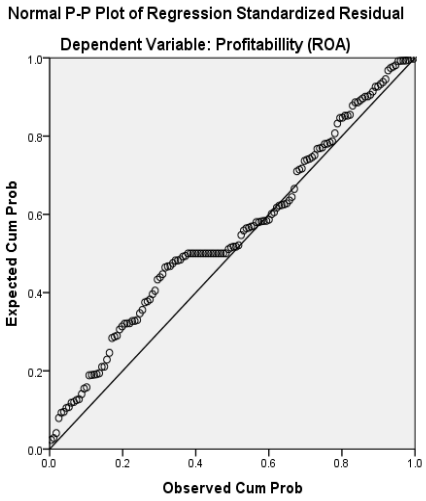

Figure 15: PP Plot

Comment: we can see deviations in the data so it is not completely normal but according to central limit theorem as the data exceeds 30 units it is approximately normal

Fourth auto correlation

Model Summary[c,d]

Model	R	R Square[b]	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.734[a]	.539	.529	.087675842947545	.660

Comment: K=3 , n= 144 from Durbin Watson table dL=1.584 and du=1.665 the Durbin-Watson= 0.660 as it relies between 0 and dL so this indicates positive auto correlation so this assumption is not valid.