International Journal of Innovative Research in Engineering and Management (IJIREM)
ISSN (Online): 2350-0557, Volume-10, Issue-6, December 2023
https://doi.org/10.55524/ijirem.2023.10.6.9
Article ID IJIR2484, Pages 60-65
www.ijirem.org

Relative Importance of Openness Index and Crude Oil Price in Explaining Growth of India's GDP During 1970-71 to 2019-20

Dipankar Pradhan[1] and Debasish Mondal[2]

[1] Research Scholar, Department of Economics, Vidyasagar University, Midnapore, India
[2] Professor, Department of Economics, Vidyasagar University, Midnapore, India

ABSTRACT- This paper seeks to evaluate the relative importance of the Openness Index (OPI) and Crude Oil Price (COP) to explain the growth of the GDP of India in the period from 1970-71 to 2019-20. For this purpose, this study uses a very popular and important method of evaluating the relative importance of predictors in the context of linear multiple regression analysis, viz., Budescu's Average Dominance (BAD) method (Budescu [1]. This study observes that two aspects of multicollinearity, viz., enhancement-synergism (ES) and change-in-sign (CS), not properly addressed in the existing literature, though observed in the majority of illustrations, are also not well taken by the above method. As a result, mainly for the second aspect, BAD methods lead to an improper evaluation of their relative importance. Under these circumstances, this study proposes a new method for evaluating True Relative Importance (TRI) of the predictors that is also able to assign their role. Using this new method this paper makes the concluding observation that openness index is more important and significant factor stimulating growth than crude oil price.

KEYWORDS- True Relative Importance, Ortho-partial Correlation, GDP. JEL Classification- C18, F43.

I. INTRODUCTION

How is the relative importance of different factors in explaining growth of India's GDP evaluated? Without a proper criterion for evaluating relative importance, the inquiries are meaningless as importance of a factor depends not only on the partial explanatory power of the explanatory variables, but also on their semi partial as well as simple explanatory powers. This paper diverges from the conventional approach of the multiple regression analysis, which typically employs marginal explanations to represent partial explanations also, and the relative importance of independent variables. Instead, this study utilizes the squared ortho-partial correlations [4], along with squared ortho-semi-partial correlations and squared simple correlations of the factors, as opposed to their marginal correlations. Understanding the nature of multicollinearity among factors is crucial when trying to determine the relative importance of different factors in social science data. This is because the method used to evaluate relative importance is dependent on the presence, nature and strength of multicollinearity. Special attention is needed to be given to Enhancement-Synergism (ES) and/or Change-in-Sign (CS). The present work tries to establish a proper criterion differing from the existing criteria in the literature for evaluating relative importance of different factors and to use it for evaluating true relative importance of different factors in explaining growth and fluctuation of India's GDP. This work proposes to use two most important factors for this purpose, viz., Openness Index (OPI) – the catch all external factor and the Crude Oil Price (COP) – the most crucial factor for economic stability.

II. REVIEW OF LITERATURE AND DISCUSSIONS

Budescu [1] has defined the General Dominance Index as the average increment in the coefficient of determination associated with predictor across all possible sub-models. Chao et. al [2] have tried to make a comparative study between squared zero-order correlation, squared semi-partial correlation, Product Measure (i.e., Pratt's Index), General Dominance Index, and Johnson's Relative Weight.

The task becomes further difficult in the presence of Enhancement-Synergism (ES) and/or Change-in-Sign (CS), none of which has been considered in the series of works mentioned here. In the present paper, we propose to determine the relative importance based on the path of explanation of any explanatory variable. This task of having the relative importance based on the path of explanation becomes further difficult as we move on to more than two explanatory variables, because in that case we have more than one path of explanatory power for each of them, and each path contains some interim values in between the starting value and the ending value, and there may be CS or ES or both in the points of those interim values.

III. OBJECTIVES

- To observe some relevant factors that can explain the variability of India's GDP
- To identify the contribution of these factors in explaining the growth of GDP.
- To evaluate the relative importance of the factors using very popular and interesting methods available in the literature and a proposed method, and to make a comparative analysis of them.

IV. DATABASE AND METHODOLOGY

The work is based on secondary data taken from the RBI handbook of statistics and the Our world in data, the crude oil price measured at a constant (2011–12) price in billions of dollars. GDP, and for the openness index, the export and import prices are measured at a constant (2011–12) price in crores of rupees.

To measure annual growth rate, we have used sei-log-linear trend regression given by,

$ln\,Y_t = a + bt + u_t \ldots \quad \ldots \quad \ldots$ (1)

where 'Y_t' is the relevant variable, 'ln' stands for natural logarithm, 'b' stands for the constant exponential annual growth rate (EAGR) of the variable (to express in percentage terms, it is to be multiplied by 100), 't' stands for time, and 'u_t' is the random error term.

Factor analysis for explaining the variability of an effect variable in terms of several factors is not an easy task. Multiple regression analysis is the commonly used method in this respect. In this method the relative importance of different factors is judged in terms of their partial correlations. However, partial correlations used in the existing literature succeed to explain neither the relative importance of the factors nor the partial significance of them; they can explain only the marginal significance of them Mondal [4] Moreover, in the presence of multicollinearity of several types among the factors, the application of this method fails to be satisfactory, and several methods are proposed in the literature to evaluate relative importance of the predictors. Among several methods of evaluating relative importance of the explanatory variables developed in the 80s and 90s of the 20th centuries, the method developed by Budescu [1] based on Average Dominance is found be very popular and relevant. The main idea behind average dominance method is to decompose the total variance explained by a regression model into the contributions of each predictor variable. Budescu [1] has done this by comparing the performance of all possible subsets of the predictors. The dominance of a predictor is then defined as the average increase in R-square when that predictor is added to all possible subsets of other predictors.

Additional contribution of			
	ρ^2	X1	X2
		$\rho_{y.x1}^2$	$\rho_{y.x2}^2$
K=0 average		$\rho_{y.x1}^2$	$\rho_{y.x2}^2$
X1	$\rho_{y.x1}^2$	----	$\rho_{y.x1x2}^2 - \rho_{y.x1}^2$
X2	$\rho_{y.x2}^2$	$\rho_{y.x1x2}^2 - \rho_{y.x2}^2$	----
K=1 average		$\rho_{y.x1x2}^2 - \rho_{y.x2}^2$	$\rho_{y.x1x2}^2 - \rho_{y.x1}^2$
X1X2	$\rho_{y.x1x2}^2$		
Overall average		Overall Average of X1	Overall Average of X2

(Source: own calculation based on Budescu [1])

Figure 1: Schematic Representation of Budescu's (1993) Average Dominance Analysis for Two Predictors

For a regression on two explanatory variables, K=0 average is also explained as the individual dominance, K=1 average as the interactional dominance of the variable for two cases but for three variable cases it is called partial dominance. The overall average of all these averages is defined as average dominance of the predictors, henceforth referred as Budescu's Average Dominance (BAD) and is used for measuring relative importance of them.

It is quite established in the existing literature that the presence of multicollinearity in the multiple regression model sometimes leads to the change-in-sign (CS) of some variables. As we shall see in the results and discussions section, this change-in-sign (CS) of any variable has its bearing on the relative importance of not only that variable, but also others. However, the above two methods, being based on squared correlations, fail to capture the effect of this change-in-sign (CS).

To overcome this problem, rather than approach the problem more directly and correctly, we propose a new method for evaluating the True Relative Importance (TRI) of the predictors. This method is like that of Budescu to some extent, and in some cases the results of this new method coincide with those of Budescu, but the approach is a bit different and improved.

The method for True Relative Importance (TRI) of the predictors can also be explained by figure 2. Under the assumption that the area of the whole rectangle is equal to 1 (one), the area covered by that designated as (1) can be explained by the incremental r-square when variable X1 is added to X2 as is explained by international dominance but for three variable or more variable cases it is partial dominance in Budescu's Average Dominance (BAD) method. This area (1) can also be obtained as the r-square when Y is regressed on $e_{1.2}$ ($e_{1.2}$ is the error variable obtained in the regression of X1 on X2, or $e_{1.2}$ is that part of X1 which is not linearly explained by X2) and this r-square is the true squared partial correlation between Y and X1 when X2 is held unchanged. This r-square is defined by Mondal [4] as the squared ortho-partial correlation (ortho means correct) between Y and X1 when X2 is held unchanged. This is done because the term squared partial correlation is defined in the existing literature by the r-square when $e_{Y.2}$ is regressed on $e_{1.2}$. However, as explained by Mondal [4], this is not the true partial correlation.

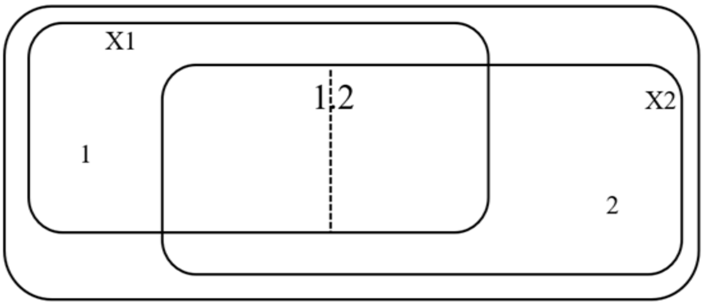
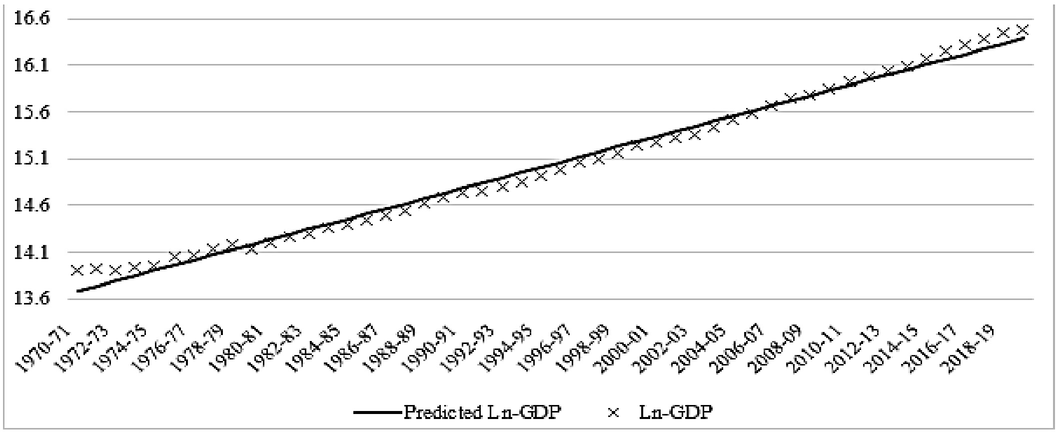


Figure 2: Venn Diagram Explaining the Variability of the Explained Variable by two Correlated Explanatory Variables

The True Relative Importance (TRI) of variable X1 is obtained as the average of the above squared ortho-partial correlation, the squared simple correlation but for three or more variable cases it is squared ortho-semi-partial correlations. When viewed in terms of ortho-partial correlation, ortho-semi-partial correlations and simple correlation, and not in terms of incremental squared correlations, we shall be able to observe any change in sign. If there is no change in sign, the result of the True Relative Importance (TRI) method coincides with the average dominance obtained in Budescu’s Average Dominance (BAD) method. But if there is change in sign of a variable, which can be identified for the True Relative Importance (TRI) method, then the above averaging of squared ortho-partial correlation, squared ortho-semi-partial correlations and squared simple correlation is done via 0 (zero), and the result differs from the average dominance obtained in Budescu’s Average Dominance (BAD) method. For example, if the ortho-partial correlation of a variable is negative and the simple correlation is positive, then whatever be the sign of theortho-semi-partial correlations, the averaging is needed to be done via 0 (zero) to obtain the relative importance of the variable. The relative importance of other variables is to be adjusted accordingly. The principle will be completely clear in our results and discussions section.

V. RESULTS AND DISCUSSIONS

This section presents results regarding the nature of growth in India's GDP during the period from 1970-71 to 2019-20. Growth is measured by the exponential annual growth rate (EAGR) obtained from the regression of ln-GDP on time (T).
The estimated regression line is: Ln-GDP = 13.6281 + 0.0552t with R-Square at 0.9904, Adjusted R-Square at 0.9902, F-value at 4969.75 with P-value = 3.97E-50. The estimated equation gives that the trend exponential annual growth rate (EAGR) of India's GDP is 5.52 percent for the period from 1970-71 to 2019-20. The nature of growth in GDP are presented in Figure 3.



(Source: authors’ own calculation)

Figure 3: India’s ln-GDP, its linear trend and fluctuation: 1970-71 to 2019-20

A. *Proposed Factors and their relevance in Growth of GDP*

To explain the growth of GDP, this paper proposes two factors, viz., Openness Index or OPI (X1), and Crude Oil Price or COP (X2) and tries to find out their relative importance. The relative importance analysis has been done through the Budescu’s Average Dominance (BAD) method [1] and the proposed methodology of evaluating True Relative Importance (TRI).
Stationarity tests on ln-GDP, and two proposed factors are considered to check for valid long-run relations. We have used the ADF and PP tests to check stationarity. Ln-GDP, X1 and X2 are found stationary at 1st deference (I(1)). So, there may exist a valid long-run relationship between Ln-GDP and the two variables X1 and X2.

To examine the degree and nature of the simple explanations of the variability Ln-GDP by the two explanatory variables (OPI (X1) and COP (X2)), the correlations among Ln-GDP and two explanatory variables are calculated and are presented in Table 1. It is observed that the simple correlations for all two explanatory variables with the explained variable are positive.

Table 1: Correlations Among Ln-GDP and Two Predictors

	Ln-GDP	OPI (X1)	COP (X3)
Ln-GDP	1		
OPI (X1)	0.92931	1	
COP (X2)	0.473079	0.676541	1

(Source: authors' own calculation)

B. Relative importance of OPI and COP in explaining the variability of Ln-GDP

To evaluate the relative importance of two explanatory variables using Budescu's Average Dominance method, we have performed a series of regressions to evaluate the incremental squared correlations and have taken the relevant averages. To explain the True Relative Importance of two explanatory variables we have performed a series of some new regressions to evaluate their squared ortho-partial correlations, squared ortho-semi-partial correlations up to two joint and squared simple correlations, have examined their change in sign effect and based on all these information their relative importance is calculated.

Two variables taken together explain 90.83% of the variability of the dependent variable which is significant with less than 1% level of significance. In the individual regression, both variables are significant with less than 1% level of significance. In Table 2, for variable X1 there is no change in sign. It has positive effects both in the simple regression and in the ortho-partial regression. But, the coefficient of X2 has a negative sign, which indicates that in the ortho partial section the variable X2 affects Y inversely, though as a whole X2 affects Y directly. The sign of the ortho partial correlation of X2 can also be identified from the regression of Y on e2.1 (where e2.1 is the residual of X2 when X2 is regressed on X1). Thus, there is a change in sign of the coefficient of X2 when we move from the simple regression to the multiple regression or from the simple regression to the ortho-partial regression. Complete results of the multiple regression are presented in Table 2.

Table 2: Regression Results of Ln-GDP on OPI (X1) and COP (X2)

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	13.99	0.07	188.21	2.73E-69
OPI (X1)	5.81E-02	3.10E-03	18.73	1.96E-23
COP (X2)	-0.001	0.00	-4.78	1.74E-05
R Square	Adj. R Square	Standard Error	*F*	*Sig. F*
0.9083	0.9044	0.250	232.73	4.1E-25

(Source: authors' own calculation)

This study first uses the BAD method to determine the average dominance of the two predictors and the results are presented in Table 3. The average dominance for X1 is found to be 0.7740, while the average dominance for X2 is observed to be 0.1342 in explaining the variability of Ln-GDP.

Table 3: Multiple Regression Results of Ln-GDP on OPI (X1) and COP (X2) for Budescu's Average Dominance Analysis

		Additional contribution of	
	$\rho_{y.x}{}^2$	OPI (X1)	COP (X2)
K=0 average		0.8636	0.2238
OPI (X1)	0.8636		0.0447
COP (X2)	0.2238	0.6845	
K=1 average		0.6845	0.0447
X1X2	0.9083		
Overall average		0.7740	0.1342

(Source: authors' own calculation)

Now, we present the results of the proposed method for evaluating the True Relative Importance (TRI) of the two predictors using ortho-partial and simple correlations in table 4. The squared ortho-partial correlation of X1 is significantly high, but for X2, it is very small, explaining very small parts of the variability of the dependent variable, but the relative importance values of these variables are not that small as the simple correlations are very high (Table 4).

Table 4: Multiple Regression Results of Ln-GDP on OPI (X1) and COP (X2) Using Proposed True Relative Importance (TRI) Method

	Squared ortho-partial correlations with sign	Squared simple correlations with sign	Unsigned Average	Signed Average	Relative Importance with sign
	1	1+1.2			
OPI (X1)	0.6845	0.8636	0.7740	0.7740	0.8187
	2	2+1.2			
COP (X2)	-0.0447	0.2238	0.1342	0.0896	0.0896
Total			0.9083	0.8636	0.9083

(Source: authors' own calculation)

Values in the first three columns of the table look very similar to those given in Table 3 implied for Budescu's Average Dominance analysis except the negative sign put before the squared ortho partial correlation of variable X2. In Table 2, we have seen that the coefficient of X2 has a negative sign, which indicates that the ortho-partial section of the variable X2 affects Y inversely, though as a whole X2 affects Y directly. The sign of the ortho-partial correlation of X2 can also be identified from the regression of Y on e2.1 (where e2.1 is the residual of X2 when X2 is regressed on X1). The value of the squared ortho-partial correlation of the variable, which comes out to be 0.0447, can also be obtained from the above regression of Y on e2.1. This squared correlation can also be evaluated as the incremental correlation as has been done by Budescu [1]. A negative sign is put before this 0.0447 to indicate that the unsquared ortho-partial correlation of this variable is negative. This will serve the purpose of evaluating the True Relative Importance (TRI) of the regressors in an interesting way.

From the results of Table 4, it is observed that out of 0.9083, a part amounting to 0.2685 is explained jointly by X1 and X2 and it is the outcome of a multicollinearity between them. This multicollinearity may be due to a near linear dependence of a part of X1 on X2 and no dependence of X2 on X1, or a near linear dependence of a part of X2 on X1 and no dependence of X1 on X2, or a two-way partial dependence of X1 on X2 as well as X2 on X1. When X1 is nearly dependent on X2, the explanatory power of X1 is just 0.6845 and that of X2 is 0.2238. On the other hand, when X2 is nearly dependent on X1, the explanatory power of X2 is 0.0447 and that of X1 is 0.8636. Finally, when there exists a two-way partial dependence of X1 on X2 as well as X2 on X1, the explanatory power of X1 will be greater than 0.6845, but less than 0.8636 and that of X2 will be greater than 0.0447, but less than 0.2238. Thus, depending on the nature of interdependence between X1 and X2, the explanatory power of X1 will vary from 0.6845 to 0.8636 and that of X2 will vary from 0.0447 to 0.2238.

For the average explanatory power or the relative importance of variable X2 will not be (0.0447+0.2238)/2=0.1342 as is given in the column of unsigned average of Table 4, or as is given in Budescu[1], because the coefficient of this variable changes its sign from negative to positive as we move from the ortho-partial regression to simple regression. The explanatory power of X2 does not vary directly from 0.0447 to 0.2238 (0.0447→ 0.2238), but it varies from 0.0447 to 0.2238 via 0 (0.0447→ 0 →0.2238), or, 0.0447 to 0 and then 0 to 0.2238. Thus, the average explanatory power of this variable can be obtained if we subtract half of (0.0447+0.2238), i.e., 0.1342 from 0.2238 and it comes out to be 0.0896. This can also be obtained in another interesting way. If we just put a negative sign (-) with 0.0477 and take a simple average of -0.0477 and 0.2238 as (-0.0447+0.2238)/2= 0.0896 which is significant with less than 5% level of significance, we shall have the same average explanatory power of X2, which is also the relative importance of the variable, hence, the concept of signed average as shown in the 5th column of Table 4.

However, for variable X1 the average explanatory power or the relative importance of the variable is not be (0.6845+0. 8636)/2=0.7740 as is given in the columns of unsigned average or signed average of Table 4, or as is given in Budescu, because the explanatory power of X1 does not vary directly from 0.6845 to 0.8636 (0.6845→ 0.8636), but it varies from 0.6845 to 0.8636 via 0.9083 (0.6845→ 0.9083 →0.8636), or, 0.6845 to 0.9083 and then 0.9083 to 0.8636. Thus, the average explanatory power of this variable can be obtained if we add 0.1342 with 0.6845 and it comes out to be 0.8187 which is significant with less than 1% level of significance. This can also be obtained in another way. If we compare the value of signed average with that of unsigned average for variable X2, we find that the value of signed average is less than that of unsigned average by 0.0447. Therefore, if we add this 0.0447 with the unsigned or signed average of X1, we have the required average explanatory power or the relative importance of variable X1. This is why this method can also be called the adjusted signed average method.

VI. CONCLUSION

The present study highlights the use of multiple regression method in evaluating the relative importance of the predictor variables in the presence of multicollinearity of different types. The study uses two different methods for evaluating relative importance of predictors, viz., the Budescu's Average Dominance (BAD) method and the presently introduced True Relative Importance (TRI) method to explain how in the presence of change in sign and a joint multicollinearity in excess of pair wise multicollinearity. The presently introduced True Relative Importance (TRI) method is not only able to assign correct average explanatory power to the predictors, but also able to evaluate the direction in which the predictor variable exerts its explanatory power.

By using the method of evaluating True Relative Importance, this study finds that two predictors, viz., Openness Index (OPI) and Crude Oil Price (COP) play significant direct role in explaining the growth of India's Gross Domestic Product in the period from 1970-71 to 2019-20, of which OPI is the most significant followed by COP. This has a clear policy implication that if the country can raise its openness, then growth of GDP will be significantly augmented.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

REFERENCES

[1] Budescu, D.V. (1993), Dominance analysis: A new approach to the problem of relative importance of predictors in multiple regression. Psychological Bulletin, 114(3), 542–551 https://doi.org/10.1037/0033-2909.114.3.542

[2] Chao YC, Zhao Y, Kupper LL, & Nylander-French LA. (2008), Quantifying the Relative Importance of Predictors in Multiple Linear Regression Analyses for Public Health Studies', Journal of Occupational and Environmental Hygiene, 5:8, 519-529, https://doi.org/10.1080/15459620802225481

[3] Hamilton, D.C. (1987), Sometimes R2 > ryx12 + ryx22 - Correlated Variables are not Always Redundant, The American Statistician, 41(2), 129-132. https://doi.org/10.1080/00031305.1987.10475460

[4] Mondal, D., (2008), On the test of significance of linear multiple regression coefficients, Communications in Statistics: Simulation and Computation, Vol. 37(4), https://doi.org/10.1080/03610910701753853