"DO PUBLIC EXPENDITURE AND DOMESTIC CONSUMPTION BECOME SIGNIFICANT IN GDP GROWTH OF INDIA AFTER 2011? "

Aditya Garg, Tushar, Sudhir Attri, Nitigya Pant

Introduction

1. The Gross Domestic Product (GDP) growth of a country is one of the most important measures of performance of an economy, and just as important is the accuracy of estimation of GDP for a country, economists are always trying to improve the methodology for estimating GDP. The government of India regularly revises the methods by periodic revisions so as to get a better picture of how the economy is performing. In the period 2011-2012 the data sources and methodology for GDP estimation was changed significantly, this was done to expand data and to move from volume based estimates of gross value added to value based estimates. After the results appeared from this change, the accuracy of these new changes were questioned by Pramit Bhattacharya (2019) providing evidence for problems in data used for new estimations. Arvind Subramanian (2019) published a detailed paper to test if the new methodology has caused misestimation in GDP Growth, This paper's objective is to further extend the study by Arvind Subramanian in verifying the same.
2. The original study (Subramanian, 2019) concludes that methodology changes introduced for the post 2011 GDP estimates have led to an overestimation supported by various robust and broad evidence. The paper is first establishes the hypothesis that new methodology has resulted in mis-estimation and then quantifies the mis-estimation.For this hypothesis, India is tested whether it follows the broad pattern of relationship between some indicators that co-move with GDP growth of a country, an outlier country suggest probable mis-estimation in calculation of GDP growth. These countries are selected on the basis of their similarity between their and India's economies so that a similar trend is expected in relation between indicators mentioned above and GDP growth. After dividing the data into pre-2011 and post-2011 and doing the same analysis India is found to be an outlier for only post-2011 period, implying in the direction that GDP estimation methodology changes are responsible for this out of line behaviour in the 2nd period. (Arvind Subra,2019)
3. The PMEAC published comments on the above mentioned paper (2) by Arvind Subramanian (2019) in the note GDP Estimation in India - Perspective and Facts (2019). In the same note, it is suggested that GDP growth drivers may have shifted to domestic consumption and public expenditure in post global financial crisis (2008) as an explanation for the mismatch between high frequency indicators and GDP growth. (Economic Advisory Council to the Prime Minister, 2019)
4. This paper re-formulate the analysis given in (2) using indicators mentioned in PMEAC note (3) and the original variables used by Arvind Subramanian. If India is found to be a normal country in post-2011 Period. This will serve as a hypothesis test for statement by PMEAC that indeed these are major growth drivers in the post 2011 period, and eventually this will give more confidence to the estimation of GDP done by govt. of India through the new methodology.

5. Data is collected for all the new indicators from World Bank's World Development Indicators (WDI) database[1] (Our World in Data, 2001-2016)

Our Approach

1. Arvind Subramanian's Model was taken into account and analysed after adding other growth indicators which might be the main indicators for the GDP growth and adding them might make the model more robust and consistent. The indicators included were related to public expenditure and domestic consumption.
2. Time series data corresponding to our new indicators was collected and appropriate regressions were generated to obtain different results from them.
3. OLS regression was used to generate the regression.
4. Different Statistical tests such as - robustness checks on choice of covariates , time span of study , data frequency , multicollinearity tests were calculated to judge the authenticity of the model. The detailed explanation is in Section-2.
5. Hypothesis Testing on the coefficient of India-Dummy to determine whether India is an outlier in a particular time period or not. Significant India-Dummy would lead to India becoming an outlier. And non-significant Dummy would lead to India not being an Outlier.
6. Concluded the outlier status of India by statistically inferring the new regression model which intricate the new indicators and commented on whether Arvind Subramanian's Report's result are intact or not.

Statistical Analysis

The main regression equation of the model is as follows, here the Gdp growth is being regressed with the following covariates. The β's are their corresponding coefficients and the ε_i is the error of the regression. The equation is:-

$$\text{gdp\_growth}_i = \beta_0 + \beta_1 credit\_\text{growth}_i + \beta_2 electric\_\text{growth}_i + \beta_3 \text{edu\_growth}_i + \beta_4 import\_\text{growth}_i + \beta_5 India + \beta_6 \text{household\_growth}_i + \beta_7 \text{military\_growth}_i + \beta_8 \text{employee\_growth}_i + \beta_9 \text{out\_growth}_i + \beta_{10} \text{health\_growth}_i + \beta_{11} \text{subs\_growth}_i + \beta_{12} \text{r.d\_growth}_i + \varepsilon_i \quad \text{--1}$$

The difference in difference model has also been taken into our analysis with the Dataset - 2 (Consists only of the BRICS nations) it has the following equation:

$$\text{gdp\_growth}_{it} = \beta_0 + \beta_1 credit\_\text{growth}_{it} + \beta_2 electric\_\text{growth}_{it} + \beta_3 \text{edu\_growth}_{it} + \beta_4^* import\_\text{growth}_{it} + \beta_5^* India + \beta_6 \text{household\_growth}_{it} + \beta_7 \text{military\_growth}_{it} + \beta_8 \text{employee\_growth}_{it} + \beta_9 \text{out\_growth}_{it} + \beta_{10} \text{health\_growth}_{it} + \beta_{11} \text{subs\_growth}_{it} + \beta_{12} \text{r.d\_growth}_{it} +$$

$\beta_{13}credit\_growth_{it}{}^{*}T + \beta_{14}electric\_growth_{it}{}^{*}T + \beta_{15}{}^{*}edu\_growth_{it}{}^{*}T + \beta_{16}{}^{*}import\_growth_{it}{}^{*}T + \beta_{17}{}^{*}India{}^{*}T + \beta_{18}household\_growth_{it}{}^{*}T + \beta_{19}military\_growth_{it}{}^{*}T + \beta_{20}employee\_growth_{it}{}^{*}T + \beta_{21}out\_growth_{it}{}^{*}T + \beta_{22}health\_growth_{it}{}^{*}T + \beta_{23}subs\_growth_{it}{}^{*}T + \beta_{24}r.d\_growth_{it}{}^{*}T + \beta_{29}{}^{*}T + \varepsilon_{it}$ --1'

Where the value of T=1 for post 2012 and zero otherwise and i denotes the countries

Multicollinearity

Before we run the regression on the selected variables, we needed to make sure that none of the independent variables are dependent on each other. If there is collinearity in independent variables, regression model ends up being inconsistent and standard errors of estimators may shoot up. We used some libraries already available in R to make sure that that there is no multicollinearity in our model, we checked pairwise correlation coefficients of all the variables, ran different multicollinearity tests and individual multicollinearity test of each variable.
The variables import_growth and export_growth were found to be correlated with a significant value, thus to make the estimation consistent, the variable export_growth was excluded from the model as import_growth is already contributing to the effect of export_growth as well. After that, we found no multicollinearity in our model in approximately all tests. Further data related to multicollinearity tests is in the Appendix. (Daoud, 2017)

Dataset for our Model

The dataset used consists of around 80 countries and is for the 2002 to 2016. To fillup the missing values for some parameters for a particular time period for a country is done by taking the average of growth of that variable by that country thus maintaining a statistically consistent model in terms of missing dataset values. The above method is adapted only when the country has at least 7 entries for that particular variable. The sources for data collection is mentioned in the data source table.

Regression Results

TABLE 1
2002-2011

India dummy	Household expenses Growth	Health expenditure Growth
1.338 [1.080]	0.552 *** [0.025]	-0.066 *** [0.015]

Min=-32.732, Max=44.060, Median=-0.029, Standard error=3.54
R^2=0.47, N=801
2012-2016

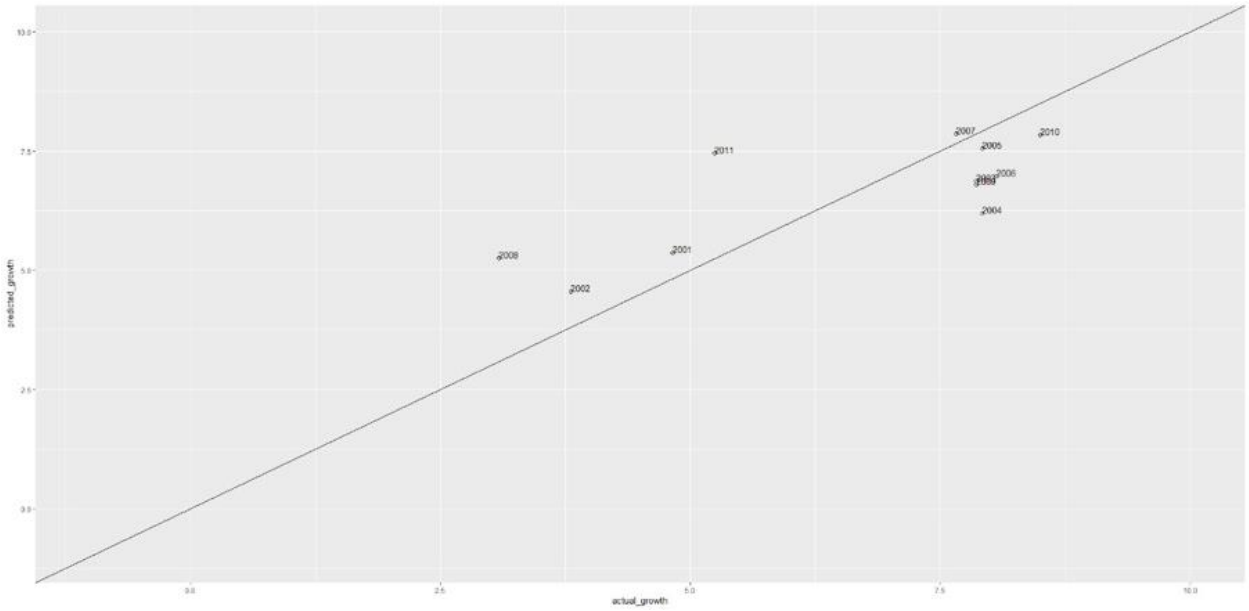
India dummy	Household expenses Growth	Health expenditure Growth
1.794 & [0.915]	0.509 *** [0.031]	-0.032 * [0.015]

Min=-6.7462, Max=10.6856, Median=-0.0043, Standard error=1.99
R^2=0.53, N=365
Significance codes: "***" 0.001, "**" 0.01, "*" 0.05, "&" 0.1

TABLE 2 (Pooled Regression)
Complete Table omitted, brief results :

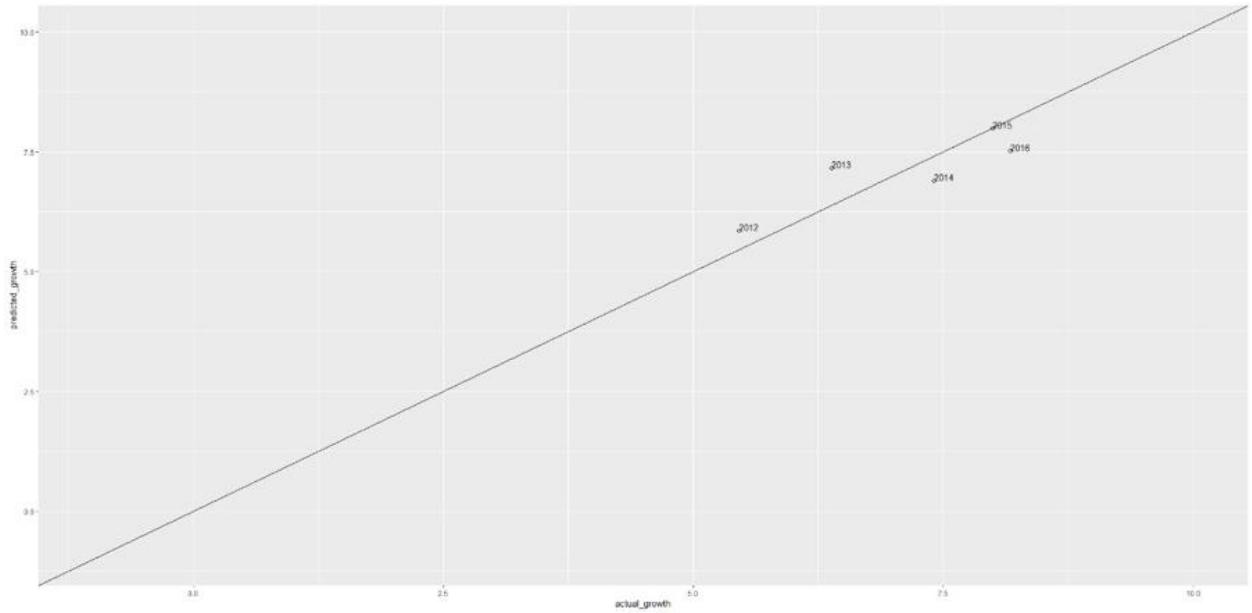
India dummy*T
0.663196 [1.191556]

Actual India GDP and fitted GDP in post 2011



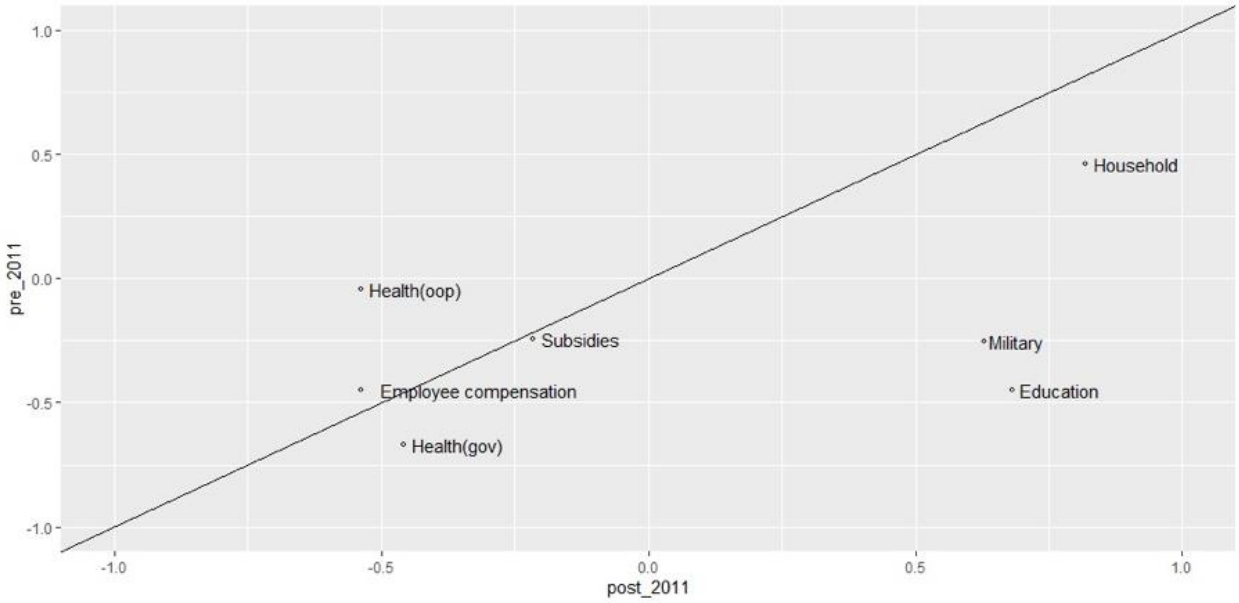
Graph 1(a) (line x=y drawn for readability for this and all following graphs)

Actual India GDP and fitted GDP in post 2011



Graph 1(b)

Correlation of selected indicators and GDP in pre 2011 and post 2011



Graph 2 (Correlation of gdp with indicators in post 2011 is plotted on X axis against the same but on pre 2011 period on X axis)

Robustness

A) Robustness with respect to Variables:-

Robustness is one of the most desirable properties of a Regression Model. Robustness helps to determine whether the model will change drastically if a covariate is added/removed from the model. A Robust model tend to change slightly in such a condition which leads to precise and consistent results. To check the robustness of our model, each covariate was omitted from the model and its effect was observed by observing the significance and the t-value of the other variables. The model turned out to be robust on all the variables except for the household_growth whose omission had a significant effect on the model which meant either houshold_growth is correlated with other variables or it is highly significant for our model . The second case turned out to be true.

B) Overall Robustness of our Model (Robust Linear Regression):-

A non-robust model tend to give misleading results when the data consists of high number of outliers as they have a significant effect on our coefficient estimation. Thus a robust-model is desirable, some tests already in R where used to check whether the model is robust or not. We compared our OLS model with that of the robust regression model and concluded that the estimation and the standard errors are almost similar thus our model is not misled by outlier points in our data. Further data related to Robust regression is in the Appendix. (Bruin, 2006)

Conclusions

1. From table 1(a) and 1(b) (Cross sectional regression), from the India (Dummy, $\beta_5{}^{*}India$) data we can see that India is a normal country for period 2002-2011 and an outlier with 90% confidence in period 2012-2016, this indicates that indeed india falls out of the line on the basis of relation between indicators and GDP but this inference is weaker than Subramanian's model for which the the $\beta_5{}^{*}India$ significant with 99% confidence in post 2011 period.
2. From table 2 (Pooled cross sectional regression), from the India (Dummy, $\beta_5{}^{*}India$) data, both $\beta_5{}^{*}India$ and $\beta_{17}{}^{*}India^{*}$T remains insignificant, implying that India is normal country with similar pattern of relationship between indicators and GDP growth, or rather we can say that introducing public expenditure and domestic consumption have resulted in this change in $\beta_{17}{}^{*}India^{*}$T.
3. Graph 1 indicates that there is no abrupt change in India's position in actual growth vs predicted graph in the post-2011 graph from pre-2011, significant deviation from x=y line would indicate abnormality. Graph 2 shows that five out of 7 new indicators introduced in our model become more correlated with GDP after 2012, this favours the hypothesis by PMEAC that these become major growth drivers in the post 2012 period.
4. Interpreting both (1) and (2) together, we can say that there is not enough conclusive evidence that adding indicators mentioned in PMEAC(2019) would result in India's GDP and indicator variables similar to other selected comparable countries. For table 2 the $\beta_{17}{}^{*}India^{*}$T is

significant only in **40%** confidence and also the $\beta_5{}^{*}India$ is shifted from 99% to 90% confidence in new model, therefore introducing these new indicators shifted the dummy variables to become more insignificant but they still have contribution, So change in methodology made India more different than other countries, but it has not shifted significantly to have conclusive evidence. Finally (3) supports the comment by PMEAC(2019) adding even more cloudiness in the conclusion. Taking all these factors together, Adding these particular variables did mitigated some of the problems explained by Arvind Subramanian (2019) but these are not sufficient to conclusively answer them.

APPENDIX

Complete Regression Table

2002-2011

India dummy	Household expenses Growth	Health expenditure Growth	Imports Growth	Electric consumption Growth	Credit Growth Growth	Military expenses Growth	Subsidies Growth	Research and Development Growth	Employee compensation Growth	Out of pocket health expenditure Growth	Education Growth
1.338 [1.080]	0.552 *** [0.025]	-0.066 *** [0.018]	-0.025 * [0.011]	0.131 *** [0.018]	0.00006 [0.0001]	-0.002 [0.008]	-0.012 [0.005]	-0.001 [0.003]	-0.025 & [0.013]	-0.065 [0.012]]	-0.046 ** [0.016]]

Min=-32.732, Max=44.060, Median=-0.029, Standard error=3.54

R^2=0.47, N=801

2012-2016

India dummy	Household expenses Growth	Health expenditure Growth	Imports Growth	Electric consumption Growth	Credit Growth Growth	Military expenses Growth	Subsidies Growth	Research and Development Growth	Employee compensation Growth	Out of pocket health expenditure Growth	Education Growth
1.794 & [0.915]	0.547 *** [0.031]	-0.032 * [0.015]	-0.049 ** [0.015]	-0.00002 [-0.00004]	0.007 [0.007]	-0.009 [0.01]	0.006 [0.007]	0.006 [0.007]	-0.013 [0.013]	0.023 * [0.012]	-0.013 ** [0.013]]

Min=-6.7462, Max=10.6856, Median=-0.0043, Standard error=1.99

R^2=0.53, N=365

Significance codes: “***” 0.001, “**” 0.01, “*” 0.05, “&” 0.1

Robust Regression analysis data:

2002-2011

India dummy	Household expenses Growth	Health expenditure Growth	Imports Growth	Electric consumption Growth	Credit Growth Growth	Military expenses Growth	Subsidies Growth	Research and Development Growth	Employee compensation Growth	Out of pocket health expenditure Growth	Education Growth
1.198 [0.064]	0.572 [0.015]	-0.078 [0.0093]	-0.12 [0.0063]	0.11 [0.001]	0.00001 [0.0001]	0.0046 [0.0047]	-0.013 [0.0075]	-0.0005 [0.0018]	-0.0336 [0.0078]	-0.0131 [0.0075]	-0.026 [0.0093]

Min=-37.26, Max=47.732, Median=-0.071, Standard error=1.839

N=801

2012-2016

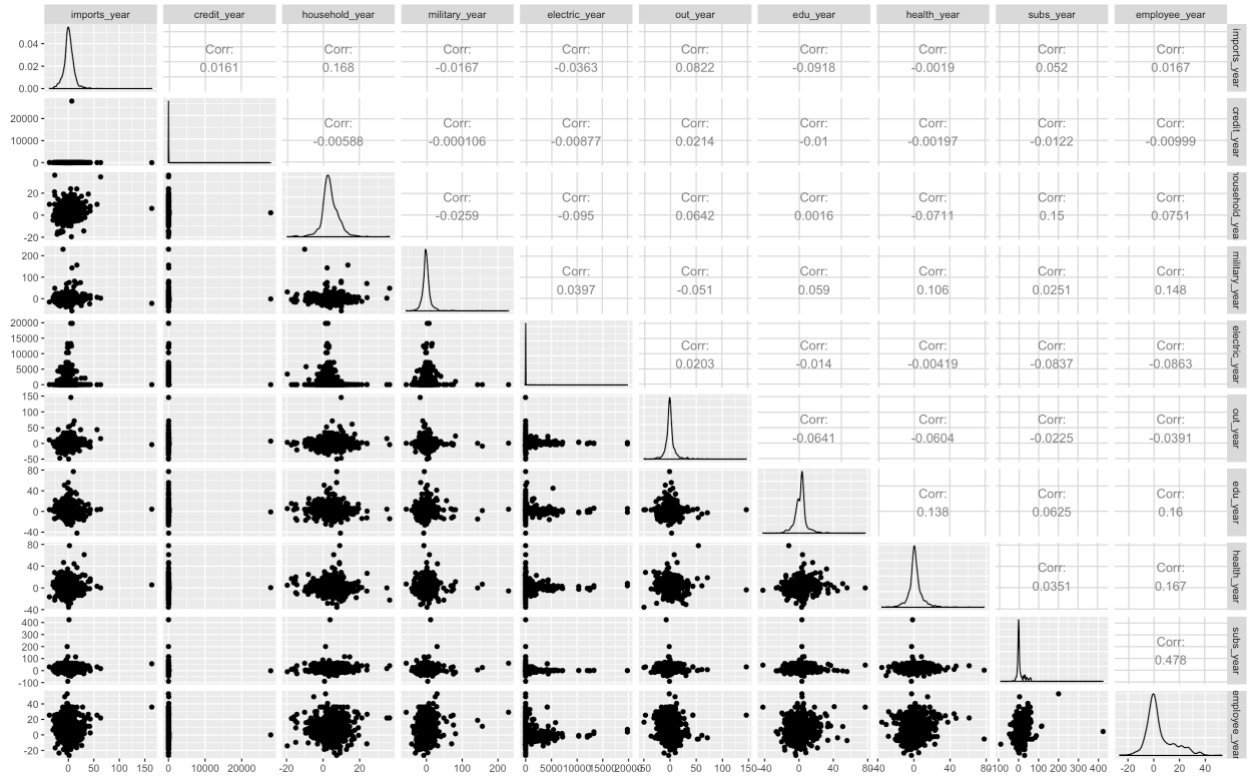
India dummy	Household expenses Growth	Health expenditure Growth	Imports Growth	Electric consumption Growth	Credit Growth Growth	Military expenses Growth	Subsidies Growth	Research and Development Growth	Employee compensation Growth	Out of pocket health expenditure Growth	Education Growth
1.634 [0.715]	0.602 [0.024]	-0.0352 [0.0122]	-0.0389 [0.0119]	0.00 [0.00]	0.0013 [0.007]	-0.009 [0.01]	0.0023 [0.0057]	0.0088 [0.0057]	-0.0125 [0.0103]	0.0191 [0.0080]	0.0073 [0.0115]

Min=-7.088, Max=10.770, Median=-0.0078, Standard error=1.373
N=365

Multi - Collinearity Data

Pairwise correlations:

The graph shows the pairwise correlation between all the variables in our model



Multicollinearity tests:

	MC Results	detection
Determinant |X'X|:	0.6252	0
Farrar Chi-Square:	545.1640	1
Red Indicator:	0.1021	0
Sum of Lambda Inverse:	11.0190	0
Theil's Method:	-3.1804	0
Condition Number:	2.7854	0

1 --> COLLINEARITY is detected by the test
0 --> COLLINEARITY is not detected by the test

DATA SOURCES

Older variables

Credit growth	Credit growth can simply be defined as the expansion of credit or money in the form of Loans to individuals or organizations which helps in maintaining the liquidity of the businesses in a country. Credit growth often means a healthy economy.	https://data.worldbank.org/indicator/FS.AST.PRVT.GD.ZS
Exports	The goods and services provided / sold by a country to other nations account for the countries exports. Exports is indeed the major contributor to the gdp of a country and thus essential in this problem.	https://data.worldbank.org/indicator/ne.imp.gnfs.zs

Variable - Public expenditure

Variable	Description	Source
Health	Health has always been a key factor in public expenditure throughout the world , government schemes for health improvement account for a large % of GDP spending by the government.	https://data.worldbank.org/indicator/SH.XPD.CHEX.GD.ZS
Education	Education is also considered one of the most important public expenditure factors throughout the world. After the right to education law in india , government expenditure on education increased by almost 1% of gdp	https://data.worldbank.org/indicator/SE.XPD.TOTL.GD.ZS
Employee Compensation	Government spending on Salaries and benefits of government employees account for a large % of gdp.	https://data.worldbank.org/indicator/GC.XPN.COMP.CN
Military Expense	Many countries spend a huge % of their gdp on military expenditure. Military expenditure is a crucial indicator of how powerful a country Is which is seen as a status symbol for a country.	https://ourworldindata.org/government-spending
Research and development	Government spends a high % of gdp on R&D sector to become technologically self reliant and provide advanced technologies to another nation's as well which in turn acts as a large part of foreign income for the country.	https://stats.oecd.org/Index.aspx?datasetcode=SOCX_AGG
Subsidies	Government provides incentives to both consumers as well as the producers with the aim of promoting social and economic policy.	https://data.worldbank.org/indicator/GC.XPN.TRFT.CN

Variable - Domestic consumption

Variable	Description	Source
Household consumption	Household expenditure means spending by an individual on day to day necessities such as food,clothing,shelter,etc. Which is an important factor in domestic spending.	https://data.worldbank.org/indicator/ne.con.prvt.cd
Health	Expenditure on health also accounts for a large part of domestic consumption. Insurance, mandatory checks are the major factors of health expenditure by the people.	https://data.worldbank.org/indicator/SH.XPD.OOPC.CH.ZS
Electricity	Important factor for the sustainability of the life.Thus a key part of the domestic consumption.	https://data.worldbank.org/indicator/EG.USE.ELEC.KH.PC

Works Cited:

Subramanian, A. (2019, June). *"India's GDP Mis-estimation: Likelihood, Magnitudes, Mechanisms, and Implications",CID Faculty Working Paper No. 354*. Retrieved from https://www.hks.harvard.edu/centers/cid/publications/faculty-working-papers/india-gdp-overestimate

Economic Advisory Council to the Prime Minister, B. D. (2019, June). *"GDP estimation in India-Perspectives and Facts"*. Retrieved from https://eacpm.gov.in/wp-content/uploads/2019/06/EAC-Paper_GDP-estimation_19-June-2019.pdf

Our World in Data. (2001-2016). *Data source*. Retrieved from 1. https://ourworldindata.org/government-spending

Daoud, J. I. (2017). *Multicollinearity and Regression Analysis*. Retrieved from IOP Science: https://iopscience.iop.org/article/10.1088/1742-6596/949/1/012009

Bruin, J. (2006). Retrieved from UCLA:Statistical Consulting Group. : https://stats.idre.ucla.edu/r/dae/robust-regression/