

You have shown most improvement from the last report

Good job. I appreciate your efforts

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

Aditya Garg, Tushar, Sudhir Attri, Nitigya Pant

### Introduction

- writing needs work  
- should go up

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. Consider rephrasing to clearly identify what you're doing. [SS]
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 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) or in an. [SS] suggesting [SS]
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.

↳ government [SS]

5. Data is collected for all the new indicators from World Bank's World Development Indicators (WDI) database<sup>1</sup> (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.

can be excluded [ss]

You could combine into one sentence. Significant - it does not need to be capitalized [ss]

## Statistical Analysis

The main regression equation of the model is as follows, here the <sup>GDP</sup>Gdp growth is being regressed with the following covariates. The  $\beta$ 's are their corresponding coefficients and the  $\varepsilon_i$  is the error of the regression. The equation is:-

$$gdp\_growth_i = \beta_0 + \beta_1 credit\_growth_i + \beta_2 electric\_growth_i + \beta_3 edu\_growth_i + \beta_4 import\_growth_i + \beta_5 India + \beta_6 household\_growth_i + \beta_7 military\_growth_i + \beta_8 employce\_growth_i + \beta_9 out\_growth_i + \beta_{10} health\_growth_i + \beta_{11} subs\_growth_i + \beta_{12} r.d\_growth_i + \varepsilon_i$$

--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:

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



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

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

Perfectly linearly related or collinear.

### 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 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)

VIF (SS)?

### 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.

statistical consistency is not directly related to how you fill in missing values. [SS]

### 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<sup>2</sup>=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<sup>2</sup>=0.53, N=365

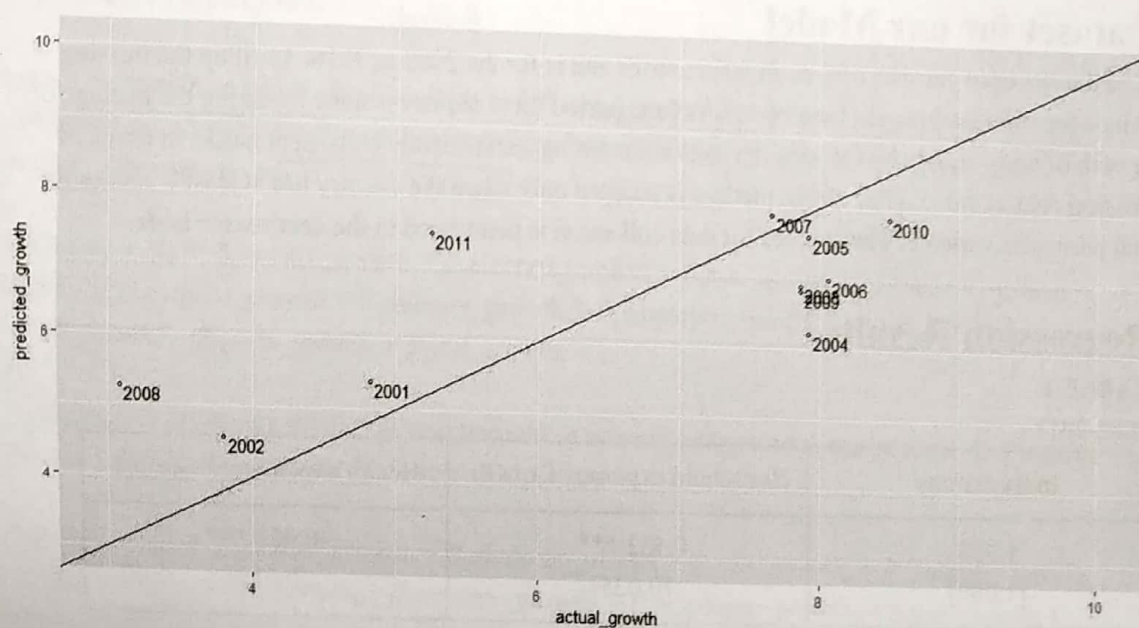
Significance codes: “\*\*\*” 0.001, “\*\*” 0.01, “\*” 0.05, “&” 0.1

**TABLE 2 (Pooled Regression)**

Complete Table omitted, brief results :

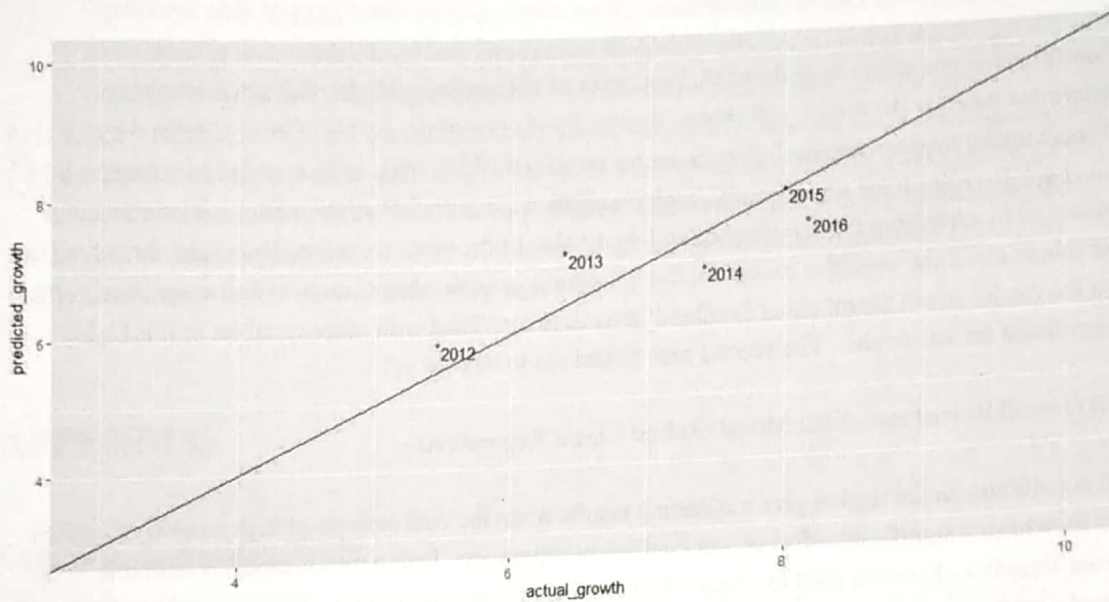
India dummy*T
0.663196 [1.191556]

### Actual India GDP and fitted GDP in post 2011



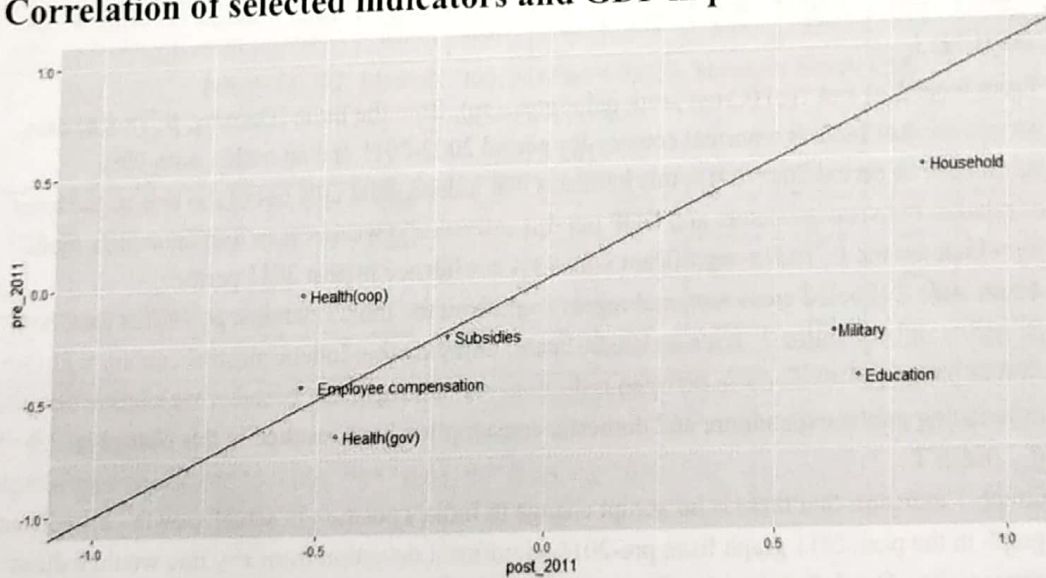
Graph 1(a) (line x=y drawn for readability for this and all following graphs)

## 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. 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 household\_growth is correlated with other variables or it is highly significant for our model. The second case turned out to be true.

[A] → consider replacing eq. [SS]

### 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<sup>2</sup>=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<sup>2</sup>=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	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	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]

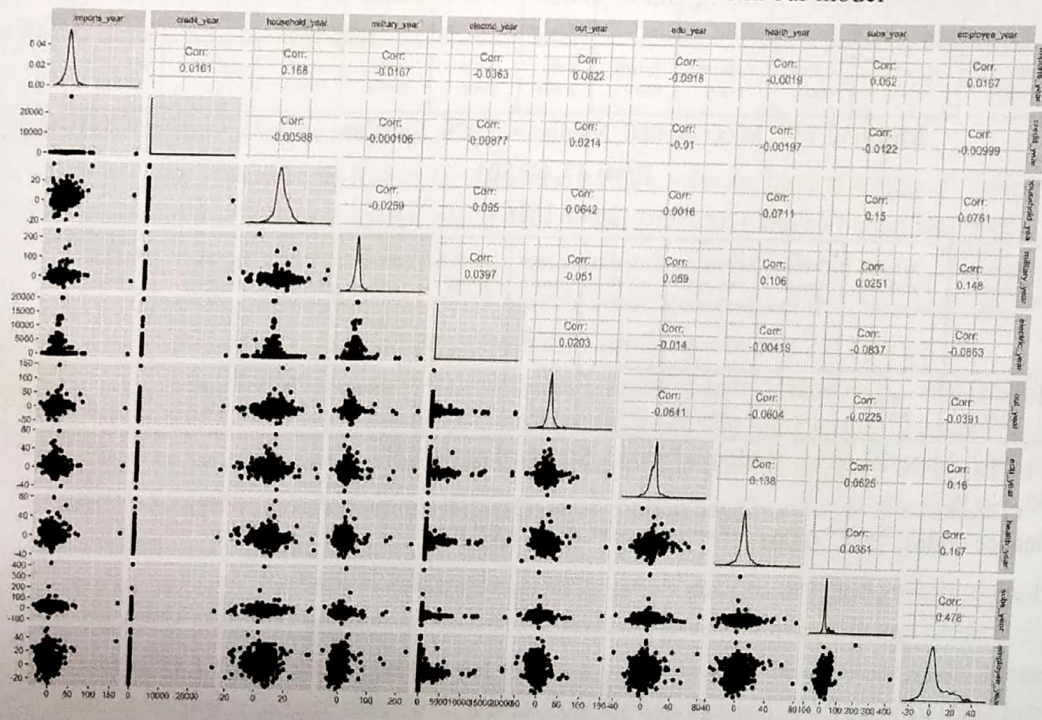
Min=-7.088, Max=10.770, Median=-0.0078, Standard error=1.373

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.	<a href="https://data.worldbank.org/indicator/FS.AST.PRVT.GD.ZS">https://data.worldbank.org/indicator/FS.AST.PRVT.GD.ZS</a>
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.	<a href="https://data.worldbank.org/indicator/ne.imp.gnfs.zs">https://data.worldbank.org/indicator/ne.imp.gnfs.zs</a>

### 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.	<a href="https://data.worldbank.org/indicator/SH.XPD.CHEX.GD.ZS">https://data.worldbank.org/indicator/SH.XPD.CHEX.GD.ZS</a>
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	<a href="https://data.worldbank.org/indicator/SE.XPD.TOTL.GD.ZS">https://data.worldbank.org/indicator/SE.XPD.TOTL.GD.ZS</a>
Employee Compensation	Government spending on Salaries and benefits of government employees account for a large % of gdp.	<a href="https://data.worldbank.org/indicator/GC.XPN.COMP.CN">https://data.worldbank.org/indicator/GC.XPN.COMP.CN</a>

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.	<a href="https://ourworldindata.org/government-spending">https://ourworldindata.org/government-spending</a>
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.	<a href="https://stats.oecd.org/Index.aspx?datasetcode=SOCX_AGG">https://stats.oecd.org/Index.aspx?datasetcode=SOCX_AGG</a>
Subsidies	Government provides incentives to both consumers as well as the producers with the aim of promoting social and economic policy.	<a href="https://data.worldbank.org/indicator/GC.XPN.TRFT.CN">https://data.worldbank.org/indicator/GC.XPN.TRFT.CN</a>

#### 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.	<a href="https://data.worldbank.org/indicator/NE.CON.PRVT.CD">https://data.worldbank.org/indicator/NE.CON.PRVT.CD</a>
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.	<a href="https://data.worldbank.org/indicator/SH.XPD.OOPC.CH.ZS">https://data.worldbank.org/indicator/SH.XPD.OOPC.CH.ZS</a>
Electricity	Important factor for the sustainability of the life. Thus a key part of the domestic consumption.	<a href="https://data.worldbank.org/indicator/EG.USE.ELEC.KH.PC">https://data.worldbank.org/indicator/EG.USE.ELEC.KH.PC</a>



## 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](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/>

① summary stats are missing.

② some Ds in writing (as suggested).



(14)

### Testing the hypothesis: "In the process, growth drivers shifted toward domestic consumption, and public expenditure."

#### Members

- Nitigya Pant - 2017355
- Sudhir Attri - 2018267
- Tushar - 2018201
- Aditya Garg - 2018124

#### Introduction

The Economic Advisory Council to the Prime Minister released a note, in which the contributors provided point to point rebuttals to the misestimation paper.

Consumers have proved to be messiahs for the Indian economy in the last three years, propping up demand for goods and services with their spending, even as the investment leg of the economy has stalled (2. Consumption expenditure is losing steam). Public investment have been the main contributors to growth in India in the last decade, while exports and private investment (and credit) have played a muted role (3. Bibek Debroy, Rathin Roy, Surjit Bhalla, Charan Singh, Arvind Virmani). Arvind Subramaniam claims that GDP growth is overestimated since 2011.

Bibek debroy claims hypothetically assuming the validity of Subramaniam's claim that a major explanation after NPA crisis of 2011 was that as global trade weakened, countries provided fiscal stimulus to their economies, growth thus shifted towards domestic consumption and public expenditure, a trend sharper for India. Thus contributors to growth changed significantly post global financial crisis towards domestic consumption and public expenditure.

The Arvind Subramaniam model attempts to test the India dummy and its behaviour in 2 different periods - pre 2011 and post 2011. The "i" in the equation is the different time periods.

$$GDP\ Growth_i = \beta_0 + \beta_1 Credit\ Growth_i + \beta_2 Electricity\ Growth_i + \beta_3 Export\ Growth_i + \beta_4 Import\ Growth_i + \beta_5 India + \epsilon_i$$

*Did he say it to you privately? Please provide proper citation.*

*Not sure if this is accurate*

*need to be complete in sentence*

*Subramaniam's*

*This not how you cite! Please see ~~the~~ research paper again to learn how to cite!!*

Table 1. Estimating India Dummy in Baseline Cross-Sectional Regressions

	Baseline		Baseline with outliers		Only MIC		All Countries	
	With Electricity	Without Electricity	With Electricity	Without Electricity	With Electricity	Without Electricity	With Electricity	Without Electricity
<b>2002-11</b>								
India	0.002	0.004	0.004	0.003	0.018	0.019	0.001	0.004
t-stat	0.37	0.63	0.80	0.48	4.51	4.06	0.22	0.71
R-sq.	0.64	0.50	0.64	0.53	0.43	0.35	0.64	0.51
<b>2012-16</b>								
India	0.028	0.040	0.036	0.045	0.029	0.037	0.027	0.039
t-stat	6.40	14.59	5.68	11.59	6.40	11.44	6.72	12.49
R sq.	0.74	0.65	0.63	0.59	0.68	0.62	0.72	0.65
# Observations	71	71	74	74	40	40	75	75
<b>Diff-in-Diff</b>								
India*Time	0.025	0.037	0.031	0.042	0.011	0.018	0.026	0.035
t-stat	3.25	5.94	3.73	6.42	1.78	3.18	3.59	5.88
R sq.	0.71	0.61	0.65	0.58	0.63	0.56	0.70	0.61
# Observations	142	142	148	148	80	80	150	150
# Countries	71	71	74	74	40	40	75	75

From the observations of table 1, the paper claims that India is an outlier among other countries because we can observe from the t-stat values that the dummy India is statistically insignificant for 2002-11 Period at 1% confidence level and significant at the 1% in 2012-16. Also Arvind Subramanian gives further assurity for the claim by doing the same analysis including and excluding 4 outlier countries.

As for robustness of the analysis, he claims to be robust to variation in measurement and placebo check only.

### Analytical Problem

*Bad stance by a researcher*

We want to verify or refute the claim in the note by The Economic Advisory Council to the Prime Minister that "In the process, growth drivers shifted toward domestic consumption, and public expenditure". To support this argument the contributors claim that domestic consumption and public investment have contributed more to growth in India in the last decade. Which implicitly implies that the difference between the values of India dummy for 2002-11 and 2012-16 is may be due to the fact that contributors to growth changed between these two time periods, which can be explained as a result of post-financial crisis for the reason stated earlier.

In order to verify this claim, we need to establish that fact that India is an outlier because of the selection of inaccurate growth indicators, in other words India dummy would be insignificant if correct

*Is it true though?*



indicators were used, therefore we estimate values of dummy India with correct indicators mentioned in the note and do the same.

The course of verification process that we choose will be:

1. (a). Finding the growth indicators which are the main contributors to growth post 2011 period claimed by rebuttal note. viz. indicators related to public expenditure and domestic consumption.  
 (b). Collecting time series data corresponding to indicators in 1(a).
2. Defining new regression model - modifying the equation (1) and (1)' by adding more indicator explanatory variables (As selected on step 1) and replicating Table 1 of Dr. by using the new Subramanian's paper by using OLS estimation of regression coefficient of India dummy variable.
3. Conducting robustness checks on choice of covariates (indicators in 1(a)), time span of study, data frequency. Tests for these checks explained in section 4.
4. Obtaining inference - Conduct a null hypothesis that the coefficient of dummy India is statistically similar in both cases within some confidence interval. Based on this test either verify or refute the rebuttal's statement.  
OR Conduct a null hypothesis that coefficient dummy India is statistically similar to zero, failure to reject the null hypothesis implies that dummy India is statistically insignificant and that india is not an outlier among other countries.

The original analysis stated above uses

**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.	<a href="https://data.worldbank.org/indicator/SH.XPD.CHEX.GD.ZS">https://data.worldbank.org/indicator/SH.XPD.CHEX.GD.ZS</a>
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	<a href="https://data.worldbank.org/indicator/SE.XPD.TOTL.GD.ZS">https://data.worldbank.org/indicator/SE.XPD.TOTL.GD.ZS</a>

Defer spend

Don't include infra spending if it's feasible

Employee Compensation	Government spending on Salaries and benefits of government employees account for a large % of gdp.	<a href="https://data.worldbank.org/indicator/GC.XPN.COMP.CN">https://data.worldbank.org/indicator/GC.XPN.COMP.CN</a>
Public procurement	Government spends a large amount of money on buying goods and services for the people. This covariate can be further divided into sub-covariates such as the expenditure on infra-projects, government outsourced services/goods etc.	<a href="https://ourworldindata.org/government-spending">https://ourworldindata.org/government-spending</a>
Social Spending	Includes housing, unemployment, old-age incentives etc.	<a href="https://stats.oecd.org/Index.aspx?datasetcode=SOCX_AGG">https://stats.oecd.org/Index.aspx?datasetcode=SOCX_AGG</a>

Variable - Domestic consumption

Variable	Description	Source
Food Consumption	Being a necessity, food plays a key role in the role in domestic consumption in every country	<a href="http://www.fao.org/figstata/en/#data">http://www.fao.org/figstata/en/#data</a>
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.	<a href="https://data.worldbank.org/indicator/SH.XPD.OOPV.CH.ZS">https://data.worldbank.org/indicator/SH.XPD.OOPV.CH.ZS</a>
Electricity and Water	Important factor for the sustainability of the life. Thus a key part of the domestic consumption.	<a href="https://data.worldbank.org/indicator/EG.USE.ALEC.KH.PC">https://data.worldbank.org/indicator/EG.USE.ALEC.KH.PC</a>
Energy and Transportation	In the developing countries, the spending on energy and transportation services increased exponentially and thus spending on them.	<a href="https://data.worldbank.org/indicator/IS.AIR.PSGR">https://data.worldbank.org/indicator/IS.AIR.PSGR</a> , <a href="https://data.worldbank.org/indicator/IS.RRS.PASG.KM">https://data.worldbank.org/indicator/IS.RRS.PASG.KM</a> , <a href="https://data.worldbank.org/indicator/EG.USE.PCAP.KG.OE">https://data.worldbank.org/indicator/EG.USE.PCAP.KG.OE</a>

is this  
expenses  
make  
sure  
that you  
include  
govt.  
spending  
as  
public  
expenditure  
and  
household  
spending as  
private expen-  
diture



### Statistical Inference of our Model:

Let the initial model be:

$$GDP\ Growth_i = \beta_0 + \beta_1 Credit\ Growth_i + \beta_2 Electricity\ Growth_i + \beta_3 Export\ Growth_i + \beta_4 Import\ Growth_i + \beta_5 India + \varepsilon_i \quad \text{-----1}$$

With considering the time as a factor:

$$GDP\ Growth_{it} = \beta_0 + \beta_1 Credit\ Growth_{it} + \beta_2 Electricity\ Growth_{it} + \beta_3 Export\ Growth_{it} + \beta_4 Import\ Growth_{it} + \beta_5 India + \beta_6 Credit\ Growth_{it} * T + \beta_7 Electricity\ Growth_{it} * T + \beta_8 Export\ Growth_{it} * T + \beta_9 Import\ Growth_{it} * T + \beta_{10} India * T + \beta_{11} * T + \varepsilon_{it} \quad \text{----- 1'}$$

The changed model after incorporating the domestic consumption and public expenditure indicators is:

$$GDP\ Growth_i = \beta_0 + \beta_1 Credit\ Growth_i + \beta_2 Electricity\ Growth_i + \beta_3 Export\ Growth_i + \beta_4 Import\ Growth_i + \beta_5 India + \beta_6 Health(govt.)\ Growth_i + \beta_7 Education(govt)\ Growth_i + \beta_8 Emp\ Compensation\ Growth_i + \beta_9 Procurement\ Growth_i + \beta_{10} Social\ Spending\ Growth_i + \beta_{11} Health(domestic)\ Growth_i + \beta_{12} Food\ Growth_i + \beta_{13} Electricity\ Growth_i + \beta_{14} Energy\ Growth_i + \beta_{15} Transportation\ Growth_i + \beta_{16} Others(consumption)\ Growth_i + \varepsilon_i \quad \text{----- 2}$$

we can do a null hypothesis for  $H_0: \beta_5 = 0$ , and through t-statistics if we found the null hypothesis to be true, we can conclude that India is not an outlier.

Taking time as a factor:-

$$GDP\ Growth_{it} = \beta_0 + \beta_1 Credit\ Growth_{it} + \beta_2 Electricity\ Growth_{it} + \beta_3 Export\ Growth_{it} + \beta_4 Import\ Growth_{it} + \beta_5 India + \beta_6 Health(govt.)\ Growth_{it} + \beta_7 Education(govt)\ Growth_{it} + \beta_8 Emp\ Compensation\ Growth_{it} + \beta_9 Procurement\ Growth_{it} + \beta_{10} Social\ Spending\ Growth_{it} + \beta_{11} Health(domestic)\ Growth_{it} + \beta_{12} Food\ Growth_{it} + \beta_{13} Electricity\ Growth_{it} + \beta_{14} Energy\ Growth_{it} + \beta_{15} Transportation\ Growth_{it} + \beta_{16} Others(consumption)\ Growth_{it} + \beta_{17} Credit\ Growth_{it} * T + \beta_{18} Electricity\ Growth_{it} * T + \beta_{19} Export\ Growth_{it} * T + \beta_{20} Import\ Growth_{it} * T + \beta_{21} India * T + \beta_{22} Health(govt.)\ Growth_{it} * T + \beta_{23} Education(govt)\ Growth_{it} * T + \beta_{24} Emp\ Compensation\ Growth_{it} * T + \beta_{25} Procurement\ Growth_{it} * T + \beta_{26} Social\ Spending\ Growth_{it} * T + \beta_{27} Health(domestic)\ Growth_{it} * T + \beta_{28} Food\ Growth_{it} * T + \beta_{29} Electricity\ Growth_{it} * T + \beta_{30} Energy\ Growth_{it} * T + \beta_{31} Transportation\ Growth_{it} * T + \beta_{32} Others(consumption)\ Growth_{it} * T + \beta_{33} * T + \varepsilon_{it} \quad \text{-----2'}$$

Where T is 1 for post 2011 period and zero otherwise.

Make sure you know that it is country not time/year.

\* Be careful about multicollinearity in your models.

Similarly, here we can do a null hypothesis for  $H_0: \beta_{21}=0$ , and through t-statistics if we found the null hypothesis to be true, we can conclude that India is not an outlier in post 2011 period.

### Robustness check

For checking the robustness of a linear regression model we will give a shock to the model by adding or removing some covariates, or changing some existing parameters and observe the nature of our model i.e. the effect of change on our coefficients. So, while incorporating the public Expenditure and domestic consumption covariates into our GDP model. We will try different permutations of those covariates and run regression on them to pick the model which turns out to be the most robust one. Mathematically:

Let the Initial Regression be  $\Rightarrow y = b_0 + b_1x + b_2z$

Let the Changed Regression after adding our variables be

$$\Rightarrow y = b_0 + b_1x + b_2z + b_3k$$

only public expenditure

only domestic consumption

So, we will check how other coefficients are changed if we add a new covariate k into our model and comment on the robustness accordingly.

- - and so on.

Hypothesis testing as a way of testing robustness - Hypothesis testing is a significant way of testing the importance of a particular covariate in our model.

In the above scenario after adding "k" covariate we can perform a hypothesis on k by taking  $H: b_3 = 0$ , if we fail to reject the null we can conclude that k is a significant covariate in our model and thus adding it will affect the other parameters much thus maintaining the robustness of our model.

- Time-span and data-frequency(robustness\*): To check the robustness of our model for different time periods we will run regression at different time intervals so as to check if there is a wide shift in our variables, what all factors are responsible for these changes and alter them appropriately so as to make our model function correctly at every time span, same with the data frequency.

Time span and data frequency(india dummy) - Let the regression equation be like:

$$\Rightarrow y = b_0 + b_1x + b_2z + b_3\text{India}$$

Where India = Dummy variable, 1 if country is India, zero otherwise.

To check the time-span consistency we will regress our model for different time spans and notice the change in the  $b_3$ , for example, if  $b_3$  comes out to be insignificant for 2001-2010, then it should also be insignificant for 2005-2008.

### References



1. <https://ourworldindata.org/government-spending>
2. Consumption expenditure is losing steam-  
<https://www.thehindubusinessline.com/opinion/editorial/consumption-expenditure-is-losing-steam/article27060728.ece#>
3. GDP estimation in India- Perspectives and Facts (Bibek Debroy, Rathin Roy, Surjit Bhalla, Charan Singh, Arvind Virmani) -  
<https://www.usebackpack.com/resources/23136/download?1568018147>
4. India's GDP Mis-estimation: Likelihood, Magnitudes, Mechanisms, and Implications:  
<https://growthlab.cid.harvard.edu/files/growthlab/files/2019-06-cid-wp-354.pdf>

APA citation style  
oxford citation style

Nice job!!

You seem to be on the right track.

I look forward to your work with  
utmost curiosity.

