# Overlooking Taxes in GDP (Mis)estimation.
## *Is this Justified ?*

**Paras Mehan**

*IIIT Delhi*
*New Delhi, India*

paras18062@iiitd.ac.in

**Ujjwal Singh**

*IIIT Delhi*
*New Delhi, India*

ujjwal18113@iiitd.ac.in

**Asmit Singh**

*IIIT Delhi*
*New Delhi, India*

asmit18025@iiitd.ac.in

**Archit Agrawal**

*IIIT Delhi*
*New Delhi, India*

archit18221@iiitd.ac.in

## Background and Motivation

"*India changed its data sources and methodology for estimating real gross domestic product (GDP) for the period since 2011-12*"[1].

Arvind Subramanian,( former Chief Economic Advisor of Indian Government and Currently, Professor at the Harvard School of Business ), published a paper "*India's GDP Mis-estimation: Likelihood, Magnitudes, Mechanism, and Implications*"[1] which proposes that the changes in the methodology of estimation have to lead to "*significant overestimation of growth*". The official estimates place annual average GDP growth between 2011-12 and 2016-17 at about 7 percent. Whereas the paper[1] that estimates "that actual growth may have been about 4½ % with a 95 % confidence interval of 3 ½ -5 ½ %. Hence, the author tries to prove that there is a misestimation of India's GDP in the Post Period.

In paper[1] the author divides the study[1] into two periods 2001-2011 and 2012-2016, which we will refer to as the Pre-2011 and Post-2011 periods. To establish the problem of overestimation of GDP, the author studies the correlation between 17 "real indicators" of economic growth, with annual GDP growth. These indicators include electricity consumption, 2-wheeler sales, petroleum consumption, cement, steel, etc.
The argument is that if there is no misestimation or overestimation of India's GDP in the post-period, then the correlation between these economic indicators and GDP growth should be similar in the Pre and Post Period.

After establishing a problem, the author identifies specific indicators that co-move with growth and are easy to produce. Then, the author relates these indicators with the GDP growth rate for a comparable set of countries for the pre-2011 and post-2011 period, to study whether India shows the same trend as compared to other countries or it is an outlier in one or both of these periods. The point that we need to note here, is that the

author has not taken tax as an indicator of GDP in his study. In the follow-up paper "*Validating India's GDP Growth Estimates*"[2], he provides an argument for not including tax in his original paper[1].

The argument in the follow up paper[2] is based on the fact that the India's Tax-GDP ratio rose in the Post-2011 period from 10% in 2011-12 to 11% in 2016-17, which suggests that the "*rising revenue-GDP ratios tend to suggest surging growth.*"[2] However, as stated in the paper[2], "*revenues are affected by more than just economic growth; they are also affected by changes in tax policies and administration*". Hence, due to factors like increase in-direct tax collection due to the fall in petroleum prices in the international markets in 2014[3] and decreasing direct tax growth rate in the post-2011 period, taxes cannot be used to infer much about GDP growth. Therefore, the author has not taken tax as an indicator of GDP in his study[1].

# Problem / Analysis Query

We propose to perform an extensive econometric analysis of these claims and test whether tax data should have been included or not included while testing the misestimation of the GDP of India.

The Economic Advisory Council to the Prime Minister (EAC-PM) in its rebuttal[4] to Arvind Subramanian paper[1] stated that :

"*Unlike many indicators, tax data is not collected through surveys or by agencies through arcane techniques, these are hard numbers and should be an important indicator of growth. Further, there have been no major changes in tax laws until the end period in the author's analysis (31st March 2017). GST was introduced on 1st July 2017. The author's logic of not using tax data appears to be a convenient argument meant to avoid inconvenient conclusions based on hard facts.*"

The EAC-PM raised concerns over the fact that Arvind Subramanian didn't include taxes in his study[1] , but as stated, taxes are "*hard numbers and should be an important indicator of growth.*" EAC-PM also argues that the tax plays an essential role in the GDP growth rate of the country, and should be considered as one of the fundamental indicators of growth of a country.

Therefore, we propose to perform an extensive econometric analysis of these claims and test whether tax should have been included or not included while testing the misestimation of the GDP of India.

1. *Prima facie*, is there a problem?

We will first test whether a problem exists in the estimation of GDP. The author has done the same analysis by calculating the correlation between India's GDP growth and Growth Indicators for two periods. We will perform the corresponding analysis, but with the addition of Tax in the growth indicators.

Using this, we can infer :

- Whether tax behaves similar to other indicators in the two periods. Which can be used to test the EAC-PM argument.
- Is there a difference in the relation between GDP and tax, in the Pre-2011 and Post-2011 period. We are thereby testing whether a problem of misestimation exists or not.

## Regressions

$$GDP\ Growth_i = \beta_0 + \beta_1 * Credit\ Growth_i + \beta_2 * Export\ Growth_i + \beta_3 * Import\ Growth_i + \beta_4 * India\ Dummy + \varepsilon_i$$

---------- (1)

$\beta_1$ - Represent the magnitude change in GDP with respect to a unit change in Credit Growth
$\beta_2$ - Represent the magnitude change in GDP with respect to a unit change in Export Growth
$\beta_3$ - Represent the magnitude change in GDP with respect to a unit change in Import Growth
$\beta_4$ - Coefficient of dummy variable

$$GDP\ Growth_i = \beta_0 + \beta_1 * Credit\ Growth_i + \beta_2 * Export\ Growth_i + \beta_3 * Import\ Growth_i + \beta_4 * India\ Dummy + \beta_5 * Tax(\%\ of\ GDP)\ + \varepsilon_i$$

---------- (2)

$\beta_1$ - Represent the magnitude change in GDP with respect to a unit change in Credit Growth
$\beta_2$ - Represent the magnitude change in GDP with respect to a unit change in Export Growth
$\beta_3$ - Represent the magnitude change in GDP with respect to a unit change in Import Growth
$\beta_4$ - Coefficient of dummy variable
$\beta_5$ - Represent the magnitude change in GDP with respect to a unit change in Tax

$$GDP\ Growth_i = \beta_0 + \beta_1 * Credit\ Growth_i + \beta_2 * Export\ Growth_i + \beta_3 * Tax(\%\ of\ GDP) + \beta_4 * Import\ Growth_i + \beta_5 * India\ Dummy*t + \beta_6 * India\ Dummy + t + \beta_7 * Credit\ Growth_i*t + \beta_8 * Export\ Growth_i*t + \beta_9 * Import\ Growth*t + \beta_{10} * Tax(\%\ of\ GDP)*t + \varepsilon_i$$

---------- (3)

$\beta_1$ - Represent the magnitude change in GDP with respect to a unit change in Credit Growth
$\beta_2$ - Represent the magnitude change in GDP with respect to a unit change in Export Growth
$\beta_3$ - Represent the magnitude change in GDP with respect to a unit change in Tax
$\beta_4$ - Represent the magnitude change in GDP with respect to a unit change in Import Growth
$\beta_5$ - Coefficient of dummy variable for post period
$\beta_6$ - Coefficient of dummy variable for India
$\beta_7$ - Represent the magnitude change in GDP with respect to a unit change in Credit Growth in post period
$\beta_8$ - Represent the magnitude change in GDP with respect to a unit change in Export Growth in post period
$\beta_9$ - Represent the magnitude change in GDP with respect to a unit change in Import Growth in post period

$\beta_{10}$ - Represent the magnitude change in GDP with respect to a unit change in Tax in post period

---

## 2. Omitted Variable Bias

Omitted variable bias occurs when a relevant variable is excluded from a statistical model.
We will test for omitted variable bias in the regression (1), due to the exclusion of tax. In this, we will calculate correlation of Tax Revenue(% of GDP) with the GDP Growth Rate, Domestic credit to the Private sector, Exports of Goods and Services and Imports of Goods and Services.
We will analyze these correlation values, and see if these conditions for Omitted variable bias exist.
   1. The tax variable is significantly related to one of the predictors.
   2. Tax variable is a significant determinant of GDP growth.
If the two conditions above follow, it would result in the violation of OLS assumption E(UX)=0, which would result in an omitted variable bias.

We will also compare the adjusted R-squares of the regression models (1) and (2), to find how well the variation in the GDP growth is explained with and without inclusion of Tax.

## 3. Regression Analysis

To study the relation of taxes with GDP, the author has studied the regression (1) and regression (3)( with the exclusion of taxes). We will perform the same regression(1) and regression (2), (which is the same as regression (1), but with the inclusion of taxes), for the period ( Pre-2011 and Post 2011) for a set of countries.

For the regression (3) the author states that :
"*Statistically speaking we are deploying the spirit of a "difference-in-differences" technique. Here the treatment is the methodology change in India; the treatment period is post-2011. We are then testing whether the treatment had a differential impact on the relationship between the indicators and GDP growth in the post-2011 period: put differently, was India differentially affected in the post-2011 period compared to countries.*"(Arvind Subramanian,2019, p. 9)

Since we are interested in understanding what happens after we add the tax to the original analysis in the paper, There we will perform the regression (3) using which we can infer whether India is still an outlier in the Post-2011 period or not.

Note:

1. Electricity: We have not selected electricity consumption along with other Indicators. The government of India has put in policies to achieve 100 percent electrification during the post-2011 period[5]. Due to which electricity consumption for India could behave differently from the other countries selected. Electricity can, therefore, became a proxy indicator of India dummy variable, since it behaves very differently as compared to other countries. Due to which we would not be able to test whether India is an outlier, using India's dummy variable.

2. Period: We have selected the post-2011 period from 2011 to 2015 and excluded 2016 from our analysis as demonetization happened in India in November 2016, which served as an internal shock to the Indian economy. Therefore, we have decided no to include 2016 in our analysis.

3. Countries Selection: We have selected the top 50 countries, listed by their GDP as per the United Nations. Since we needed a group of countries whose GDP is comparable to India, hence we have selected the top Countries ranked by their GDP in our analysis.

4. Tax: To include tax as an indicator of GDP, in the original analysis, we have included Tax Revenue(% of GDP). Because it is a percentage, not an absolute value, therefore can be used to compare Tax revenue of Different countries easily.

## Variables and Their Descriptions.

| S No. | Variable | Description | Notation |
|---|---|---|---|
| 1 | GDP growth (annual %) | The annual GDP growth of a country | GDP Growth |
| 2 | Domestic credit to private sector (% of GDP) | Domestic credit to private sector refers to financial resources provided to the private sector by financial corporations, such as through loans, purchases of non-equity securities, and trade credits and other accounts receivable, that establish a claim for repayment.[7] | Credit Growth |
| 3 | Exports of goods and services (annual % growth) | The annual growth rate of exports of goods and services based on constant local currency. Aggregates are based on constant 2010 U.S. dollars. Exports of goods and services represent the value of all goods and other market services provided to the rest of the world. They include the value of merchandise, freight, insurance, transport, travel, royalties, license fees, and other services, such as communication, construction, financial, information, business, personal, and government services. They exclude compensation of employees and | Exports Growth |

| 4 | Imports of goods and services (annual % growth) | The annual growth rate of imports of goods and services based on constant local currency. Aggregates are based on constant 2010 U.S. dollars. Imports of goods and services represent the value of all goods and other market services received from the rest of the world. They include the value of merchandise, freight, insurance, transport, travel, royalties, license fees, and other services, such as communication, construction, financial, information, business, personal, and government services. They exclude compensation of employees and investment income (formerly called factor services) and transfer payments.[9] | Import Growth |
| 5 | Tax revenue (% of GDP) | Tax revenue refers to compulsory transfers to the central government for public purposes. Certain compulsory transfers such as fines, penalties, and most social security contributions are excluded. Refunds and corrections of erroneously collected tax revenue are treated as negative revenue.[10] | Tax Growth |

**Note :**
 - All Variables in the above table were acquired from WDI Database([Link](#)).

## Summary Statistics

| Sno | Variable Name | Mean | Standard Deviation | Minimum | Maximum |
|-----|---------------|------|--------------------|---------|---------|
| 1. | GDP growth (annual %) | 3.315 | 4.2 | -33.10 | 54.16 |
| 2. | Domestic credit to private sector (% of GDP) | 82.54 | 51.8 | 0 | 206.67 |
| 3. | Exports of goods and services (annual % growth) | 4.94 | 8.9 | -30.01 | 85.61 |

| 4. | Imports of goods and services (annual % growth) | 5.46 | 10.74 | -50.05 | 84.74 |
| 5. | Tax revenue (% of GDP) | 14.6 | 8.90 | 0 | 36.50 |

# Results

## 1. Relation Between GDP And Indicators

The correlations of GDP and indicators is depicted in figure.
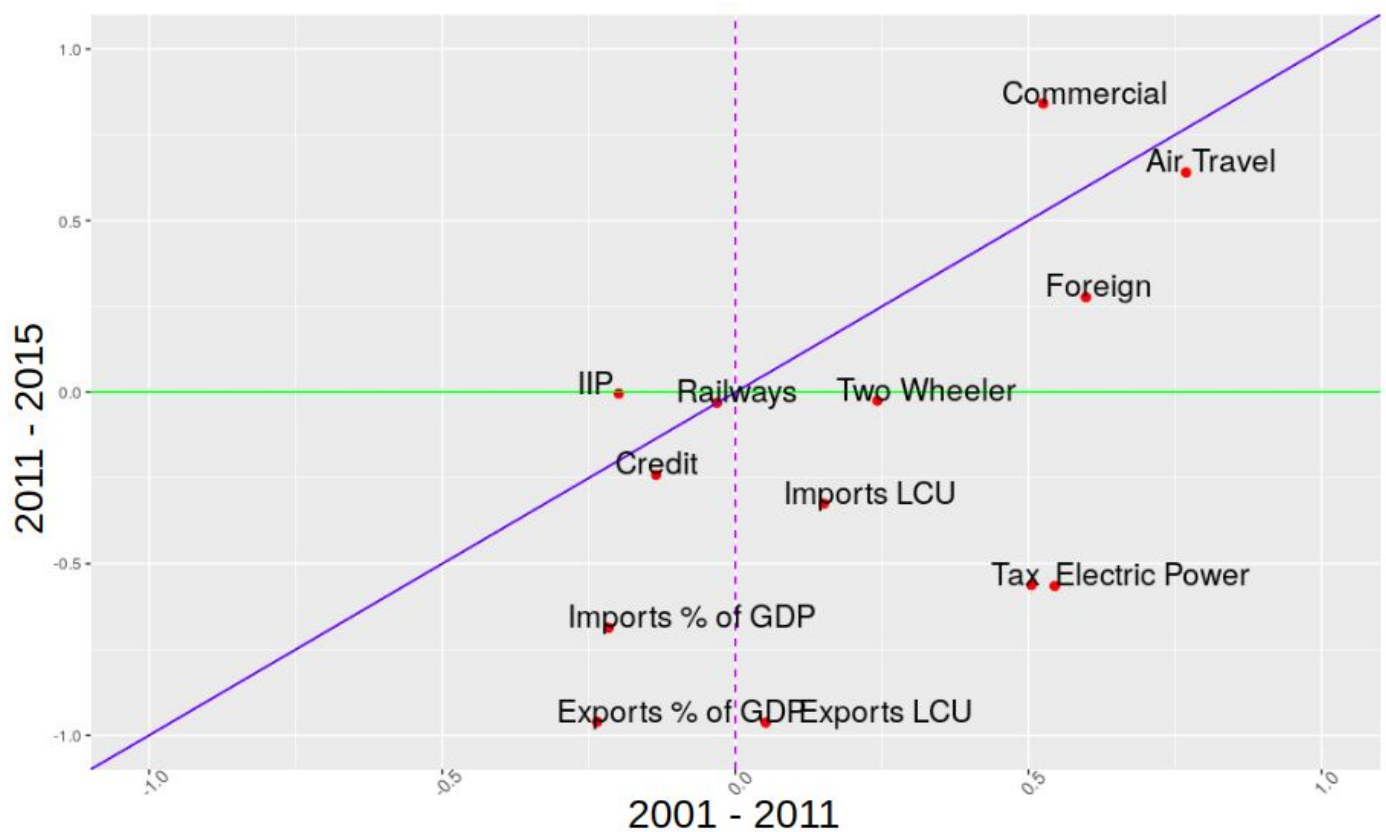
- Each point X-coordinate depicts the correlation of GDP Growth rate and the growth of the indicator rate for the Pre-2011 period, and the Y-coordinate depicts the same for Post 2011 period.
- What we find is that 8 out of 13 indicators are positively correlated with the GDP growth in the Pre-2011 period, whereas 10 of these 13 indicators have a negative correlation with the GDP growth in the Post-2011 period.
- The Tax Revenue(% of GDP) (referred to as Tax hereafter), has +0.5 correlation with GDP Growth Rate in Pre-2011 period with p-value of 0.11 at 95% confidence level. However, with -0.5 correlation with GDP Growth Rate with p-value of 0.44 at 95% confidence level.
- Thus, we can infer that Tax is behaving similar to other growth indicators, which were changing like Electricity, Steel etc.
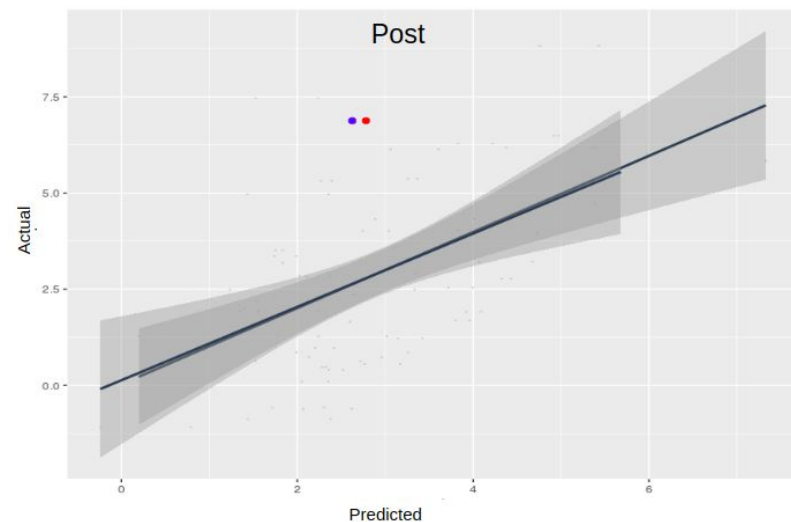
*Figure 1*



*Figure 2(a)*            *Figure 2(b)*
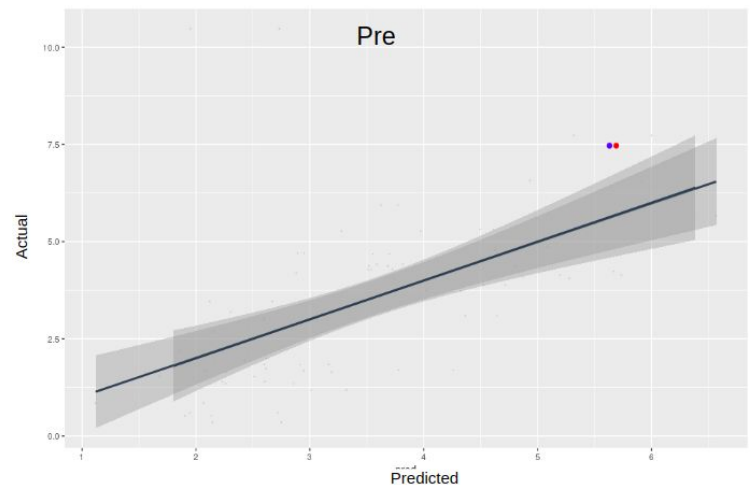
- WITHOUT TAX
- WITH TAX

## 2. Omitted Variable Bias

- Tax is a significant predictor of GDP growth, which is evident from Table no. 3 and Table no. 4, where tax variable is significant at 0.1% confidence level for both the periods.

- The tax is correlated with Domestic Credit to the private sector variable, with a correlation coefficient of 0.46, with high significance (refer table no. 6) for the Pre-period and with a correlation coefficient of 0.44, with high significance for the Post-period.

Both these observations confirm the correlation between tax and domestic credit variables. This implies that the basic OLS assumption that E(X/U)=0, is not satisfied when tax is not included. Thus excluding tax leads to an omitted variable bias in the estimation of GDP.

Moreover, on comparing R-squared values of Regressions (1) and (2) (see Table5), it is clear that including taxes in the regression helped to explain more about the variation in the GDP growth.

1. **Regression Analysis**
   a. The estimated $\beta_5$ in the regression(2) is -0.08 for the pre period and -0.10 for post period (refer Table 3 and Table 4) , significant at 0.1% confidence level, which implies that GDP growth rate and tax revenue are inversely related.
      i. If the tax revenue of a country decrease by 1(% of GDP), then its GDP Growth rate should increase by 0.1
      ii. This should not happen because, if a country has higher tax revenue, then the country's government has higher income, has more power to spend.
      iii. This result supports Dr. Subramanian's decision of not including tax, since tax are not behaving normally, as other indicators.
   b. The estimated $\beta_5$ variable of Regression (3) is 2.39 at 1% significance level, which represents that India behaved as an outlier in the post period, which was our treatment period for diff-in-diff analysis.
   c. Moreover, Fig -2(a) and Fig -2(b), which depict the plot for actual versus predicted GDP growth for the two periods, clearly show that India behaved as an outlier for the post period, irrespective of whether tax was included or not. India lies far up from the predicted-actual GDP growth line.

## Conclusion

- EAC-PM was valid to suggest the inclusion of taxes in the estimation of GDP.

- There are structural changes in the GDP in the Post-2011 period, which was more apparent after the inclusion of taxes.
- There is an Omitted variable bias in the regression.
- The author's decision of the exclusion of taxes is also verified from our study.

## References

[1]  (n.d.). India's GDP Mis-estimation - Harvard Kennedy School. Retrieved November 7, 2019, from
https://www.hks.harvard.edu/centers/cid/publications/faculty-working-papers/india-gdp-overestimate
[2] (n.d.). Validating India's GDP Growth Estimates | Harvard Kennedy .... Retrieved November 12, 2019, from
https://www.hks.harvard.edu/centers/cid/publications/faculty-working-papers/india-gdp
[3] (2014, December 8). Why the oil price is falling - The Economist explains. Retrieved November 12, 2019, from
https://www.economist.com/the-economist-explains/2014/12/08/why-the-oil-price-is-falling
[4] (2019, June 19). EAC-PM Paper on GDP Estimates in India - Economic .... Retrieved November 12, 2019, from
https://eacpm.gov.in/wp-content/uploads/2019/06/EAC-Paper_GDP-estimation_19-June-2019.pdf
[5] (2018, April 30). As Modi Claims 100% Electrification Feat, Govt Data Show 94 .... Retrieved November 12, 2019, from
https://www.news18.com/news/india/as-modi-claims-100-electrification-feat-govt-data-shows-94-villages-were-already-electrified-1734107.html
[6] (2018, September 1). India's Demonetization Drive: A Necessary Jolt Towards A .... Retrieved November 12, 2019, from
https://www.forbes.com/sites/nusbusinessschool/2018/09/01/indias-demonetization-drive-a-necessary-jolt-towards-a-more-digital-economy/
[7] (n.d.). Domestic credit to private sector (% of GDP) by ... - IndexMundi. Retrieved November 7, 2019, from
https://www.indexmundi.com/facts/indicators/FS.AST.PRVT.GD.ZS
[8] (n.d.). Exports of goods and services (annual % growth ... - IndexMundi. Retrieved November 7, 2019, from
https://www.indexmundi.com/facts/indicators/NE.EXP.GNFS.KD.ZG
[9] (n.d.). Imports of goods and services (annual % growth ... - IndexMundi. Retrieved November 7, 2019, from
https://www.indexmundi.com/facts/indicators/NE.IMP.GNFS.KD.ZG
[10]  (n.d.). Tax revenue (% of GDP) by Country - IndexMundi. Retrieved November 7, 2019, from
https://www.indexmundi.com/facts/indicators/GC.TAX.TOTL.GD.ZS

# Tables

| 2001-2011 Without Tax | | | | |
|---|---|---|---|---|
| Variables | Estimate | Std. Error | t-value | Pr(>|t|) |

| | | | | |
|---|---|---|---|---|
| Intercept | 3.26 | 0.35 | 9.41 | 2e-16*** |
| India_Dummy | 1.07 | 1.20 | 0.89 | 0.000549 *** |
| Domestic credit to the private sector | -0.01 | 0.003 | -3.48 | 0.000549 *** |
| Exports of goods and services | 0.06 | 0.02 | 2.68 | 0.007527 ** |
| Imports of goods and services | 0.12 | 0.02 | 6.49 | 1.99e-10 *** |

Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.892 on 534 degrees of freedom
Multiple R-squared:  0.22,     Adjusted R-squared:  0.22
F-statistic: 38.24 on 4 and 534 DF,  p-value: < 2.2e-16

Table - 1

| 2011-2015 Without Tax | | | | |
|---|---|---|---|---|
| Variables | Estimate | Std. Error | t-value | Pr(>\|t\|) |
| Intercept | 3.22 | 0.37 | 8.661 | 1.95e-15 *** |
| India_Dummy | 4.35 | 1.32 | 3.29 | 0.001 ** |
| Domestic credit to the private sector | -0.01 | 0.003 | -3.58 | 0.0004 *** |
| Exports of goods and services | 0.08 | 0.03 | 2.42 | 0.02 * |
| Imports of goods and services | 0.18 | 0.03 | 5.99 | 1.03e-08 *** |

Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.845 on 533 degrees of freedom
Multiple R-squared:  0.2428,  Adjusted R-squared:  0.2357
F-statistic: 34.19 on 5 and 533 DF,  p-value: < 2.2e-16

Table - 2

| 2001-2011 With Tax | | | | |
|---|---|---|---|---|
| Variables | Estimate | Std. Error | t-value | Pr(>|t|) |
| Intercept | 3.94 | 0.39 | 10.18 | 2e-16 *** |
| India_Dummy | 1.01 | 1.19 | 0.85 | 0.40 |
| Domestic credit to the private sector | -0.01 | 0.003 | -1.86 | 0.06 |
| Exports of goods and services | 0.05 | 0.02 | 2.38 | 0.018 * |
| Imports of goods and services | 0.12 | 0.02 | 6.70 | 5.4e-11 *** |
| Tax revenue | -0.08 | 0.02 | -3.77 | 0.000183 *** |

Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.845 on 533 degrees of freedom
Multiple R-squared:  0.24,     Adjusted R-squared:  0.24
F-statistic: 34.19 on 5 and 533 DF,  p-value: < 2.2e-16

Table - 3

| 2011-2015 With Tax | | | | |
|---|---|---|---|---|
| Variables | Estimate | Std. Error | t-value | Pr(>|t|) |
| Intercept | 4.24 | 0.42 | 10.02 | <2e-16 *** |
| India_Dummy | 4.19 | 1.26 | 3.32 | 0.03 * |
| Domestic credit to the private sector | -0.01 | 0.003 | -2.23 | 0.03 * |
| Exports of goods and services | 0.07 | 0.03 | 2.26 | 0.02 * |
| Imports of goods and services | 0.20 | 0.03 | 6.81 | 1.26e-10 *** |

| Tax revenue | -0.10 | 0.02 | -4.43 | 1.58e-05 *** |

Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.474 on 190 degrees of freedom
Multiple R-squared:  0.3883,  Adjusted R-squared:  0.3722
F-statistic: 24.12 on 5 and 190 DF,  p-value: < 2.2e-16

Table - 4

| R-Sq Comparison | | |
| --- | --- | --- |
| Time Period | Without Tax R-sq (Adjusted) | With tax R-sq (Adjusted) |
| 2002 - 2001 | 0.21 | 0.23 |
| 2011 - 2015 | 0.31 | 0.37 |

Table - 5

| Time Period | Correlation Value | GDP Growth | Domestic credit to private sector (% of GDP) | Exports of goods and services (annual % growth) | Imports of goods and services (annual % growth) |
| --- | --- | --- | --- | --- | --- |
| Pre (2001 -11) | Correlation coefficient | -0.24 | 0.46 | -0.15 | -0.11 |
| | P Value | 5.124e-08 | < 2.2e-16 | 0.0009 | 0.01 |
| Post (2011 -15) | Correlation coefficient | -0.36 | 0.44 | -0.04 | 0.04 |
| | P Value | 1.054e-06 | 1.075e-09 | 0.57 | 0.58 |

Table 6

| Regression 3 | | | | |
|---|---|---|---|---|
| Variables | Estimate | Std. Error | t-value | Pr(>|t|) |
| Intercept | 3.92 | 0.36 | 11.04 | <2e-16 *** |
| Domestic credit to the private sector | -0.006 | 0.003 | -2 | 0.45 * |
| Exports of goods and services | 0.05 | 0.02 | 2.58 | 0.01 * |
| Imports of goods and services. | 0.12 | 0.02 | 7.30 | 7.60e-13 *** |
| Tax.revenue....of.GDP. | -0.08 | 0.03 | -4.08 | 4.86e-05 *** |
| India_Dummy | 1.86 | 1.09 | 1.70 | 0.08 . |
| T | 0.30 | 0.70 | 0.43 | 0.66 |
| India_Dummy:T | 2.38 | 2.10 | 1.13 | 0.25 ** |
| Domestic credit to the private sector.:T | -0.002 | 0.006 | -0.23 | 0.81 |
| Exports of goods and services:T | 0.02 | 0.05 | 0.36 | 0.71 |
| Imports of goods and services:T | 0.07 | 0.04 | 1.63 | 0.10 |
| Tax revenue:T | -0.02 | 0.04 | -0.65 | 0.52 |
| Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 | | | | |
| Residual standard error: 2.474 on 190 degrees of freedom<br>Multiple R-squared:  0.3883,  Adjusted R-squared:  0.3722<br>F-statistic: 24.12 on 5 and 190 DF,  p-value: < 2.2e-16 | | | | |