The Effect of Air Pollution on India's GDP Growth Rate

Karthik Sunkesula
The College of New Jersey
March 26, 2016
Dr. Michele Naples

Abstract: India is considered the largest democracy in the world. Despite the associated organizational issues, India's GDP growth has risen over the last 50 years due to trade liberalization, human capital growth, and foreign direct investment. However, in light of the concern over global warming, it is necessary to determine the economic consequences of an increase in the accompanying CO_2 emissions. Recent literature (Holtz-Eakin and Selden) on air pollution and economic development suggests diminishing marginal propensity to emit (MPE) harmful gases in the short run, while predicting the opposite in the long run. This paper will help provide insight into the relationship between air pollution and India's GDP growth rate using empirical analysis.

Table of Contents

I. Introduction

In recent years, global warming has become one of the most pressing areas of concern in society. Increases in climate change will cause more negative consequences for the whole of humanity; natural effects could range from coastal erosion to droughts, and far worse. In addition to the adverse natural effects of global warming, there can be many economic costs that follow.

As the world's fourth-biggest carbon emitter, India has much of the burden on its shoulders. In addition to the ambitious goal of deriving more energy from renewable resources, the Indian government needs to enact more stringent legislation and rule of law that ensures compliance with environmental regulation. India has recently developed a National Action Plan on Climate Change to strategize its role in cleaning the environment. These efforts range from reforestation to enhanced solar power initiatives, and will include a "series of mandated efficiency standards for vehicles, buildings, and appliances" (WorldWatch). However, it is not enough for only India to take upon their fair share of responsibility to addressing climate change; all countries need to maximize their commitment to lowering their associated per-capita emissions, in order to help save the environment and global economy.

The paper commences with the motivation behind the research question and topic, and will then flow into a review of the various works surrounding the matter at hand. Next, I delve into the different strategies that were employed to create each version, after which, I cover the data used in the study. To conclude the paper, I present and analyze the econometric results, and interpret them in order to determine how India can tackle the problem of global warming.

II. Research Question and Motivation

In light of the increased focus on global warming and air pollution, I will be researching the effects of carbon dioxide (CO_2) emissions, among other macroeconomic variables, on India's Gross Domestic Product (GDP) growth rate. There has been much debate about the sustainability of growth for developing countries; unfortunately, there has not been conclusive research into the relation between air pollution and the economic growth of a developing country. Using supply/demand side and PPF independent variables, I wrestle with multiple versions to measure the effects of those selected variables, and especially CO_2 emissions, on India's GDP growth rate.

This research topic aims to reaffirm the known effects of various macroeconomic variables on economic growth (measured via GDP growth rate), and uncover the magnitude of effect air pollution has on India's GDP growth rate. This research paper looks to answer the question of *whether increases in air pollution diminishes economic growth in developing countries* and strives to develop this insight with a macroeconomic perspective.

If, given an irrefutable relation between GDP growth rate and air pollution, I hope to expand upon Holtz-Eakin and Selden's work and bring some understanding into the economic effects of air pollution in the short and long run. Holtz-Eakin and Selden determined four things via their study: 1) as economies develop, there seems to be a diminishing marginal propensity to emit (MPE) greenhouses gases (CO_2), 2) despite the diminishing MPE, global emissions of CO_2 will continue to grow at close to 2% annually, 3) the growth in greenhouse gas emissions stems from economic and population growth in lower-income nations with the highest MPEs, and 4) the rate at which global economies are developing does not drastically change the future annual/cumulative flow of CO_2 emissions. This paper utilizes the studies cited, and pairs it with the results attained through this paper, to illustrate the need for anti-pollution legislation in developing countries.

III. Literature Review

There has been much research done addressing the economic nature and growth of developing countries. Moreover, there are a few studies that concentrate on the effect of pollution on economic growth. With the information attained from both types of research, one can put the pieces of the puzzle together to discern the hypothesized conclusion that relates to my thesis question.

Ahluwalia (1995) looks at the different reforms in the financial sector that have affected India's economic growth. Her paper describes some of these reforms and includes that "the interest rates on government securities are increasingly market determined… and that deposit rates for different maturities are only subject to a single ceiling" (Ahluwalia 1995: p. 8). With these reforms, among many others, the Indian Government looks to focus on "establishing a framework of regulation to ensure transparency of trading practices, speedy settlement procedures, enforcement of prudential norms, and full disclose for investor protection" (Ahluwalia 1995: p. 9). With these efforts taken, it will be interesting to see how these policies have impacted the interest rate and investment throughout the years. Similarly, I can determine the effects of these changes in interest rates and investment on India's GDP growth rate.

Asafu-Adjaye (1979) looks to approximate the pivotal relationships between energy consumption and income for developing countries. His paper uses cointegration and error-correction Versioning techniques. With respect to India, Asafu-Adjaye concludes that energy consumption and income are not neutral with one another in the long run; however, they seem to be neutrally related in the short-run. To compare, he cites Kraft and Kraft (1979) in which there was evidence found for causality between GNP and energy consumption in the United States. Unfortunately, all those who have ventured to find similar correlations have not found these results, and instead have discovered no relation between energy consumption and variables that derive themselves from income.

Borensztein, De Gregorio, and Lee (1998) aimed to "examine empirically the role of foreign direct investment (FDI) in the process of technology diffusion and economic growth in developing countries" (Borensztein, De Gregorio, Lee 1998: p. 116). They found that FDI is important only when there is enough advanced technology in the host economy to accommodate the increased productive value that the FDI brings into the economy. Thus, "in a typical Version of technology diffusion, the rate of economic growth of a backward country depends on the extent of adoption and implementation of new technologies that are already in use in leading countries" (Borensztein, De Gregorio, Lee 1998: p. 116). This realization is quite essential and relevant to India's growth; in the past year, India has been in discussion with other countries to build high-tech developments for India's development efforts. Countries such as Japan and the United States have been involved in helping develop high-speed railway systems and aircraft carrier fleets. Moreover, this paper uses "a Version of endogenous growth, in which the rate of technological progress is the main determinant of the long-term growth rate of income" (Borensztein, De Gregorio, Lee 1998: p. 116). Borensztein indicates that rate of tech progress was used as the main component to determine the rate of economic growth in developing countries. The rate of human capital/tech progress is just as important to the growing FDI, as the latter is dependent on the former to truly be productive to a growing economy's growth. In addition to the FDI, the Version uses stock of human capital and initial GDP capital to determine the extent of the casual relationship between FDI and human capital stock. As iterated in the paper, "there needs to be a certain level of absorptive capability for advanced technologies to have an impact on the growth rate of income" (Borensztein, De Gregorio, Lee 1998: p. 117). This relationship helps to identify the importance of the relationship between capital stock and FDI in a developing country, and solidifies our use of capital stock in our Version. Similar to the rate of technology growth, the capital stock of a developing country is of utmost importance in relation to the FDI, as the latter is dependent on the former to truly be productive in its use of expanding a developing economy. Together with an accelerating rate of technology, human capital can be of great use to FDI in its mission to expand a country's economic growth.

Bosworth and Collins (2007) compared the economic performances of India and China by using "estimations of the contribution of labor, capital, education and total factor productivity" (Bosworth, Collins 2007: Abstract). From the data collected, it is shown that China's output growth was double India's between 1993 and 2004. China's growth was due to the boost in its industrial sector, whereas India's growth stemmed from its services industry. However, both countries gained in productivity growth when reallocating labor from agriculture to industry and services. Their research points to the fact that labor efforts in industry and services are much more economically advantageous than concentrated labor in agriculture.

Holtz-Eakin and Selden (1995) "used panel data to estimate the reduced form relationship between per-capita income and emission, and forecasted aggregate emissions and their distribution among countries" (Holtz-Eakin, Selden 1995: p. 86). Given the increase in global warming concerns, Holtz-Eakin and Selden uncovered four results: as economies develop, there will be a diminishing marginal propensity to emit (MPE) CO_2; despite the diminishing MPE, global carbon dioxide emissions will continue to grow at a rate of close to 2% annually; middle-to-lower income nations with high MPEs will be the cause of continued growth; pace of economic growth will not drastically alter current or future cumulative CO_2 emissions. Using these conclusions on developed countries, I try to gather CO_2 data for emerging markets to determine any similarities or differences in the relation between $C0_2$ emissions and economic development.

K.L Krishna (2004) analyzes the various determinants of economic growth and instabilities that distinguish developing countries. Krishna points to urbanization, manufacturing sector, and literacy rates as the main engines for economic growth. "In regards to the 14 major states that made up India during the period 1960-2000, the interstate variations in income and growth increased over time; however, the relative positions of the states remained quite constant over time" (Krishna 2004: p. 1-3). Despite the lack of conclusive evidence, the studies suggest a negative relation between long-term growth rates and income levels, implying conditional divergence in the long run.

Krishna and Mitra (1998) look to investigate the "effects on competition and productivity of the dramatic 1991 trade liberalization in India" (Krishna, Mitra 1998: p. 448). Through their work, it is found that trade liberalization resulted in increased competition and welfare, contributing to a spurt in growth rate of productivity" (Krishna, Mitra 1998: Abstract). However, it seems that trade policy could "result in both increased and negative growth rates" (Krishna, Mitra 1998: p. 448). They allow for a change in their returns to scale after liberalization, making their work differ from their predecessors'. With the trade liberalization that went on in 1991, there came significant reductions in the price-marginal cost markups in the post 1991 era, and some weaker evidence of an increase in the growth rate of productivity" (Krishna, Mitra 1998: p. 449). While trade liberalization has had its good and bad effects on growth rates, there seems to be general agreement that it positively affects developing economies' growth rates. Productivity seems to have increased in the past, and this can be a sign that an increase in exports can stimulate greater productivity and growth within a developing economy. Trade policy does not necessarily have to target both import and export legislation, but could rather just narrow in on the level of exports targeted in the future. This paper bolsters our use of the export growth rate in the Version to determine the significance it has on the India's GDP growth rate.

Kunal Sen (2007) strives to illustrate the reasons for India's growth acceleration in the late 1970s. Despite the widespread attitude that the growth acceleration was due to the inception of revolutionary economic policies, the study points to other factors that seem to be more plausible causes of the growth spurt. Sen asserts that the "increase in growth was a result of private equipment investment, public fixed investment, while noting that the shift in focus to the private sector was a result of the prior increases in investment vehicles, rather than an independent factor to India's rapid development" (Sen 2007: Abstract).

Levin and Raut (1997) examine the "role of exports and human capital in the determination of long-run economic growth" (Levin, Raut 1997: p. 155). Despite the known effect of international trade on GDP growth, the channel through which the growth is accelerated is unclear. Moreover, it has been determined that enrollments/ measures of human capital do not have a huge influence on GDP growth in developing countries. To address these issues, Levin and Raut "use a sample of 10-year GDP growth rates for 30 semi-industrialized developing nations over time" to determine the roles exports and human capital plays in GDP growth (Levin, Raut 1997: p. 170). It seems that the results they found were inconclusive due to geographic and time differences. For example, "India may expend significant resources on educational investment, it might reap insignificant benefits due to the small size of its manufactured export sector" (Levin, Raut 1997: p. 170). This might hint at a causal relationship between exports and human capital, but more research will have to be conducted to solidify this possibility.

Nidugala (2009) looks to examine thoroughly the role of exports in India’s economic growth by analyzing the relationship between the growth in exports and the growth in India’s economy in the 1980s, using data from the Central Statistical Organization (CSO) Planning Commission and DGCIS (Nidugala 2009: p. 6). Through this research, it has been found that the growth of manufactured exports had a significant influence on GDP growth during the 1980s with exceptions in 1961-1962, and 1979-1980. Moreover, Nidugala determined that the increase in correlation between export growth and GDP growth is due to stronger inter-industry linkages and a shift towards a concentration of manufactured exports because of their price responsiveness. In conclusion, Nidugala assumes that an export-led growth strategy, which revolves around manufacturing goods, can be used to stimulate the Indian economy.

Posner (1998) emphasizes the benefits of legal reform in an economy, as “modest expenditures on law reform increase the rate of economic growth, in turn generating resources that will enable more ambitious legal reforms to be undertaken in the future” (Posner 1998: p. 1). This paper lets us remember that there is more to economic growth than the macroeconomic variables that make up the Aggregate Expenditure equation. For a country to realize the benefits of economic policy, legal reforms must be passed to facilitate economic progress. For a developing country like India, this is especially relevant and essential as India faces corruption and legal issues.

IV. Theoretical Version Development

To start the research, I first needed to compile the data required to develop a conclusion. Using a macroeconomic approach for the issue at hand, I added environmental data to a standard GDP equation. In essence, I included the factors that affect pollution into the factors that affect GDP growth.

1. GDP growth rate = f (L_g, K_g, $Pollution_g$)
2. $Pollution_g$ = g (GDPg, CARg, DTHg, RAINg, TEMPg, POPg, CARg)
3. GDP growth rate = h (L, K, EMITg, CARg, DTHg, RAINg, TEMPg, POPg)

The L and K in the original GDP growth equation (1) stand for labor and capital. All three equations are functions, and indicate what the independent variable is affected by; the third equation is a reduced-form. I measured economic development via GDP Annual Growth Rate **(GDPg)**, and used it as my dependent variable to determine the relationships among GDPg and the independent variables. My independent variables are Export Annual Rate of Growth **(EXg)**, CO_2 Annual Rate of Growth **(EMITg)**, Net Official Development Assistance Annual Rate of Growth **(AIDg)**, Capital Stock Annual Rate of Growth (**CAPg**), Investment Annual Rate of Growth (**INVg),** Enrollment in Primary Education Annual Rate of Growth **(EDUg**), Labor Force Annual Rate of Growth **(LABg)**, Death Rate Annual Rate of Growth **(DTHg)**, Rainfall Annual Rate of Growth **(RAINg),** Population Density Annual Rate of Growth **(POPg)**, Temperature Annual Rate of Growth **(TEMPg)** and Total Vehicles Annual Rate of Growth **(CARg**).

The variables can be grouped into three categories:

- **Environmental Variables →** CO_2 emissions (EMIT), <u>death rates</u> (DTH), <u>annual rain levels</u> (RAIN), <u>car growth</u> (CAR) and average <u>temperature</u> (TEMP)
- **Other Supply-side Variables →** Enrollment in primary education **(**EDU), Labor force (LAB) and <u>population density</u> (POP)
- **Demand-side Variables →** Investment (INV), Net official development assistance (AID), and exports (EX)

Unfortunately, while it is argued that growing pollution hurts GDP growth, it is recognized that GDP growth may contribute to pollution growth. This may bias coefficient-estimates of eq. 3 above. Consequently it was decided to explore a simultaneous-equation model, as well, to estimate eqs. 1 and 2 directly. Those variables, which are underlined, are variables that will act as instrumental variables during the 2SLS regression analysis.

V. Data

Taking into account the macroeconomic perspective of the issue at hand, I chose both supply-side and demand-side variables to develop data sets that were to be used in the empirical analysis. All of the data was collected via The World Bank, Open Government Data (OGD) Platform India, Open Data for Africa (ODA), or the Federal Reserve Economic Data (FRED) database.

- **The World Bank Datasets**

The World Bank provided the annual growth rates of India's GDP and the annual level-series data for India's Exports of Goods and Services (In current U.S. dollars) from 1961 – 2014. The World Bank website also provided annual level-series data on CO_2 emissions (In metric tons per capita) from 1961 – 2011 and Net Official Development Assistance and Official Aid Received (In current U.S. dollars) from 1961 – 2013. The World Bank website also provided annual level series data for Labor Force from 1990 – 2014, Enrollment in Primary Education (Both sexes) from 1971 – 2003 and 2007 – 2012, Death Rates (per 1000) from 1961 – 2013, and Population Density (People per square kilometer of land area) from 1961 – 2014.

- **Open Government Data (OGD) Platform India Datasets**

The OGD Platform India Data provided annual level series data on the Total Number of Vehicles from 1961 – 2012, Annual Temperatures from 1961 – 2014, and Annual Weighted Rainfall (In millimeters) from 1961 – 2011.

- **Open Data for Africa (ODA) Datasets**

The ODA provided annual level series data on Total Annual Investment (% of GDP) from 1980 – 2014. Next, I calculated the Total Investment (In current U.S. Dollars) by multiplying the Total Annual Investment (% of GDP) by the associated GDP (In current U.S. dollars) for each year. This data set was then transcribed to growth rates by using the Growth formula I have previously discussed.

- **Federal Reserve Economic Data (FRED)**

The FRED database provided the annual level series data for GDP (In current U.S. dollars) from 1961 – 2013, and was used to calculate the Total Investment (In current U.S. dollars). FRED also provided the annual level series data of Capital Stock at Constant National Prices (in millions of U.S. dollars) from 1961 – 2011.

If growth rates were not already given, all annual level series data were used to create annual growth rates with the use of the Growth Formula:

$$\frac{\text{(New data point - Old data point)}}{\text{Old data point}}$$

I repeated this derivation for every variable for the longest time period possible with the associated time periods of data accessible.

There are several limitations on the data acquired through the web. There were many datasets on the World Bank that were not thoroughly captured over the years; in essence, there were many gaps, or there were strings of years, through which data had not been calculated or collected. This lack of data choice constricted our range of selection, thus impeding the ability to determine the best independent variables to explain the relation between economic development and air pollution.

From graphing GDPg (dependent variable) and EMITg (key causal independent variable), we can see that there is some correlation between the two variables.

[Insert Graph 1]

VI. Empirical Version Development and Specification

After gathering all the data, I ventured to come up with three versions, upon which I would derive my concluding results. In accordance with the regression results, I deduced that the key variables which were only available for a subset of years were INVg, LABg and EDUg. Therefore, I proposed the following three versions:

1. Without INVg, LABg and EDUg
2. With INVg and LABg (without EDUg)
3. With INVg, LABg and EDUg

Version 1 (Without INVg, LABg and EDUg) entails the years 1969 – 2011. Version 2 (With INVg and LABg [without EDUg]) involves the years 1981 – 2011. Version 3 (With INVg, LABg and EDUg) encompasses the years 1981 – 2003. Studying all three versions would provide a comprehensive study of the key causal independent variables and their respective impacts on the economic development (GDPg) of India.

After building each version based on the variables' growth rates, and running analyses on each equation, I chose to advance my research by using the 2SLS method to take a closer look at the possibility of endogeneity (two-way causation) within the versions. With the 2SLS method, I would be able to see the relations among variables explicitly. Endogeneity occurs when the dependent variable's error terms are correlated with the independent variables. Thus, to account for the possibility of endogeneity, I added DTHg, RAINg, TEMPg, POPg, and CARg as instrumental variables that would be related to EMITg, but not significantly to GDPg.

Version 1 (n = 43) →

GDPg = $\beta_0 + \beta_1$ (EXg) - β_2 (EMITg) + β_3 (AIDg) + β_4 (CAPg) - β_5 (DTHg) + β_6 (RAINg) +β_7 (TEMPg) + β_8 (POPg) + β_9 (CARg) + u

Version 2 (n = 31) →

GDPg = $\beta_0 + \beta_1$ (EXg) - β_2 (EMITg) + β_3 (AIDg) + β_4 (CAPg) + β_5 (INVg) + β_6 (LABg) - β_7 (DTHg) + u

Version 3 (n = 23) →

GDPg = $\beta_0 + \beta_1$ (EXg) - β_2 (EMITg) + β_3 (AIDg) + β_4 (CAPg) + β_5 (INVg) + β_6 (EDUg) + β_7 (LABg) - β_8 (DTHg) + β_9 (RAINg) + u

The assumed signs on each of the independent variables are: (+) for EXg, AIDg, CAPg, RAINg, POPg, CARg, LABg, INVg, and EDUg. The assumed signs on EMITg, TEMPg and DTHg are (-). I included DTHg, RAINg, TEMPg, POPg and CARg since each of these added independent variables has a specified relationship with EMITg, and GDPg. It was assumed that TEMPg, DTHg, RAINg, POPg and CARg had a stronger relationship with EMITg than with GDPg, while having an inverse, inverse, positive, positive, and positive relationship to GDPg, respectively. This is due to the proven natural or macroeconomic relation each variable has in relation to economic growth. In addition, all betas are assumed to be (+).

As each equation was regressed and analyzed, I found three variables to be highly correlated: CAPg, EXg, and INVg.

[Insert Table 2: Correlation Matrices]

In version 1, 2 and 3, CAPg and EXg had a correlation of 0.35, 0.47, and 0.52. Moreover, CAPg had high correlations with INVg in versions 2 and 3 with correlations of 0.47, and 0.54. Lastly, EXg and INVg were collinear with correlations of 0.65 and 0.66 in versions 2 and 3, respectively.

As a result of the collinearity among these variables, I ran GDP as a function of the explanatory variables, while excluding the highly correlated independent variables (CAPg, INVg and EXg).

[Insert Table 4: Regressions without collinear variables]

Unfortunately, no significance was found with either of the versions. And so, I went back to including CAPg, while excluding INVg and EXg, as it was the most useful prior to the exclusion of the collinear variables.

Thereafter, when using the 2SLS method, I used various instrumental variables in each version. Instrumental variables can be used to address the problem of omitted-variable bias or solve the classic errors-in-variables problem. For a variable to be instrumental in these versions, it must be:

1. Exogenous
2. Correlated with the endogenous variable EMITg, and uncorrelated with GDPg

In version 1, I instrumented EMITg against DTHg, RAINg, TEMPg, POPg, and CARg. Similarly, in version 2, I used DTHg as the only instrumental variable, whereas in version 3, I used DTHg and RAINg as the only instrumental variables

[Insert Table 4: 2SLS Outputs]

VII. Econometric Results and Interpretations

After running the regressions on each version, I noticed that the R^2 increased as n (number of observations) increased from .304 in version 1 to .91 in version 3. The R^2 denotes the explanatory power of the version, illustrating how well the data fits with the regression line. Although the independent variables in each version differed, this pattern needed supplementary examination. And so, I looked to the parameter coefficients (PEs) to determine which variables were significant, and how each version fared in regards to how well the equations worked.

Unfortunately, CAPg and INVg were the only two variables that seemed to have any significance within all three versions. Unhappy with the lack of significant variables, I ventured to determine the possibility of collinearity within the versions. This called for further examination of the correlation matrices; I found that three variables were correlated with each other: EXg, CAPg, and INVg. In version 1, CAPg and EXg are correlated at 0.35. In version 2, CAPg and EXg are correlated at 0.47, whereas, INVg and CAPg are correlated at 0.65. Lastly, in version 3, CAPg and EXg are correlated at 0.51, whereas INVg and EXg are correlated at 0.66. Given these high correlations and the possibility of endogeneity in each version, I adjusted each version by extracting the collinear variables (EXg, CAPg and INVg), and re-ran the regressions.

After re-running the regression for Version 1, I found that none of the variables were significant at either the 1%, 5% or 10% levels, hinting at no change compared to the first regression run of version 1. Moreover, the sign of the parameter estimate for EMITg was positive for the re-runs of version 3 (significant at the 15% level) , and thus was not in accordance with the original hypothesis. When I ran the regressions again on version 2 and 3, none of variables were significant at either the 1%, 5% or 10% levels. There was also a decrease in both R^2 and Adjusted R^2 across the board in the second regression run.

[Insert Table 3: Old and New Regression Statistics]

However, given that increases in air pollution signify increases in business, this would, therefore, indicate that the increase in GDPg hints at increased business in India, and as a result, contributes to the increase in air pollution. To account for this issue of two-way causation, I included instrumental variables such as DTHg, RAINg, TEMPg, POPg, and CARg because these variables had a larger correlation with air pollution than with GDPg. I performed 2SLS tests for each version, while accounting for the highly correlated variables: CAPg, EXg, and INVg, by running 2SLS tests where only oee of the highly collinear variables were included, with the other two left out. And so, I found that CAPg is significant at the 1% level for version 1 (CAPg)[1], and at the 5% level for version 2 (CAPg). Unfortunately, there were no significant variables for version 3 when I ran the 2SLS. The F-statistics were significant at the 1% level for version 1 (CAPg) and version 3 (CAPg), and at the 10% level for version 2 (CAPg). In regards to the versions with only EXg (leaving out the other two highly correlated variables: CAPg and INVg), I discovered that none of the independent variables or F statistics was meaningful at any significant level for any of the versions. This signified that EXg was not at essential to the equation as I had previously imagined it to be. Fortunately, I found that the 2SLS runs with only INVg yield some significant results: version 2 (INVg) generated INVg and an F-statistic that were significant at the 1% level, and a constant that was significant at the 5% level. Moreover, in version 3 (INVg), INVg and the F-statistic were found to be significant at the 5% and 10% levels, respectively. Since INVg and the F-statistics were significant at some levels, the results shows that INVg can be relied upon in future studies. Unfortunately, data limitations did not allow for a version 1 (INVg); consequently, I did not have the necessary data to compare version 1 (CAPg) and version 1 (INVg). Given this dilemma, I chose to move forward with the 2SLS outputs of the versions 1, 2 and 3 with CAPg as their included correlated variable.

[Insert: Table 4: 2SLS Results]

To compare with these 2SLS outputs of versions 1, 2 and 3 with CAPg as their main correlated variable, I ran adjusted OLS regressions to compare how meaningful each variable was

---

[1] Version 1 (CAPg) signifies version 1 with only one of the highly correlated variables: CAPg (excludes the other two highly correlated variables: EXg and INVg), in its equation for the 2SLS runs. In essence, if a highly correlated variable is in parentheses, it indicates that specific variable is included in the equation for the 2SLS run, while the other two highly correlated variables are not.

in each version. The adjusted OLS regressions included only EMITg, AIDg, and CAPg, while being in accordance with each version's data limitations. The adjusted regression output showed that CAPg was significant at the 1% level for both version 1 and version 2. However, despite the comparison with the 2SLS outputs, I was unable to find additional significant explanatory variables. Consequently, the macroeconomic method used failed to establish the relationship between economic development and air pollution.

[Insert: Table 5: 2SLS (CAPg) Comparison with Adjusted OLS Regression]

VIII. Conclusions and Suggestions Moving Forward

Despite the rigorous data collection and analysis methods used, the study was unable to confirm the relation between GDPg and EMITg that was hypothesized. As a result, I was unable to either accept the hypothesis that EMITg and GDPg have an inverse relationship.

The problems faced can be due to multiple sources: the inaccuracy of data, lack of data for variables needed, and lack of data for variables used. While the data was provided by verifiable sources, data collection in India can be very misleading: data can be erroneous, misleading or misinterpreted, thereby resulting in inaccurate data. If the data is inaccurate, this can result in misleading output, and conclusions. Moreover even though there were data for the variables used in the study, the data was not comprehensive and did not cover every single year. More often than not, there were many gaps between years, making the level-series data unreliable. Lastly, while there were many variables available, there were also many variables that would have served as better estimates of the influences I attributed to the equation. With a bigger database of data available, one can utilize the needed variables to come up with more innovative versions for testing.

Moving forward, I would use a microeconomic outlook, rather than a macroeconomic method. I would test the differences in GDPg between two similar neighboring states, accounting for the differences in EMITg between the two states. This implies the necessary condition that the two states, while being similar in economic advancement, must be polar opposites (or as close to) in regards to air pollution emissions. Perhaps, this would yield better results as I can focus in on the issue, and disregard external variables that can be rendered irrelevant.

Despite the lack of conclusive results, I have laid out the general macroeconomic version one can use when developing an equation to explain the relationship between GDPg and EMITg. In addition, this study has led to further brainstorming, and helped determine which variables work and which do not, in accordance with any limiting factors arising from the data sources.

IX. Bibliography

1. Ahluwalia, Montek S. "India's economic reforms." *India: The Future of Economic Reform, Oxford University Press: Oxford, United Kingdom* (1995).

2. Asafu-Adjaye, John. "The relationship between energy consumption, energy prices and economic growth: time series evidence from Asian developing countries." *Energy economics* 22.6 (2000): 615-625.Kira, Alex Rueben. "The Factors Affecting Gross Domestic Product (GDP) in Developing Countries: The Case of Tanzania." European Journal of Business and Management 5.4 (2013): 1-11. IISTE. Web. 10 Nov. 2015.

3. Borensztein, Eduardo, Jose De Gregorio, and Jong-Wha Lee. "How does foreign direct investment affect economic growth?" *Journal of international Economics* 45.1 (1998): 115-135.

4. Bosworth, Barry, and Susan M. Collins. *Accounting for growth: comparing China and India*. No. w12943. National Bureau of Economic Research, 2007.

5. Holtz-Eakin, Douglas, and Thomas M. Selden. "Stoking the Fires? CO2 Emissions and Economic Growth." *Stoking the Fires? CO2 Emissions and Economic Growth*. Journal of Public Economics, 1995. Web. 11 Apr. 2016.

6. Krishna, K. L. "Patterns and determinants of economic growth in Indian states." (2004).

7. Krishna, Pravin, and Devashish Mitra. "Trade liberalization, market discipline and productivity growth: new evidence from India." *Journal of Development Economics* 56.2 (1998): 447-462.

8. Levin, Andrew, and Lakshmi K. Raut. "Complementarities between exports and human capital in economic growth: Evidence from the semi-industrialized countries." *Economic development and cultural change* 46.1 (1997): 155-174.

9. Nidugala, Ganesh Kumar. "Exports and economic growth in India: an empirical investigation." *Indian Economic Journal* 47.3 (2000): 67.

10. Posner, Richard A. "Creating a legal framework for economic development." *The World Bank Research Observer* 13.1 (1998): 1-11.

11. Sen, Kunal. "Why did the elephant start to trot? India's growth acceleration re-examined." *Economic and Political Weekly* (2007): 37-47.

X. Appendix A

Graph 1: GDPg v. EMITg

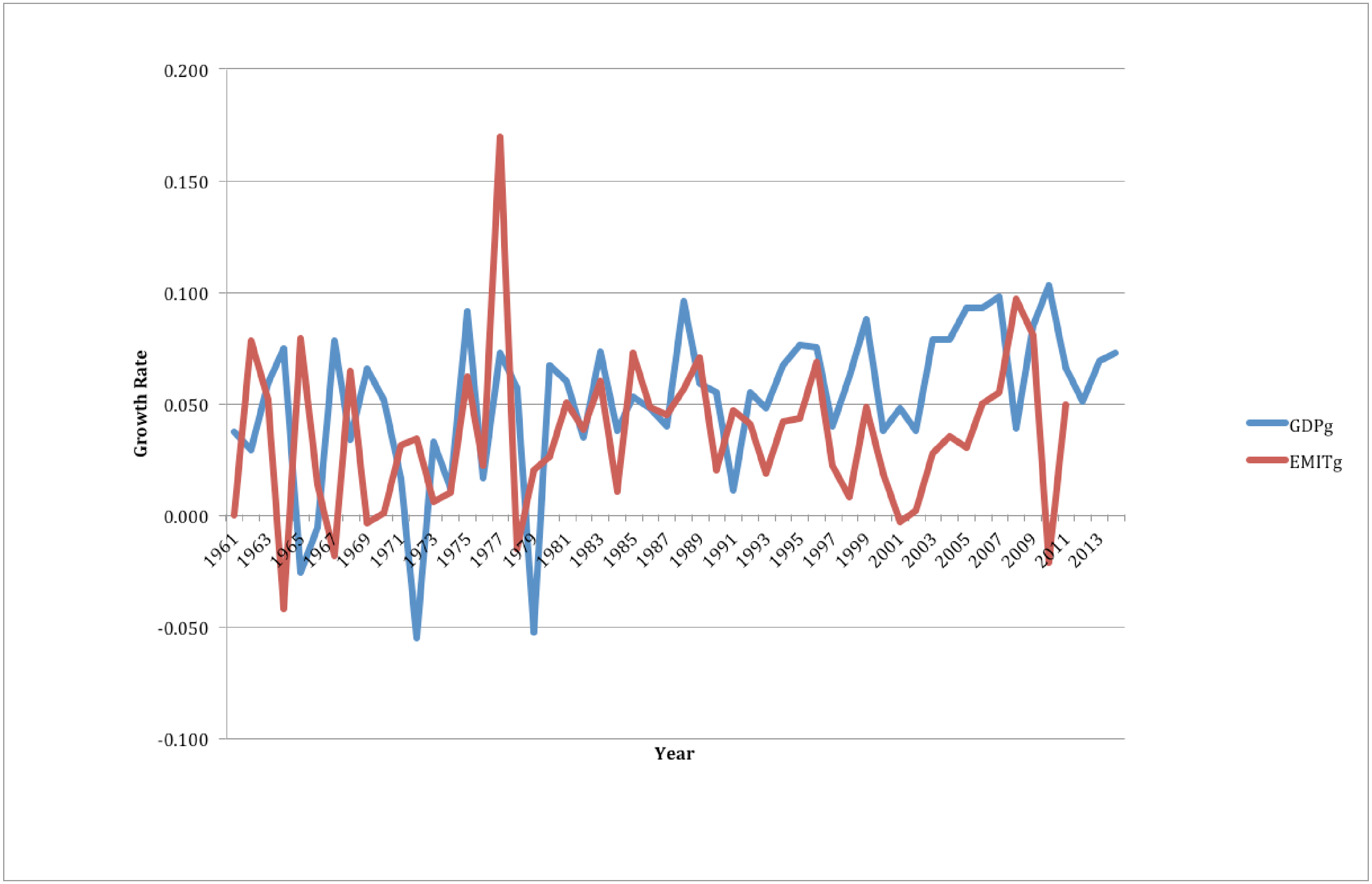

Table 1: Summary Statistics

Variable	Definition	Mean	Standard Deviation	Minimum	Maximum	Source
GDPg	Annual Growth Rate of GDP	0.052	0.034	-0.055	0.103	World Bank
EXg	Annual Growth Rate of Exports of Goods and Services	0.118	0.109	-0.080	0.39	World Bank
EMITg	Annual Growth Rate of CO2 Emissions	0.037	0.035	-0.042	0.169661	World Bank
AIDg	Annual Growth Rate of Net Official Development Assistance and Official Aid Received	0.080	0.355	-0.585	1.423439	World Bank
CAPg	Annual Growth Rate of Capital Stock	0.056	0.022	0.029	0.113372	FRED
INVg	Annual Growth Rate of Total Investment as a % of GDP	0.102	0.150	-0.296	0.428973	ODA
LABg	Annual Growth Rate of Total Labor Force	0.017	0.010	0.000	0.029973	World Bank
EDUg	Annual Growth Rate of Enrollment in Primary Education	0.022	0.023	-0.025	0.090057	World Bank
DTHg	Annual Growth Rate of Death Rates per 1000	-0.039	0.135	-1.000	-0.00819	World Bank
RAINg	Annual Growth Rate of Weighted Rainfall	0.070	0.507	-1.000	1.236025	OGD
TEMPg	Annual Growth Rate of Temperatures	0.001	0.012	-0.029	0.033933	OGD
POPg	Annual Growth Rate of Population Density	0.020	0.003	0.012	0.023582	World Bank
CARg	Annual Growth Rate of Total Number of Vehicles	0.114	0.033	0.031	0.207407	OGD

Table 2: Correlation Matrices

Version 1

	year	gdpg	exg	c02g	aidg	capg	dthg	raing	tempg	popg	carg
year	1.0000										
gdpg	0.4789	1.0000									
exg	0.2267	0.0931	1.0000								
c02g	0.0832	0.2038	-0.1833	1.0000							
aidg	0.0678	0.0100	0.0875	-0.0858	1.0000						
capg	0.8704	0.5146	0.3518	0.1638	0.0958	1.0000					
dthg	0.8236	0.4285	0.1748	0.0647	0.0396	0.7720	1.0000				
raing	0.0764	-0.0272	0.1342	-0.2589	0.0326	0.0320	0.0870	1.0000			
tempg	-0.0359	-0.1030	-0.0519	-0.0942	-0.2096	-0.0067	-0.0600	0.0041	1.0000		
popg	-0.9370	-0.4654	-0.2826	-0.0233	-0.0675	-0.9261	-0.7852	-0.0685	0.0326	1.0000	
carg	-0.1959	0.0435	-0.3147	0.3745	-0.1712	-0.2192	0.0230	-0.0616	-0.1002	0.3055	1.0000

Version 2

	year	gdpg	exg	c02g	aidg	capg	invg	labg	dthg
year	1.0000								
gdpg	0.5096	1.0000							
exg	0.4265	0.5231	1.0000						
c02g	0.1088	0.0510	-0.2406	1.0000					
aidg	0.0637	-0.1233	-0.0513	0.1612	1.0000				
capg	0.8953	0.6497	0.4762	0.3385	0.1491	1.0000			
invg	0.4149	0.7739	0.6509	-0.0511	-0.1757	0.4752	1.0000		
labg	-0.6331	-0.3650	-0.1161	-0.3923	0.0783	-0.7225	-0.1050	1.0000	
dthg	0.9492	0.4442	0.4542	-0.0224	0.0229	0.7928	0.4147	-0.4229	1.0000

Table 2: Correlation Matrices (Cont.)

Version 3

	year	gdpg	exg	c02g	aidg	capg	invg	labg	dthg	raing	edug
year	1.0000										
gdpg	0.2462	1.0000									
exg	0.3475	0.3951	1.0000								
c02g	-0.5298	0.3599	0.0213	1.0000							
aidg	-0.4102	-0.4969	-0.4712	0.1986	1.0000						
capg	0.7192	0.6959	0.5195	-0.0064	-0.5012	1.0000					
invg	0.3383	0.7149	0.6623	0.0809	-0.4119	0.5414	1.0000				
labg	0.3247	-0.1891	0.0118	-0.4406	0.0547	0.0416	0.0890	1.0000			
dthg	0.9900	0.1467	0.2960	-0.5465	-0.3513	0.6792	0.2495	0.3376	1.0000		
raing	0.3582	-0.0767	0.5180	-0.4157	-0.2793	0.1743	-0.0259	0.0152	0.3365	1.0000	
edug	0.1900	0.2122	0.5585	0.1174	-0.3904	0.2103	0.3661	0.3686	0.1689	0.1454	1.0000

Table 3: Old and New Regression Statistics

	Variable	EXg	EMITg	AIDg	CAPg	INVg	LABg	EDUg	DTHg	RAINg	TEMPg	POPg	CARg	Constant	F-Statistic	Adj. R2	R2	N
Version 1																		
	Coefficient	-0.018	0.044	-0.003	0.81	-	-	-	0.01	-0.001	-0.232	-0.1	0.111	-0.002	1.6	0.114	0.304	43
	t-value	-0.33	0.24	-0.22	1.17	-	-	-	0	-0.12	-0.61	-0.02	0.6	0.02				
Version 1*																		
	Coefficient	-	0.133	0	-	-	-	-	0.38	-0.001	-0.15	-4.78	0.118	0.138	1.9	0.13	0.27	43
	t-value	-	0.8	-0.02	-	-	-	-	0.23	-0.14	-0.4	-1.63	0.65	4.18				
Version 2																		
	Coefficient	-0.04	-0.239	-0.009	1.272	0.076	0.444	-	-2.48					-0.072	6.45	0.65	0.77	21
	t-value	-0.95	-1.45	-1.02	2.34	2.69	0.71	-	-1.87					-1.16				
Version 2*																		
	Coefficient	-	-0.002	-0.006	-	-	-0.481	-	1.65					0.103	1.32	0.06	0.24	21
	t-value	-	-0.01	-0.51	-	-	-0.75	-	1.46					5.21				
Version 3																		
	Coefficient	-0.196	0.246	-0.009	2.002	0.1	-0.547	0.113	-1.65	0.012	-	-	-	-0.075	3.5	0.65	0.91	13
	t-value	-1.74	0.64	-0.66	1.84	2.48	-0.34	0.41	-0.95	1.16	-	-	-					
Version 3*																		
	Coefficient	-	0.772	-0.03	-	-	0.897	-0.2	1.47	0	-	-	-	0.0467	1.26	0.11	0.55	13
	t-value	-	1.7	-1.8	-	-	0.37	-0.58	0.95	-0.06	-	-	-	0.66				
	Legend	1%	5%	10%			* indicates the regressions in which the colinear variables (EXg,CAPg and INVg) were excluded											

Table 4: 2SLS Outputs

2SLS	Variable	EXg	EMITg	AIDg	CAPg	INVg	LABg	EDUg	DTHg	RAINg	TEMPg	POPg	CARg	Constant	F-Statistic	Adj. R2	R2	N
Version 1 (CAPg)																		
	Coefficient	-	0.26	-0.001	0.703	-	-	-	-	-	-	-	-	0.003	5.04	0.2	0.26	43
	t-value	-	1.02	-0.09	3.27	-	-	-	-	-	-	-	-	0.26				
Version 1 (EXg)																		
	Coefficient	0.062	0.604	0.004	-	-	-	-	-	-	-	-	-	0.022	0.97	-	-	43
	t-value	1.11	1.61	0.26	-	-	-	-	-	-	-	-	-	1.16				
Version 1 (INVg)																		
	Coefficient	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	t-value	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Version 2 (CAPg)																		
	Coefficient	-	0.602	-0.022	0.99	-	-	-	-	-	-	-	-	-0.05	2.42	-0.22	0.02	21
	t-value	-	1.09	-1.46	2.41	-	-	-	-	-	-	-	-	-1.01				
Version 2 (EXg)																		
	Coefficient	0.032	-1.02	0.007	-	-	-1.914	-	-	-	-	-	-	0.131	0.88	-	-	31
	t-value	0.16	-0.37	0.2	-	-	-0.57	-	-	-	-	-	-	0.7				
Version 2 (INVg)																		
	Coefficient	-	-0.54	0.002	-	0.103	-0.738	-	-	-	-	-	-	0.068	8.57	0.6	0.68	21
	t-value	-	-0.09	0.23	-	4.98	-1.03	-	-	-	-	-	-	2.06				
Version 3(CAPg)																		
	Coefficient	-	0.729	-0.026	1.33	-	1.32	-0.218	-	-	-	-	-	-0.068	3.07	0.43	0.66	13
	t-value	-	1.85	-1.46	1.88	-	0.63	-0.76	-	-	-	-	-	-1.17				
Version 3 (EXg)																		
	Coefficient	0.073	0.617	-0.033	-	-	1.388	-0.297	-	-	-	-	-	0.003	1.15	0.204	0.536	13
	t-value	0.8	1.05	-1.65	-	-	0.48	-0.71	-	-	-	-	-	0.04				
Version 3 (INVg)																		
	Coefficient	-	0.229	-0.018	-	0.081	-0.435	-0.125	-	-	-	-	-	0.056	2.95	0.492	0.704	13
	t-value	-	0.45	-0.99	-	2.51	-0.18	-0.41	-	-	-	-	-	0.91				

Legend	1%	5%	10%

Version 1 (CAPg): 2SLS with CAPg; excludes EXg and INVg
Version 1 (EXg): 2SLS with EXg; excludes CAPg and INVg
Version 1 (INVg): 2SLS with INVg; excludes CAPg and EXg

Table 5: 2SLS (CAPg) Comparison with Adjusted OLS Regression

2SLS	Variable	EXg	EMITg	AIDg	CAPg	INVg	LABg	EDUg	DTHg	RAINg	TEMPg	POPg	CARg	Constant	F-Statistic	Adj. R2	R2	N
Version 1																		
	Coefficient		0.26	-0.001	0.703									0.003	5.04	0.2	0.26	43
	t-value		1.02	-0.09	3.27									0.26				
Version 1.1																		
	Coefficient	-	0.121	-0.002	0.738	-	-	-	-	-	-	-	-	0.006	5.06	0.22	0.28	43
	t-value	-	0.87	-0.2	3.59	-	-	-	-	-	-	-	-	0.53				
Version 2																		
	Coefficient		0.602	-0.022	0.99									-0.05	2.42	-0.22	0.02	21
	t-value		1.09	-1.46	2.41									-1.01				
Version 2.1																		
	Coefficient		0.022	-0.008	0.624									0.021	5.42	0.3	0.37	31
	t-value		0.16	-0.96	4									0.011				
Version 3																		
	Coefficient		0.729	-0.026	1.33		1.32	-0.218						-0.068	3.07	0.43	0.66	13
	t-value		1.85	-1.46	1.88		0.63	-0.76						-1.17				
	Legend	1%	5%	10%			Version 1.1: Regression with only EMITg, AIDg, and CAPg (used to compare with 2SLS of V1)											
							Version 2.1: Regression with only EMITg, AIDg, and CAPg (excludes LABg)											

XI. Appendix B: Advice for Future Students

- Start early on your thesis planning (the earlier, the better)
- Learn how to organize your data, so that you can easily refer back to it when necessary
- Become proficient in Excel and Word prior to starting your thesis
- Plan a schedule and try to stick to it (add buffer time for unforeseen circumstances)
- Meet as many times as you can with your thesis advisor, and ask for advice, tips and recommendations regularly