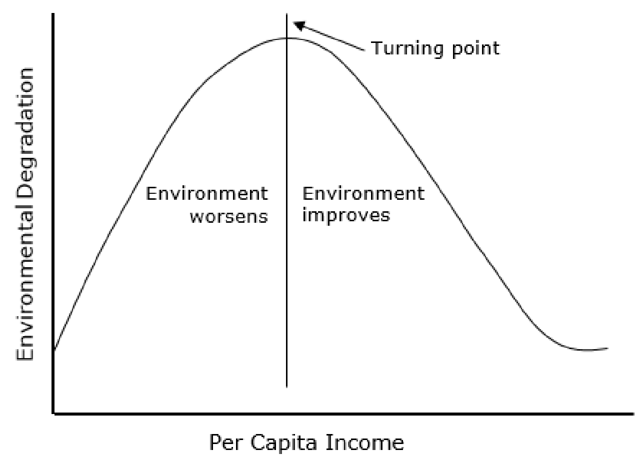
The Environmental Kuznets Curve: An analysis of CO_2 and GDP

The goal of this research paper is to analyze the legitimacy of the Environmental Kuznets Curve (EKC), a hypothesis derived from the works of Simon Kuznets in the 50s and 60s during Grossman and Krueger's research into the potential impacts of NAFTA in the early 90s, with a specific focus on C02 emissions. The main assumptions of the EKC are that pollution indicators will increase in the early stages of economic development, hit a point, and then decrease as a country gets more and more developed. So, the relationship between pollutants and GDP per capita for any country should follow an inverse U pattern. Grossman and Krueger's research argued that NAFTA would have a positive impact on the environment, because wealth and economic growth were believed fundamental in improving it (Stern). Much debate has followed, with research trying to disprove or prove the EKC's existence. I will be analyzing the relationship between GDP/Capita and CO_2/Capita.

In 1992, Wilfred Beckerman wrote a thesis in which he stated that "in the longer run, the surest way to improve [the] environment is to become rich" (491). Beckerman and many other supporters of the EKC theory believe -

> *At higher levels of development, …increased environmental awareness, enforcement of environmental regulations, better technology and higher environmental expenditures, result in [the] levelling off and gradual decline of environmental degradation.*
>
> (Panayatou, 1993, pg 1)

While this theory does sound like it has some rationality, and while it certainly does for some specific countries, others would argue that the scale effect has a larger impact on the relationship between pollutants and economic development than does the effect of changes in the structure and technology of an economy. The scale effect says that pure growth in an economy should result in pure growth in determinants of pollution. These two effects, for the most part, drive the EKC debate. Proponents of the EKC say that the scale effect dominates for developing countries, but then, when a country becomes developed, structural and technological changes overcome the scale effect, leading to decreased pollution and the inverted U shape shown below.



What I have found, and many others have found as well, is that, in regards to CO2 emissions, there is a monotonic relationship between CO2 emissions per capita and GDP per capita – as the GDP of such a country rises, CO2 emissions are predicted to rise also. This theory is contradictory to the EKC model but does not invalidate its hypothesis. The whole theory of the EKC model uses many different dependent variables, including CO2 emissions, NO2, SO2, water quality, and others. When attempting to see the validity of the EKC hypothesis, it is necessary to understand that environmental health is multifaceted and not one dependent variable (such as the CO2 emissions I use in my analysis) can prove or disprove its existence. That being said, my analysis will provide a critique of the EKC, focus on both developed and developing nations and will attempt to show that the common relationship between GDP and CO_2 emissions is monotonic for developing nations but is more complex for developed nations. I will also characterize this relationship.

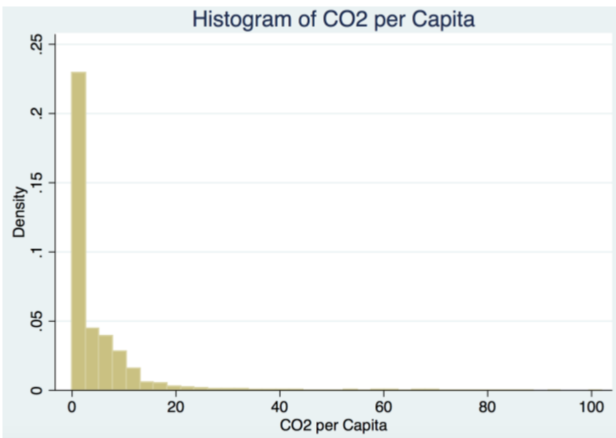
Using the World Bank for data, I downloaded the GDP per capita and CO_2 (metric tons) per capita for 215 countries from every year between 1960 to 2013. US dollar values are gathered from end of year 2015 as the World Bank has them. I have arranged all the data as panel data. For the purpose of this analysis, my variables, after transforming them into natural logarithms, are stationary. While I will be doing individual analysis on a few interesting countries, the bulk of my analysis and research is focused on developed and developing countries. I use the cutoff point of $10,000 GDP per capita as the determinant for developed or developing and a GDP dummy value of 1 means GDP/Capita is greater than $10,000, a GDP dummy value of 0 means GDP/Capita is less than $10,000. I used this cutoff point, which is less than the current cutoff point of $12,475 established by the World Bank, to account for the constant change in its value – it was $6,000 back in 1987.

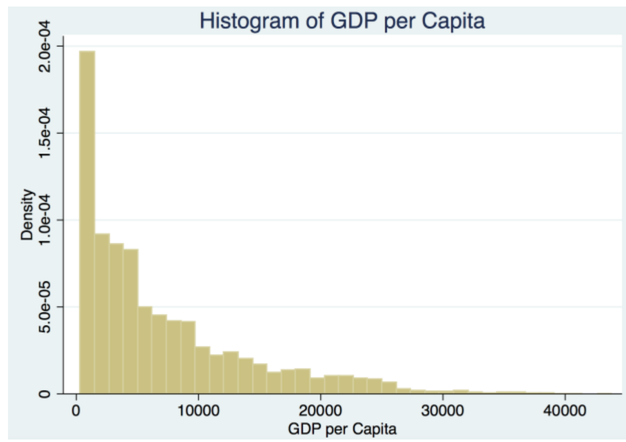
Variable	Obs	Mean	Std. Dev.	Min	Max
t_year	11610	1986.5	15.58646	1960	2013
country_id	11610	108	62.06716	1	215
gdppercapita	11417	3764.672	6131.662	0	43819.1
ln_gdpcapita	6106	8.343611	1.090171	5.86678	10.68783
ln_gdpcapi~2	6106	70.80413	18.00445	34.41911	114.2296
co2_capita	8844	4.357763	7.598609	.0005803	102.0307
ln_co2capita	8844	.2505218	1.796508	-7.451965	4.625274
GDPdummy	11610	.1458226	.3529434	0	1

There are a few things important to note when looking above. Of the 215 countries I downloaded data on, some did not have a reported GDP or CO_2 output. This is why the GDP per capita minimum is 0. Obviously this significantly throws off the true mean for GDP per capita and the true mean for the GDP dummy. I take this into consideration when making my graphs.

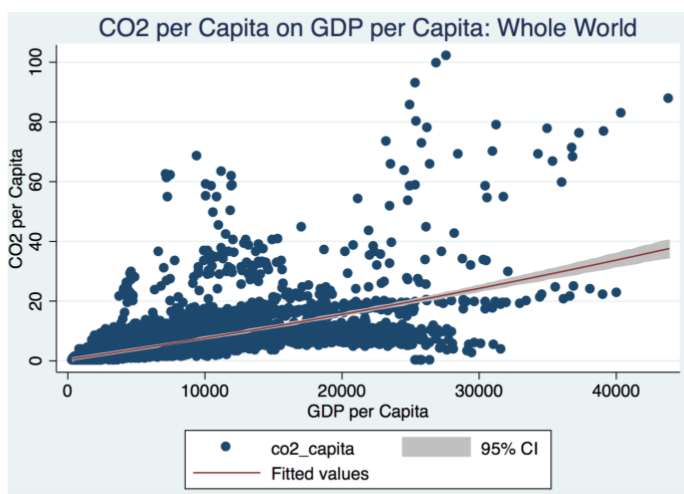
So the observations were limited to 6,106 for GDP per capita and 8,844 for CO2 per capita. When running the regressions for the econometric model I will present later, Stata automatically considers the missing data, so the differences in observations are not of concern. For my regressions, there are a little more than 5,400 observations that I will be relying on.

Here are the histograms for CO2 per capita and GDP per capita. I made them insuring that the zero values for GDP were not included.





As you can see, for both histograms, most of the data is on the left hand side. Most of the countries are still developing and have GDP per capita less than $10,000. And most of our countries produce lower levels of CO2 per capita.



Before I go further, I would like to show a graph of the data I have collected. This graph on the right shows the simple trend between GDP/Capita and CO2/Capita. Higher levels of GDP/Capita are associated with higher levels of CO2/Capita.

The early models of the EKC were simple polynomial equations measuring a pollution determinant with income levels (Stern). The EKC hypothesis was measured by this model:

$$y_{it} = \alpha_i + \beta_1 x_{it} + \beta_2 {x_{it}}^2 + \beta_3 {x_{it}}^3 + \beta_4 z_{it} + \varepsilon_{it}$$

Where Y is CO_2 per capita, alpha is an intercept, x is GDP per capita and z relates to other variables impacting environmental degradation (Stern and Dinda). When trying to make this regression model fit requirements, I found that there was a very low R-squared, and, according to the augmented Dicky-Fuller test using the xtfisher command testing for a unit root in the data, my variables were not stationary. By changing the model into a log-log regression, the data still is not stationary but has a high r-squared and many researchers have worked with this model in recent years. For the first model, I decided to use the standard EKC regression model proposed by many, but brought to my attention by David Stern in his analysis of the EKC.

$$\ln(E/P)_{it} = \alpha_i + \gamma_t + \beta_1 \ln(GDP/P)_{it} + \beta_2 (\ln(GDP/P))^2{}_{it} + \varepsilon_{it}$$

However, this model is still not stationary. The first two terms are intercept parameters that change among different countries (i) and years (t). Because we are dealing with panel data, there are two ways to approach the analysis – one being the fixed effects model and the other being the random effects model. In the fixed effects model, the first two terms Υt and αi are treated as regression parameters which are correlated together. In the random effects model zero correlation is assumed. I ran a Hausman test which let me know that I should be doing a fixed regression model. It is shown below.

```
             ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |     fixed        random       Difference          S.E.
-------------+----------------------------------------------------------------
ln_gdpcapita |    3.371785     3.518519       -.1467335         .0152103
ln_gdpcapi~2 |   -.1517363    -.1584502        .0067139          .000723
------------------------------------------------------------------------------
                           b = consistent under Ho and Ha; obtained from xtreg
            B = inconsistent under Ha, efficient under Ho; obtained from xtreg

    Test:  Ho:  difference in coefficients not systematic

                  chi2(2) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =      103.85
                Prob>chi2 =      0.0000
```

From this you can see that the differences in the coefficients are systematic (I reject the null hypothesis) because the p-value is less than any significance level, so I proceeded to run my regression controlling for the fixed effects of time and country. I will be running my fixed effects regression for country, only time, and then country and time. For the time fixed effects, β_1 and β_2 are the average slopes for all country per time. For the country fixed effects, β_1 and β_2 are the average slopes for all years (t) per country. And for the fixed effects of both country and time, β_1 and β_2 is per country and per year, controlling for country(i) and year(t) fixed effects. All of these fixed effect regressions are calculated with robust standard errors to control for heteroskedasticity.

Using the esttab command in STATA I generated the following regression, fixing effects for country. Taking note of the coefficients here, I can see that the model would look like this:

$$\ln\left(\frac{CO2}{Capita}\right) = \beta_0 + 3.372\ln\left(\frac{GDP}{Capita}\right) + (-0.152)(\ln\left(\frac{GDP}{Capita}\right))^2$$

Where β_0 will change depending on the specific country. Here you can see that both coefficients are statistically significant at the .001 level and that the adjusted R-squared is very high at 94.2%.

Country fixed effect

	(1) ln_co2capita	(2) ln_co2capita
ln_gdpcapita	0.807***	3.372***
	(35.40)	(15.81)
ln_gdpcapi~2		-0.152***
		(-12.34)
_cons	-5.889***	-16.65***
	(-30.79)	(-18.06)
N	5426	5426
adj. R-sq	0.939	0.942

t statistics in parentheses
*** p<0.001

A 1% change in GDP/Capita corresponds to a 3.372% change in CO2/Capita. 3.372 is equal to the elasticity of our dependent variable with respect to our independent variable. The turning point is the GDP/Capita at which CO2/Capita is supposed to be at its highest. Where, when GDP/Capita goes any higher, CO2/Capita is supposed to decrease. For this econometric model, the turning point can be calculated with e^-($\beta_1/2\beta_2$). So, the turning point here is e^(3.372/0.304) which equals $65,650.81. Following fixed effects only for country, we can see that the GDP/Capita a country should obtain before any increases in it would result in an expected decrease in CO2/Capita is $65,650.81. Now I will generate, also using the esttab command in Stata, the fixed effects model for only time.

The econometric model is:

$$\ln\left(\frac{CO2}{Capita}\right) = \beta_0 + 4.953\ln\left(\frac{GDP}{Capita}\right) + (-0.217)(\ln\left(\frac{GDP}{Capita}\right))^2$$

Time fixed effect

	(1)	(2)
	ln_co2capita	ln_co2capita
ln_gdpcapita	1.383***	4.953***
	(119.87)	(28.56)
ln_gdpcapi~2		-0.217***
		(-20.60)
_cons	-11.05***	-25.48***
	(-79.25)	(-35.11)
N	5426	5426
adj. R-sq	0.772	0.795

t statistics in parentheses
*** p<0.001

β_0 here now changes depending on the specific time in years. Both coefficients are still statistically significant but there is a change in them. β_1 went up by 1.581, meaning a 1% change in GDP/Capita corresponds to an expected 1.581% higher increase in CO2/Capita than in the country fixed effects model. β_2 is .065 less than in the country fixed effect model. The turning point here is calculated as being e^(4.953/0.434) = $90,440.04. Significantly higher than for country fixed effects.

Now I will generate the fixed effects model for both country and time together. The model and regression are shown below and to the right, respectively.

$$\ln\left(\frac{CO2}{Capita}\right) = \beta_0 + 3.125\ln\left(\frac{GDP}{Capita}\right) + (-0.150)(\ln\left(\frac{GDP}{Capita}\right))^2$$

Time and country fixed

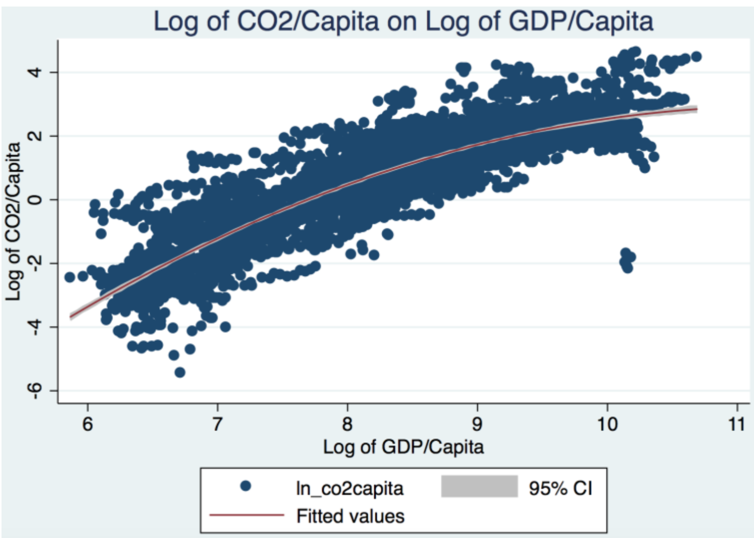
	(1) ln_co2capita	(2) ln_co2capita
ln_gdpcapita	0.606***	3.125***
	(16.02)	(15.99)
ln_gdpcapi~2		-0.150***
		(-12.89)
_cons	-4.855***	-15.36***
	(-17.10)	(-18.17)
N	5426	5426
adj. R-sq	0.946	0.949

t statistics in parentheses
*** p<0.001

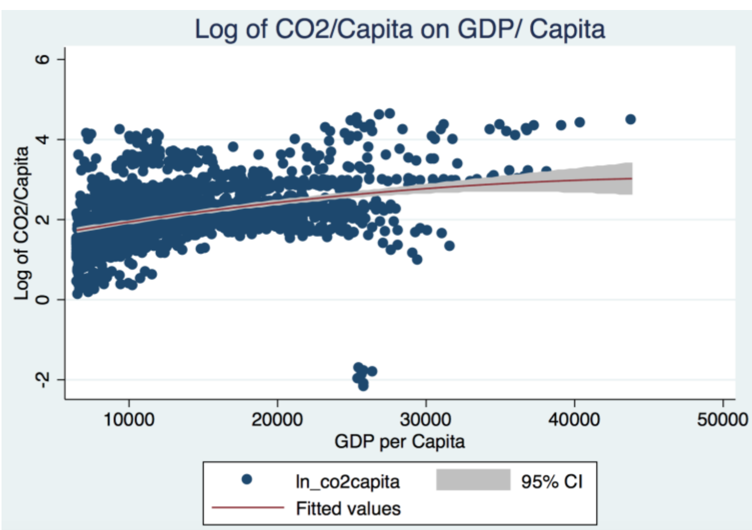
β_0 is dependent on both time and country. The coefficients of $\beta_1$is 3.125 and of β_2 is -0.15 and they are both still statistically significant at the same level. These values are lower than the previous two regressions that were fixing country or time independently. The turning point generated with time and country fixed effects is e^(3.125/0.3)= $33,411.88. Much, much lower than the values I generated previously. Which controlling variable makes sense? After some thought I believe it makes sense that fixing country doesn't work. The differences across countries, the government and the culture and the technology and the structure, have an influence on the dependent variable. Because of this, we can assume that the variation across countries is random and we can allow these differences to be absorbed by the intercept. Time, however, is time variant and is not constant. We hold time fixed because time has non-random characteristics that we don't want to be added to our error term. By not fixing for country, we are allowed to infer our results to all countries. So holding time and country fixed is not plausible. Holding only country fixed is also not plausible. So we are left with the fixed effect model that fixes solely time. Meaning the turning point of $90,440 is the most accurate model.

Which one of these actually makes sense? We can't be sure that any of them make sense. Why? Well because the model we were using doesn't track the growth of CO2/Capita.

And we are trying to figure out the economic point, the GDP per capita, at which the growth in CO2/Capita actually equals zero and where higher GDP per capita should result, not just in continued zero growth in CO2 emissions, but a decrease in CO2 emissions – where growth is negative. We are trying to see if there is a point where further growth in GDP per capita, which is our measure of development, should result in lower CO2 emissions per capita. The scale effect holding true, it is predicted that increasing GDP/Capita should result in positive CO2/Capita growth and decreasing GDP/Capita should result in negative CO2/Capita growth.



By looking at the two graphs on this page, you can see a pretty clear trend of a half arc. We don't see the full arc proposed by the EKC. When will the line start to decrease? At what GDP/Capita should countries expect decreasing CO2/Capita?



I decided to transform the model to be the growth of our variables over time. This logarithmic model would help us determine the impact of changes in our independent variable on changes in our dependent variable. This model is also stationary. Here we use the difference of the natural log of CO2 per capita as our dependent variable and the difference of the natural log of GDP per capita as my independent variables. Simply, it's the same econometric model we used before except we changed it to measure growth in CO2/Capita and growth in GDP/Capita. Our new model is:

$$d.\ln\left(\frac{CO2}{Capita}\right) = \beta_0 + \beta_1 d.\ln\left(\frac{GDP}{Capita}\right) + \beta_2 (d.\ln\left(\frac{GDP}{Capita}\right))^2$$

Here is a summary of the new variables:

Variable	Obs	Mean	Std. Dev.	Min	Max
GrCO2cap	8642	.0253861	.2136992	-3.960807	3.642912
GrGDPcap	5969	.0178107	.0574567	-1.019272	.5692601
GrGDPcap2	5969	.0036179	.018364	2.63e-10	1.038916

Now the question is whether or not we should use the fixed effect model for this model. I believed we should, until I ran the Hausman test again. This test let me know that I should in fact use the random effect model where the unique errors are uncorrelated with the regressors. The Hausman test is below:

	Coefficients (b) fixed	(B) random	(b-B) Difference	sqrt(diag(V_b-V_B)) S.E.
GrGDPcap	.7071095	.6933791	.0137304	.0115836
GrGDPcap2	.2805624	.3056417	-.0250792	.0313914

```
                 b = consistent under Ho and Ha; obtained from xtreg
  B = inconsistent under Ha, efficient under Ho; obtained from xtreg

Test:  Ho:  difference in coefficients not systematic

              chi2(2) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                      =        2.97
            Prob>chi2 =      0.2268
```

It is clear that the p-value (0.2268) is greater than the 0.05 significance level and I fail to reject the null hypothesis. Therefore, I should use the random effect model instead of the fixed effect model. The problem with this model is that it often results in omitted variable bias. However, I must use another test called the Breusch and Pagan Lagrangian multiplier test for random effects. This will determine if I should use just a simple OLS regression or the random effects. This test is the one to the right. As you can see, I fail to reject the null – no evidence of significant differences and no panel effect. So I run a standard OLS regression below.

```
Breusch and Pagan Lagrangian multiplier test for random effects

        GrCO2cap[country_id,t] = Xb + u[country_id] + e[country_id,t]

        Estimated results:
                        |       Var     sd = sqrt(Var)
                ---------+-----------------------------
                GrCO2cap |   .0291323       .1706818
                       e |    .027598       .1661265
                       u |   .0001786       .0133656

        Test:   Var(u) = 0
                             chibar2(01) =     0.22
                          Prob > chibar2 =   0.3178
```

It should be noted that we can't find the turning point value using this model of growth. The model looks like this:

$$d.\ln\left(\frac{CO2}{Capita}\right) = 0.006 + 0.69d.\ln\left(\frac{GDP}{Capita}\right) + 0.32(d.\ln\left(\frac{GDP}{Capita}\right))^2$$

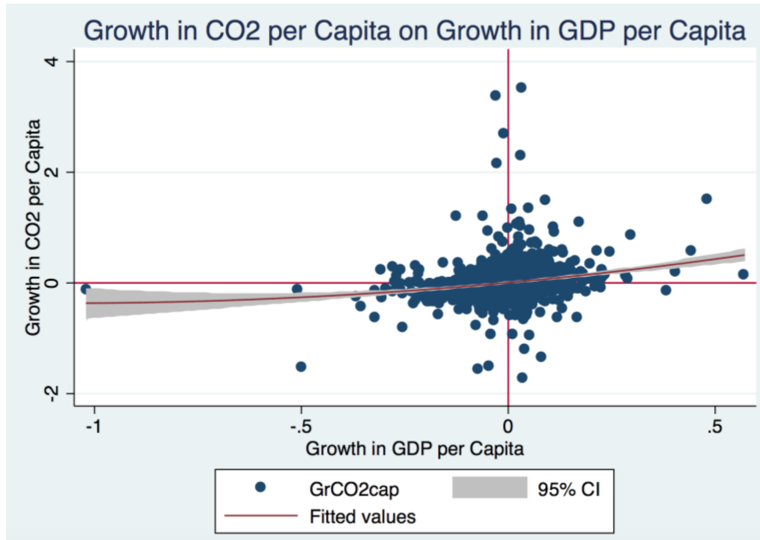
The regression shows that β_2 is not significant and that β_1 is significant to the 0.001 level. Notice how low the adjusted R-squared is. It is not necessary to have a very high R-squared in panel data.

What is very interesting is the graph of this model. It shows the very clear prediction and the scale effect in works. Negative growth in

	(1) GrCO2cap	(2) GrCO2cap
GrGDPcap	0.668***	0.690***
	(8.68)	(9.35)
GrGDPcap2		0.320
		(1.06)
_cons	0.00740*	0.00581
	(2.57)	(1.92)
N	5303	5303
adj. R-sq	0.049	0.049

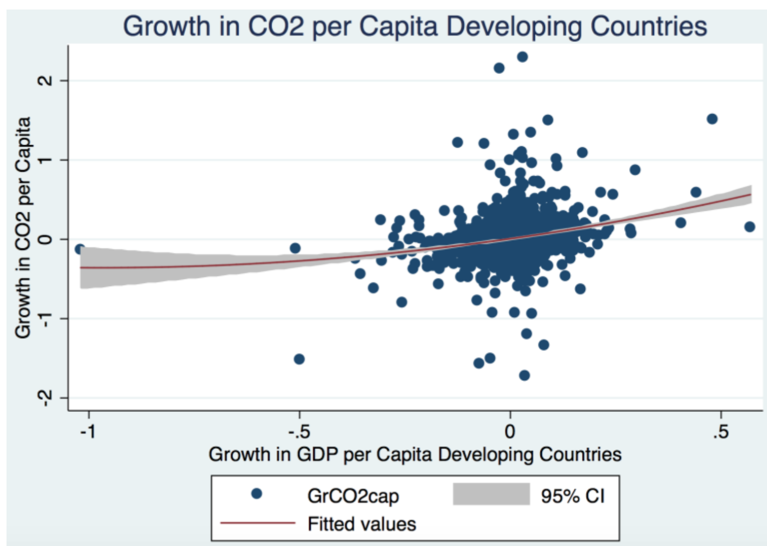
t statistics in parentheses
* p<0.05, ** p<0.01, *** p<0.001

CO2/Capita is associated with negative growth in GDP/Capita and positive growth in CO2/Capita is associated with positive growth in GDP/Capita.

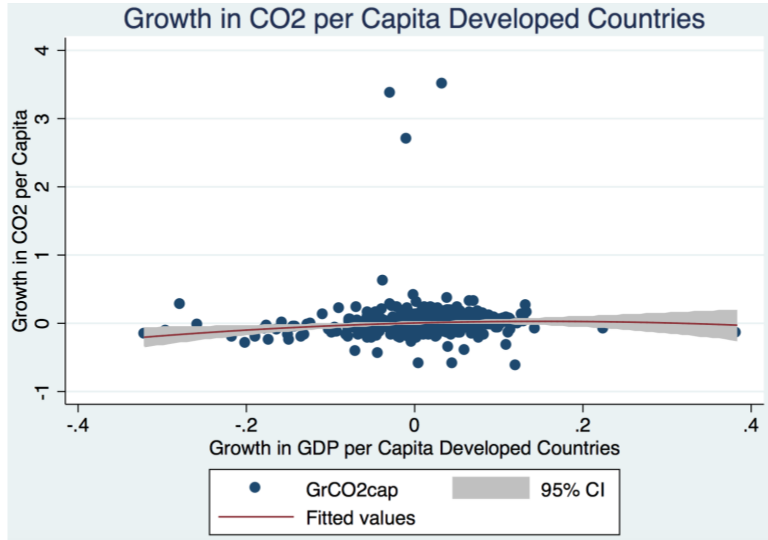


For all the countries in my data, the best fit quadratic line shows this relationship. It seems that, in general, when a country expands and increases their GDP/Capita, one should expect an increase in CO2/Capita.

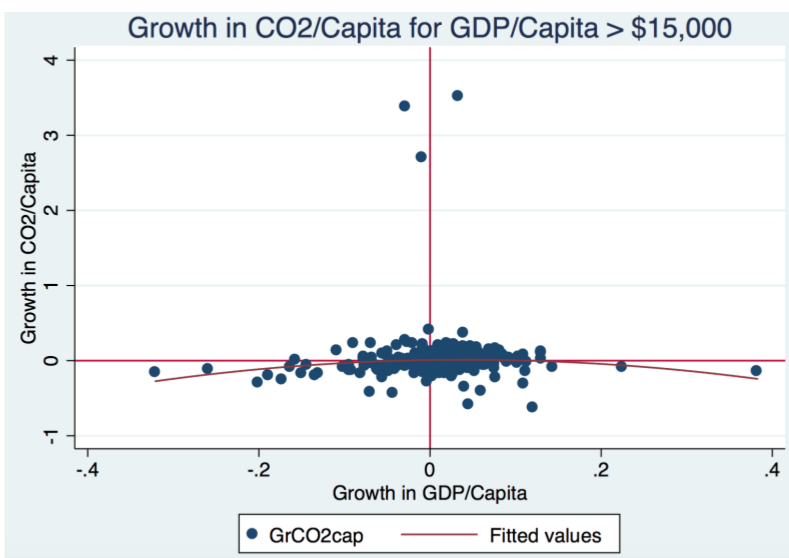
What does it look like for developing countries with GDP/Capita<$10,000? Well, as predicted, developing countries have higher growth in CO2/Capita when they are growing GDP/Capita than the rest of the world – demonstrated by a more pronounced upward angling curve. And both graphs on this page have positive second derivatives.



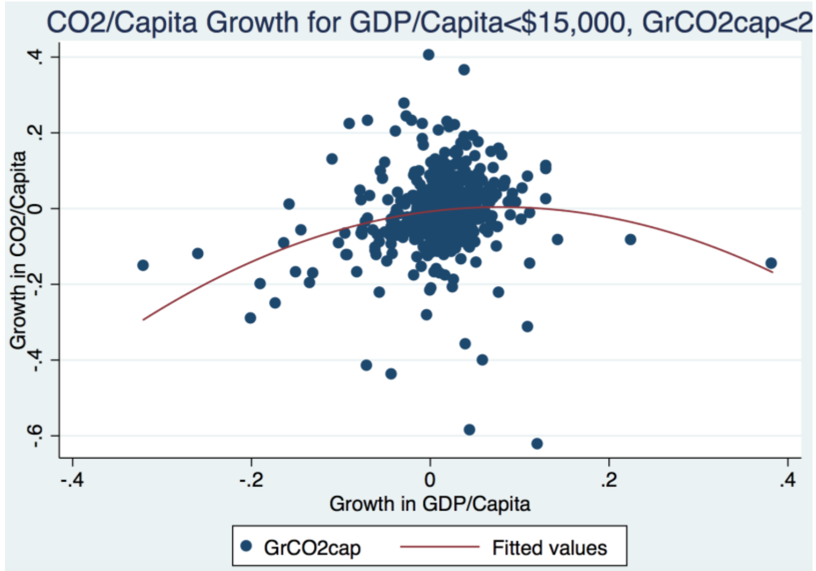
What about developed countries? This is where things get interesting. As you can see, no longer do we have a clear upward sloping line. For countries with a GDP/Capita >$10,000, it looks as though growth in GDP/Capita corresponds to almost zero growth in CO2/Capita, maybe even very slightly negative growth!



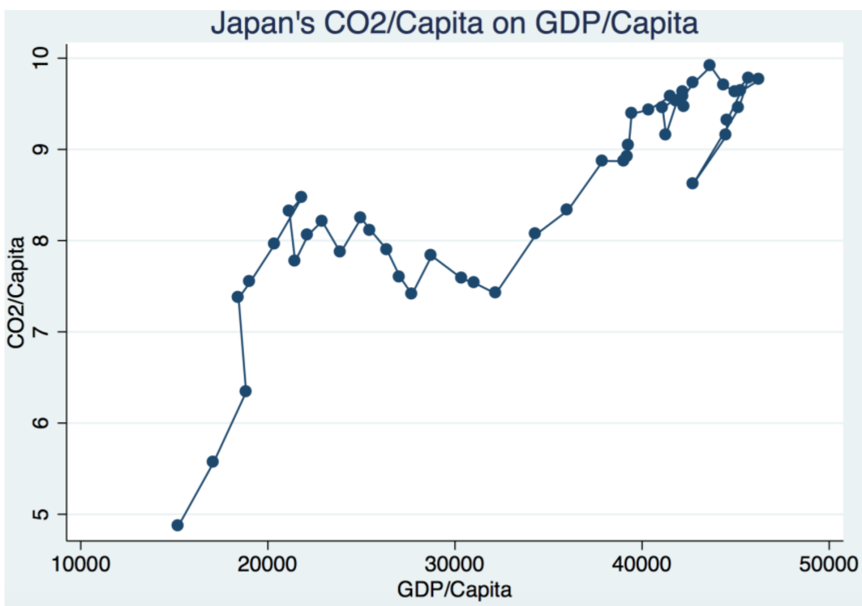
This intrigued me so I increased the threshold to being GDP/Capita > $15,000. And I generated this graph below. As you can see, for developed countries with a GDP/Capita > $15,000, there is a very clear downward quadratic fit following growth in GDP/Capita.



What would happen to the best fit quadratic if I remove the three oddities that have growth in CO2/Capita > 2? So, this is the graph I made that plots the growth in CO2/Capita against growth in GDP/Capita for developed countries with GDP/Capita greater than $15,000 and growth in CO2/Capita less than 2. The result is a graph that shows a very clear inverse U. The graph does look a bit varied and maybe tough to tell if the fitted line gives a legitimate trend line.

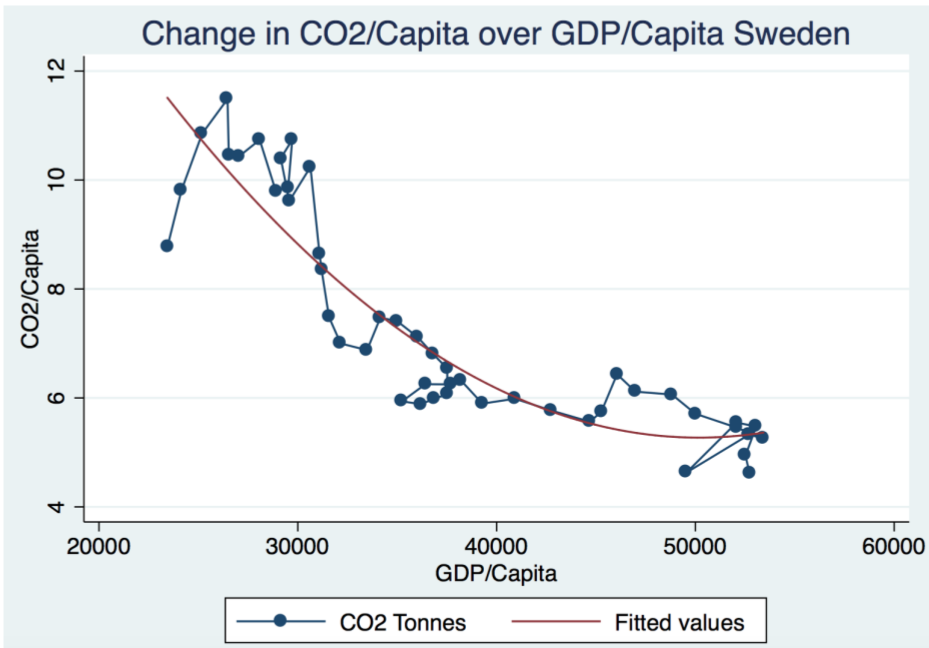


Before I go into my conclusion. I want to go into a couple specific countries – Japan and Sweden. Japan has an interesting graph that portrays the cultural and governmental changes towards pollution and the environment in the 1970s. Japan had a massive increase in CO2/Capita up until the early 1970s, before laws and regulations were put in place to combat the environmental problems. CO2/Capita decreased, then has been on the upswing mostly since the mid 80s. Japan is a good



example of how a country's specific characteristics influence the relationship between pollution and economic development, and shows how validating an EKC model can be very hard given individual country changes.

Sweden is a unique example which demonstrates that a country can manage to reduce CO2/Capita consistently whilst simultaneously growing its economy. This does not show the reported EKC curve; however, it does demonstrate that, once a country becomes developed, reducing pollution is indeed possible. Sweden is the best example I have found of a country that has consistently managed to decrease average pollution levels per capita whilst increasing GDP per capita.



To conclude, I have given a brief yet detailed background to the Environmental Kuznets Curve and explained both of my econometric models. I showed the log-log regression and its fixed effects regressions and estimated a turning point of $90,440 . I explained the logarithmic regression's short comings and hypothesized a new growth oriented model. The growth model showed how negative growth for the pollution parameter is indeed possible for developed countries who are growing their GDP/Capita and showed how developing countries, more than developed, do exhibit increased pollution associated with increased GDP/Capita. "It is a significant condition that only when income grows, the effective environmental policies can be implemented" (Dinda pg. 431). Just as Dinda explained and many other researchers today also explain, wealth does play a large role in fixing environmental barriers and structural problems. Without money, the necessary steps needed to enforce deals such as the recent Paris Agreements cannot be taken. Without money, straying away from fossil fuels in favor of renewable energy sources is difficult, if not impossible. However, at the same time as developing countries struggle, current globalization and interconnectedness provides more ways than ever before to bring our most powerful technologies to the developing world. While I characterize the relationship between CO_2/Capita and GDP/Capita as positive but with a negative second derivative, there are examples, such as Sweden, that go against my hypothesis. I cannot predict what the future will hold, and if the EKC will gain more support and foundation, but I can say that the turning point I calculated has not yet been reached, and that the second half of the arch has not yet been discovered. In the future, more research needs to be done checking for integration and cointegration of panel data and additional explanatory variables should be used. Education and government expenditures for starters could be useful variables. While my research does shed some light on the lacking EKC for carbon dioxide, there is much evidence out there in support of the EKC hypothesis for local pollutants. Is the EKC hypothesis wrong? No. It is only a hypothesis that has practical applications and prompts thought and environmental research, which is what we simply need.

Bibliography

Beckerman, Wilfred. (1992). Economic growth and the environment: Whose growth? whose environment? World Development, 20(4), 481-496.

Dinda, Soumyananda. (2004). Environmental Kuznets Curve Hypothesis: A Survey. Ecological Economics, 49(4), 431-455.

Panayotou, Theodore. (1993). Empirical tests and policy analysis of environmental degradation at different stages of economic development. 1-2011.

Stern, David. (2004). The Rise and Fall of the Environmental Kuznets Curve. World Development, 32(8), 1419-1439.