 doi:10.20944/preprints201905.0058.v1

Have greenhouse gas emissions from US energy production peaked?

State level evidence from six subsectors

Michael Cary[✓]

[✓]Division of Resource Economics and Management, West Virginia University,

email: macary@mix.wvu.edu

May 3, 2019

Abstract

Analyses of the Environmental Kuznet's Curve (EKC) hypothesis have largely focused on economy level data with occasional analyses exploring sector level data. This paper exploits a new data set which contains sector level data on greenhouse gas emissions from the US energy sector as well as subsector data from six disjoint subsectors which together comprise the entire energy sector. The data contained in this data set is annual data at the state level from 1990 through 2011. A battery of EKC models are tested and some evidence supporting the EKC hypothesis is found for the US energy sector and its subsectors. We find that aggregated subsector level estimates outperform sector level estimate on in-sample accuracy. These estimated models are then used to forecast emissions for the energy sector. We find that EKC estimations for the energy sector as a whole are consistently lower than aggregated subsector EKC estimates. Addressing the title, we find limited evidence at best that US greenhouse gas emissions are at or near a peak.

Keywords: Emissions, Energy, Environmental Kuznet's Curve, Panel Data, Subsector Analysis

JEL Codes: Q53, Q54, Q56, Q57

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2019 doi:10.20944/preprints201905.0058.v1

Contents

1 Introduction

The primary source of greenhouse gas (GHG) emissions from an economy is the energy production sector. Marrero (2010) determined that greater than 65% of the world's GHG emissions come from energy production. According to data from the WRI CAIT Climate Data Explorer (2014), the data set used for this paper, during the period from 1990 through 2011 the energy production sector accounted for more than 88% of total GHG emissions in the United States. Considering the contribution to total GHG emissions from the energy sector of the US economy, any analysis or forecast of GHG emissions in the United States must seriously consider the dynamics of GHG emissions due to energy production.

With new data from the WRI CAIT Climate Data Explorer data set, we explore for the first time the dynamics of six subsectors of the energy production sector of the US economy. By order of magnitude of GHG emissions, these six subsectors are: electric power, transportation, industrial, residential, commercial, and fugitive emissions. By not only studying the energy sector as a whole, but also by studying the behavior of its subsectors, we can improve our understanding of GHG emissions from energy production and thereby enhance forecasts of GHG emissions.

Since the quantity and composition of GHG emissions varies across sectors of the economy, and similarly across subsectors of the energy sector, no two subsectors of energy production should expect to benefit equally from any given emissions mitigation program. Moreover, technological advancement does not affect the energy production sector ubiquitously. By breaking the energy sector down into subsectors, we can increase our ability to estimate the effects of subsector specific technological advancements or emissions mitigation policies. As energy consumption habits are known to change with income, see, e.g., Caron and Fally (2018), energy production portfolios should vary across income levels and thus across states. For all of these reasons, finding an overall trend in GHG emissions from energy production will not be as accurate as finding an aggregated trend of the subsectors of energy production.

 doi:10.20944/preprints201905.0058.v1

In order to study energy subsector GHG emissions, we use the standard Environmental Kuznet's Curve (EKC) approach. Under this approach, GHG emissions are estimated primarily as a function of GDP per capita. The EKC hypothesis states that there is an inverted "U" shape relationship between GHG emissions and GDP per capita. To test this hypothesis, models using the EKC approach include both GDP per capita and GDP per capita squared (often times the third power of GDP per capita is included as well). The ecological consequences of this hypothesis are that, if true, GHG emissions are not a monotonically increasing function of production. That is to say, as production continues to increase, overall levels of emissions may actually decrease. Possible explanations for this behavior include the adaption of greener technologies and the switch to renewable sources of energy.

The two overwhelming drivers of emissions from the energy production sector are electric power generation and transportation. Both electric power generation and transportation are areas which are experiencing rapid growth in the adaptation of greener technologies, the switch to renewable sources of energy, and in the growth of regulation which imposes, through various means, emissions limitations (Fox et al., 2017). However, the methods employed to achieve emissions reductions in these two subsectors are not identical, so our method of isolating these subsectors and forecasting them independently of each other allows us to determine not only if the EKC hypothesis is likely for the energy production sector as a whole, but how valid the EKC hypothesis is for each subsector. Moreover, by assessing whether or not the EKC hypothesis is likely at the subsector level, better policy proposals and decisions can be made through this more in depth targeting of the energy production sector.

The remainder of this paper follows the following structure. After providing a review of the relevant literature, we describe our data set and the methodologies we employ in the effort to test the EKC hypothesis for the energy sector and its subsectors in the United States. At this point the results from the various regression models are presented and their results and implications for the EKC hypothesis are discussed. In particular, we discuss how well each subsector fits the EKC hypothesis in the context of each of the models. Once we have an understanding of these results, we describe our forecasting technique which comes from Selden and Song (1994) and forecast state level and total US GHG emissions for the energy production sector both independently and as an aggregate of the forecasts for each of its subsectors. To determine the validity of our aggregated forecasting technique, we show that aggregated fitted values from subsector models have greater in-sample accuracy in the energy sector level data.

We then discuss these forecasts and what they imply both economically and ecologically going forward. We conclude this paper with a discussion of our testing of the EKC hypothesis for state level GHG emissions from the energy production sector and its subsectors, our forecasting techniques, their results and implications, and add possible avenues for furthering this research.

2 Literature Review

Selden and Song (1994) popularized the EKC hypothesis as a means for studying various GHG emissions. After testing the EKC hypothesis, which they found evidence for in the case of several greenhouse gases, they forecasted future emissions and found that GHG emissions were likely to taper off, though not decrease, over the next several decades. Since then, several papers have studied and tested the EKC hypothesis in various contexts.

Early support for the confirmation for the EKC hypothesis was given by List and Gallet (1999). From there, numerous case studies on various nations, US states, and global regions were performed. Friedl and Getzner (2003) studied the EKC hypothesis in Austria and found evidence for the standard EKC model with a cubic term for GDP per capita included. Similarly, Shittu et al. (2018) studied the EKC hypothesis in Malaysia, Kunnas and Myllyntaus (2007) studied the case of Finland, and Soytas and Sari (2009) did the same in Turkey. At the state level in the United States, Konan and Chan (2010) performed a case study on Hawai'i and Yang et al. (2009) performed a case study on California. On a larger geographic scale, Zoundi (2017) studied the EKC hypoth-

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2019 doi:10.20944/preprints201905.0058.v1

esis in Africa. This analysis found ambiguous evidence, showing that in Africa CO_2 levels rise monotonically with rises in income.

However, not every study has provided the same level of support for the EKC hypothesis. Azomahou et al. (2006) found that the EKC hypothesis has more or less support in different countries and under different statistical approaches. The standard model for EKC hypothesis testing did exhibit evidence for the EKC hypothesis, but a non-parametric model found only limited evidence which varied widely by country. In the context of sulfur emissions specifically, Stern and Common (2001) found that there was support for the EKC hypothesis, but only for high income countries. The standard EKC model found monotonically increasing sulfur emissions when using a global data set.

Spatial models were first used as a tool for assessing the EKC hypothesis by Rupasingha et al. (2004). Extending the use of more sophisticated econometric approaches, Maddison (2006) studied a spatial variant of the EKC hypothesis for national emissions of SO_2 and NO_x and found evidence that national emissions are lowered by proximity to nations with high per capita income, contradicting the notion that nations must achieve increased environmental quality at the expense of their neighbors. Burnett et al. (2013a) also studied several spatial panel data models in the context of CO_2 emissions in the United States, and Ordás Criado (2008) performed a case study on Spain using non-parametric estimation techniques.

In addition to testing the EKC hypothesis with semi- and non-parametric and spatial econometrics models, recent emphasis has been placed on validating estimations and on the importance of the data itself. Chow and Li (2014) established a method using t-tests for validating conclusions supporting the EKC hypothesis using standard econometric approaches. However, Millimet et al. (2003) demonstrated that semi-parametric techniques tend to be superior to standard econometric approaches. As for the choice of data set on the part of the researcher, both Harbaugh et al. (2002) and Galeotti et al. (2006) showed that the data set chosen does not significantly impact any conclusions drawn from statistical analyses.

In lieu of testing the EKC hypothesis itself, the topic of convergence of per capita GHG emissions has also garnered recent attention in the literature. Aldy (2006) studied the convergence of per capita GHG emissions in OECD countries and found evidence supporting this idea. Criado and Grether (2011) found with non-parametric methods that convergence in per capita emissions existed only within groups of similar countries. Aldy (2007) found no evidence, in the context of the United States, for state-wise convergence in per capita emissions, while Burnett (2016) found a group of 26 states for which there is support for convergence in emissions. Relatedly, Tol et al. (2006) studied the EKC hypothesis for emissions intensities and Burnett et al. (2013b) found the economic growth drives not absolute emissions but emissions intensities.

Not all sectors of the economy are equally responsible for producing emissions. For this reason several papers have explored the EKC hypothesis for certain sectors of the economy. Using data from the industrial sector, Edelenbosch et al. (2017) forecasted decreasing energy intensities for this sector in non-OECD countries following a brief increase in industrial energy demand in the short term. Suri and Chapman (1998) first studied the EKC hypothesis in the context of energy consumption. Expanding on their results, Soytas et al. (2007) investigated energy consumption in the United States and found that increases in energy use, not increases in income, cause carbon emissions in the United States by using a Granger causality test. This result is noteworthy because it implies that economic growth alone is not a solution for long term emissions reductions.

Furthering these sector level analyses, Bruvoll and Medin (2003) looked at potential economic factors driving air pollution and found that changes in the relative size of different sectors of the economy helps drive changes in emissions. Mazzanti et al. (2008) used data on value added and capital stock per employee as covariates and found that the inclusion of these measures of labor productivity led to mixed support at the sector level for the EKC hypothesis in the case of Italy.

In the United States, another interesting way to subdivide emissions is at the state level. Aldy (2005) studied per capita emissions at the state level, finding that the estimated EKC varied by state. Using a spatial autoregressive model, Burnett and Bergstrom (2010) found further evidence for the

 doi:10.20944/preprints201905.0058.v1

EKC hypothesis in the United States. Auffhammer and Steinhauser (2012) found that, while standard EKC models did well with in sample accuracy (R^2, Akaike Information Criterion, Schwartz Information Criterion, and others), they were not the top performing models for out of sample testing. Combining the state level and sector specific analyses, Zhou and Gurney (2011) used a spatial econometric model to study sector specific emissions and the state and county level in the United States. Using surface temperature data, they found that emissions due to space cooling in warmer climates should more than offset any future emissions reductions due to reduced needs for space heating. A study by Apergis et al. (2017) found limited evidence for the EKC hypothesis at the state level (10 of 48 states fit the hypothesis) using a standard EKC model without a cubic term.

Using a similar basis, namely province level EKC testing in China, Auffhammer and Carson (2008) forecasted aggregate Chinese emissions. Of particular interest in this paper was the response by the forecasted Chinese emissions to the Kyoto Protocol. More generally, the theme of increased emissions mitigation techniques, ranging from legislation to greener technologies to socioeconomic inequality and beyond, to combat global climate change and increased air pollution have also been studied in the context of the EKC hypothesis. Magnani (2000) studied how income inequality affects investment into emissions mitigation R&D and found that income inequality leads to a gap between the willingness to pay of a country and its ability to pay.

The negative externalities associated with GHG emissions and how they will impact the economy are currently attracting much attention from interdisciplinary researchers. Fankhauser and Tol (2005) studied the effects of climate change on capital accumulation, finding that forward looking agents are likely to change their savings behavior which will lead to reduced economic growth. A growth model with endogenous emissions reduction was created by Criado et al. (2011), and by using data from European nations they were able to validate this theoretical model. Simulations of the impact of automation on emissions at the household level were performed by Ringel et al. (2019), demonstrating evidence that automation of lighting, heating, and cooling in a household can help reduce emissions. Al-Mulali et al. (2015) studied contributing factors to air pollution in Vietnam and Yates and Strzepek (1998) found that climate change had a minimal impact on agricultural production in Egypt. Beyond agricultural production, the negative externalities due to emissions are vast and include depletion of the ozone layer, primarily due to NO_x emissions (Ravishankara et al., 2009). Additionally, impacts on human health due to emissions, including increased risks of various chronic respiratory diseases and cancers (Kampa and Castanas, 2008), are also concerns which emissions mitigation policies attempt to address. The economic impact of increased health risks due to emissions are well documented (West et al., 2013).

We conclude the literature review section of this paper by acknowledging several well written survey papers on the EKC hypothesis which the reader may find helpful, including (Dinda, 2004), (Gill et al., 2018), (Stern, 2004), and (Stern, 2017).

3 Data

Our data set is from the WRI CAIT Climate Data Explorer (2014). This data set is the only data set known to the author which includes GHG emissions data not only for the energy production sector, but for its comprising subsectors (electric power generation, transportation, industrial, residential, commercial, and fugitive emissions). This data considers total GHG emissions in Mt CO_2 equivalent rather than by Mt of individual chemical species. This data is at the state level, includes estimates of GDP per capita and population, and covers the time period from 1990 to 2011. While 22 years is a relatively short period of time, the novelty and level of specification of this data set allows us to be the first to analyze GHG emissions at the state level for subsectors of the US energy production sector.

The socioeconomic data used in this study includes GDP per capita and population density. The WRI data set included GDP per capita and population. To derive population density, the given population of each state was divided the area of the state in square kilometers.

The data for the subsectors exhibit several important quantitative and qualitative properties.

 doi:10.20944/preprints201905.0058.v1

First, the sum of GHG emissions over all subsectors in a fixed year and state are equal to the total GHG emissions from the energy sector in that same state and year. Qualitatively, there are two primary drivers of the energy production sector (electric power generation and transportation), and the three largest subsectors account for nearly 90% of GHG emissions from US energy production. The average contribution of each subsector to total GHG emissions from energy production over all states and years in the data set is summarized in the table below.

Table 1:
Contribution of each subsector to
GHG emissions from energy production

Subsector	Contribution
Electric Power	36.23%
Transportation	32.16%
Industrial	18.60%
Residential	6.18%
Commercial	3.90%
Fugitive Emissions	2.91%

4 Methodologies

This paper has two primary goals: to test the EKC hypothesis for GHG emissions from the US energy production sector and its subsectors using several variants of the standard parametric model, and to forecast GHG emissions from US energy production using these results. In this section we detail the regression models used to test the EKC hypothesis on the data set.

As detailed in the literature review section of this paper, several regression techniques, including fixed effects models, random effects models, semi- and non-parametric techniques, and spatial regression models, have been used on various data sets to test the EKC hypothesis. In this paper we use four different models, each tested with and without a cubic term for GDP per capita, for a total of eight regression models applied to each of the six subsectors of the energy production

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2019 doi:10.20944/preprints201905.0058.v1

sector as well as the energy production sector itself. All regression models specified in this paper are variants of the standard EKC model which estimates GHG emissions primarily as a function of GDP per capita. Eventually, the estimated coefficients from all of these models will be used in our forecasts.

The first model we use is a basic OLS model with the following form in which the variable $m_{s,t}$ represents total GHG emissions from state s in year t, $y_{s,t}$ represents GDP per capita in state s in year t, and $p_{s,t}$ represents the population density of state s in year t.

$$m_{s,t} = \beta_0 + \beta_1 y_{s,t} + \beta_2 y_{s,t}^2 + \beta_3 y_{s,t}^3 + \beta_4 p_{s,t} + \varepsilon_{i,t} \tag{1}$$

This model serves as a baseline model for testing the relationship between GDP per capita and GHG emissions. However, as we have already discussed, it is imperative to control for spatiotemporal effects. To do this, we first add fixed effects for the year with addition of the variable δ_t.

$$m_{s,t} = \beta_0 + \beta_1 y_{s,t} + \beta_2 y_{s,t}^2 + \beta_3 y_{s,t}^3 + \beta_4 p_{s,t} + \delta_t + \varepsilon_{i,t} \tag{2}$$

We then specify a third model with fixed effects for both and state (γ_s) and year (δ_t) as follows, and the final model uses the year fixed effect but a random effect for states.

$$m_{s,t} = \beta_0 + \beta_1 y_{s,t} + \beta_2 y_{s,t}^2 + \beta_3 y_{s,t}^3 + \beta_4 p_{s,t} + \gamma_s + \delta_t + \varepsilon_{i,t} \tag{3}$$

Because the cubic term for GDP per capita is not always included in the model, and because the smaller subsectors might be more sensitive to changes in this term, we repeat all four models without the cubic term, i.e., by removing $\beta_3 y_{s,t}^3$ from each regression equation. In fact, Hüttler et al. (1998) described EKC models which include a significant cubic term as a variant of the EKC model in which the phenomenon of GHG emissions reducing with increases in income is representative of a mere transitory phenomenon. This can be seen quite easily by realizing that the cubic term

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2019 doi:10.20944/preprints201905.0058.v1

eventually leads to (particularly massive) growth in emissions volume, or what is referred to in the EKC literature as an 'N' shaped curve. Thus, including the cubic term can be seen as a sort of check on the EKC hypothesis, for if the cubic model outperforms the standard EKC model without a cubic term, then any evidence for the EKC hypothesis should be suspect at best.

After estimating each model, the estimated coefficients are used to find the point at which GHG emissions peak according to the EKC hypothesis. This is done by using the parameters β_1 and β_2 by checking that $\beta_1 > 0$ and $\beta_2 < 0$. If so, the estimated peak, or turning point as it is often called, is given by $\frac{-\beta_1}{2\beta_2}$. Cubic terms are ignored in these calculations since they are expected not to be significantly different from zero under the EKC hypothesis. We mention now that for the models with effects for year and state, it was indeed the case that these cubic terms were not significantly different from zero for both the energy sector as a whole, and for almost all subsectors including the big two subsectors (electric power generation and transportation).

5 Results

Since we are analyzing results from not only the energy sector as whole, but for six subsectors of this sector, we break this section down into seven subsections, one for the energy production sector and one for each of its six subsectors. In each subsection we will present the estimates from the regression models, provide the estimated turning point ($\frac{-\beta_1}{2\beta_2}$) representing an estimated peak emissions level, and discuss whether the results fit the EKC hypothesis.

Overall, we do find that several models support the EKC hypothesis, at least to some extent, in particular the models which include fixed effects for year or fixed effects for both year and state. The best models, in terms of fitting the data (adjusted R^2), are consistently the models which include fixed effects for both year and state (Equation 3), specifically the variant which did not include a cubic term for GDP per capita (though the inclusion of the cubic term typically did not negatively impact accuracy as this term was rarely significant, even at the 10% level).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2019 doi:10.20944/preprints201905.0058.v1

Additionally, for those models supporting the EKC hypothesis, we find in most cases reasonable estimated turning points. These turning points are typically at income levels slightly higher than what is observed for the United States in its entirety, but are often already realized or imminently realizable by higher income states. The preferred model (the model with fixed effects for year and state which did not include a cubic term) consistently predicted a higher turning point, indicating that models lacking sufficient control variables may under estimate turning points and provide unsupported evidence for the EKC hypothesis.

5.1 The Energy Production Sector

We begin by examining the US energy production sector as a whole. The data set used for this study broke down the energy production sector into six disjoint subsectors: electric power, transportation, industrial, residential, commercial, and fugitive emissions. An initial survey of the estimates from the regression models, presented below in Tables 2 and 3, yields several implications for the EKC hypothesis. First, the fixed effects model specified by Equation 3 which includes fixed effects for both year and state is by far the most accurate model, either with or without the inclusion of the cubic term for GDP per capita. Evidence for the EKC hypothesis was almost exclusively lended by the regression models which did not include cubic terms for GDP per capita, with the lone exception being the full fixed effects model given by Equation 3. Turning points (measured in 1997 USD) for GHG emissions ranged from 35,481 to 49,239. These results indicate that, with a handful of exceptions at the state level for the low end estimated turning points, we should not expect to have seen a peak in GHG emissions from energy productions, and, considering that this source on average contributed 88% of total US GHG emissions, we are yet to have seen a peak in total US GHG emissions.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2019 doi:10.20944/preprints201905.0058.v1

Table 2: Estimation Results for Energy Sector Baseline and Year Fixed Effects Models

Variable	No Fixed Effects With Cubic Term	No Fixed Effects Without Cubic Term	Year Fixed Effects With Cubic Term	Year Fixed Effects Without Cubic Term
Constant	1.928e-06*	-132.6671***	1.6168*	-195.1998***
	(9.97e-07)	(42.955)	(0.933)	(54.979)
GDP per capita	0.0019	0.0136***	-0.0003	0.0175***
	(0.001)	(0.002)	(0.002)	(0.003)
(GDP per capita)^2	1.47e-07**	-1.769e-07***	2.443e-07***	-2.17e-07***
	(6.59e-08)	(3.06e-08)	(8.55e-08)	(3.64e-08)
(GDP per capita)^3	-2.808e-12***	—	-3.811e-12***	—
	(7.8e-13)	—	(9.6e-13)	—
Population Density	0.0870*	0.0906*	0.0666	0.0636
	(0.048)	(0.048)	(0.050)	(0.051)
Adjusted R^2	0.034	0.031	0.037	0.016
Turning Point	—	38,440	—	40,755

Table 3: Estimation Results for Energy Sector Fixed Effects and Random Effects Models

Variable	State Fixed Effects With Cubic Term	State Fixed Effects Without Cubic Term	State Random Effects With Cubic Term	State Random Effects Without Cubic Term
Constant	0.5616	274.9480***	43.385*	23.627
	(0.541)	(11.301)	(23.559)	(19.272)
GDP per capita	0.0064***	0.0033***	0.002	0.003***
	(0.001)	(0.000)	(0.001)	(0.000)
(GDP per capita)^2	-9.019e-08***	-3.351e-08***	0.000	-0.000***
	(3.21e-08)	(3.8e-09)	(0.000)	(0.000)
(GDP per capita)^3	1.377e-13	—	-0.000	—
	(3.81e-13)	—	(0.000)	—
Population Density	0.0895***	0.4419***	0.386***	0.401***
	(0.017)	(0.079)	(0.076)	(0.075)
Adjusted R^2	0.888	0.995	—	—
Turning Point	35,481	49,239	—	—

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2019 doi:10.20944/preprints201905.0058.v1

5.2 Electric Power Subsector

The largest subsector of the energy production sector by volume of emissions is electric power generation which accounts for more than 36% of total energy sector emissions. Collectively, the various models tested actually lend more support for the EKC hypothesis for the electric power generation subsector than for the energy production sector as a whole. In fact, all but one model met the minimum requirement to lend evidence supporting the EKC hypothesis, with the only exception being the baseline model with a cubic term for GDP per capita. Again it seems the case that the fixed effects model specified by Equation 3 fits the data the best. Moreover, the particular variant which excludes the cubic term for GDP per capita fits the data better than any other model. Estimated turning points for emissions from this subsector range from \$20,357 to \$44,865. The lowest estimated turning point indicates that it is possible that emissions from the generation of electric power have peaked in the US, though the estimated turning point from the preferred model, the model with year and state fixed effects and no cubic term for GDP per capita, indicates that emissions have likely not peaked yet.

Table 4: Estimation Results for Electric Power Subsector Baseline and Year Fixed Effects Models

Variable	No Fixed Effects With Cubic Term	No Fixed Effects Without Cubic Term	Year Fixed Effects With Cubic Term	Year Fixed Effects Without Cubic Term
Constant	-3.817e-08	-41.4537***	-0.6497**	-11.6901
	(3.4e-07)	(14.703)	(0.319)	(18.763)
GDP per capita	0.0016***	0.0051***	0.0027***	0.0033***
	(0.000)	(0.001)	(0.001)	(0.001)
(GDP per capita)^2	1.895e-08	-7.066e-08***	-3.009e-08	-5.411e-08***
	(2.26e-08)	(1.05e-08)	(2.93e-08)	(1.24e-08)
(GDP per capita)^3	-7.388e-13***	—	-2.441e-13	—
	(2.68e-13)	—	(3.28e-13)	—
Population Density	-0.0084	-0.0075	0.0030	0.0069
	(0.017)	(0.017)	(0.017)	(0.017)
Adjusted R^2	0.042	0.042	0.050	0.034
Turning Point	—	36,088	44,865	30,493

 doi:10.20944/preprints201905.0058.v1

Table 5: Estimation Results for Electric Power Subsector Fixed Effects and Random Effects Models

Variable	State Fixed Effects With Cubic Term	State Fixed Effects Without Cubic Term	State Random Effects With Cubic Term	State Random Effects Without Cubic Term
Constant	-1.9860***	-5.3015	9.809	3.829
	(0.077)	(6.406)	(11.006)	(7.920)
GDP per capita	0.0041***	0.0015***	0.001	0.001***
	(9.18e-05)	(0.000)	(0.001)	(0.000)
(GDP per capita)^2	-1.007e-07***	-1.917e-08***	-0.000	-0.000***
	(4.55e-09)	(2.15e-09)	(0.000)	(0.000)
(GDP per capita)^3	6.54e-13***	—	-0.000	—
	(5.39e-14)	—	(0.000)	—
Population Density	0.0003	0.2279***	0.164***	0.168***
	(0.002)	(0.045)	(0.041)	(0.040)
Adjusted R^2	0.981	0.987	—	—
Turning Point	20,357	39,124	—	—

5.3 Transportation Subsector

The second largest subsector is transportation, accounting for over 32% of total emissions from the US energy production sector. Transportation is relatively elastic in how it is consumed when compared to the other subsectors. For this reason, transportation is a common target of emissions mitigation policies, legislation, and technologies (Fox et al., 2017). Furthermore, emissions from transportation often occur in heavily populated areas. These emissions contribute directly to numerous negative health consequences (Kagawa, 2002), further incentivizing reductions in emissions from this subsector.

Results from our models for the transportation subsector indicate initial support for the EKC hypothesis in different forms dependent upon whether or not the cubic term for GDP per capita is not included in the regression model. When no cubic term for GDP per capita is included, we find evidence supporting the standard EKC hypothesis, but when a cubic term is included, we see an

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2019 doi:10.20944/preprints201905.0058.v1

'N' shaped model due to the (statistically significant even at the 1% level) effects of the cubic term. While this appears to lend some support to the EKC hypothesis over what could be considered a realistic income domain, Canas et al. (2003) pointed out that the 'N' shaped curve can perhaps be better thought of as an 'augmented' EKC. Hüttler et al. (1998) went further and asserted that the 'N' shaped curve supports not the EKC hypothesis, but a perverted variant in which emissions reduction with economic growth is merely a transitory phase that precedes a final phase in which further economic growth drives a rapid growth in emissions. Because of this, Canas et al. (2003) recommended that any inference of support for EKC models relying on significant effects from the cubic term should be taken cautiously.

Estimated turning points range from \$39,962 to \$56,117, an interval which, at its lowest end, includes only the most wealthy of states by currently realized income. The upper end of this interval far exceeds any current state level values for GDP per capita.

Table 6: Estimation Results for Transportation Subsector Baseline and Year Fixed Effects Models

Variable	No Fixed Effects With Cubic Term	No Fixed Effects Without Cubic Term	Year Fixed Effects With Cubic Term	Year Fixed Effects Without Cubic Term
Constant	1.5649e-06***	-62.0515***	0.8338***	-94.2633***
	(3.39e-07)	(14.720)	(0.319)	(18.791)
GDP per capita	-0.0004	0.0051***	-0.0015**	0.0071***
	(0.000)	(0.001)	(0.001)	(0.001)
(GDP per capita)^2	8.895e-08***	-6.381e-08***	1.939e-07***	-8.451e-08***
	(2.26e-08)	(1.05e-08)	(2.92e-08)	(1.24e-08)
(GDP per capita)^3	–1.329e-12***	—	-1.848e-12***	—
	(2.67e-13)	—	(3.28e-13)	—
Population Density	0.0776***	0.0793***	0.0670***	0.0653***
	(0.017)	(0.017)	(0.017)	(0.017)
Adjusted R^2	0.065	0.059	0.072	0.050
Turning Point	—	39,962	—	42,007

 doi:10.20944/preprints201905.0058.v1

Table 7: Estimation Results for Transportation Subsector Fixed Effects and Random Effects Models

Variable	State Fixed Effects With Cubic Term	State Fixed Effects Without Cubic Term	State Random Effects With Cubic Term	State Random Effects Without Cubic Term
Constant	0.2030	163.6813***	25.867***	-4.687
	(0.357)	(4.905)	(9.353)	(7.434)
GDP per capita	0.0009**	0.0011***	-0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
(GDP per capita)^2	2.099e-08	-9.801e-09***	0.000***	-0.000***
	(2.12e-08)	(1.65e-09)	(0.000)	(0.000)
(GDP per capita)^3	-4.74e-13*	—	-0.000***	—
	(2.52e-13)	—	(0.000)	—
Population Density	0.0772***	0.3603***	0.296***	0.318***
	(0.011)	(0.034)	(0.032)	(0.033)
Adjusted R^2	0.595	0.993	—	—
Turning Point	—	56,117	—	—

5.4 Industrial Subsector

The third largest subsector is the industrial subsector. Emissions from this subsector account for nearly 19% of total emissions from US energy production, and the top three subsector emission sources combine for nearly 90% of this total.

The regression models yielded rather interesting results for this subsector, as the models with fixed effects for both year and state, while accurate within the data set itself, did not find GDP per capita to be a significant factor in predicting GHG emissions unless the cubic term was included. If we exclude the turning point for the model with year and state fixed effects and no cubic term for GDP per capita[1], estimated turning points ranged from \$29,806 to \$44,550. These values range from below the current national GDP per capita to an income level just beyond the current wealthiest state.

[1]This model estimated a turning point at \$263,922 but also did not find GDP per capita to be a significant factor for GHG emissions (even at the 10% level), so we exclude it from our discussion of the results.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2019 doi:10.20944/preprints201905.0058.v1

Table 8: Estimation Results for Industrial Subsector Baseline and Year Fixed Effects Models

Variable	No Fixed Effects With Cubic Term	No Fixed Effects Without Cubic Term	Year Fixed Effects With Cubic Term	Year Fixed Effects Without Cubic Term
Constant	-5.645e-07*	-20.1783	0.9198***	-60.0081***
	(3.33e-07)	(14.412)	(0.312)	(18.350)
GDP per capita	0.0005	0.0023***	-0.0010	0.0047***
	(0.000)	(0.001)	(0.001)	(0.001)
(GDP per capita)^2	2.47e-08	-2.889e-08***	8.71e-08***	-5.275e-08***
	(2.21e-08)	(1.03e-08)	(2.86e-08)	(1.21e-08)
(GDP per capita)^3	-4.79e-13*	—	-1.114e-12***	—
	(2.62e-13)	—	(3.21e-13)	—
Population Density	-0.0295*	-0.0289*	-0.0433***	-0.0472***
	(0.016)	(0.016)	(0.017)	(0.017)
Adjusted R^2	0.009	0.008	0.020	0.004
Turning Point	—	29,806	—	44,550

Table 9: Estimation Results for Industrial Subsector State Level Fixed Effects and Random Effects Models

Variable	State Fixed Effects With Cubic Term	State Fixed Effects Without Cubic Term	State Random Effects With Cubic Term	State Random Effects Without Cubic Term
Constant	1.0906***	78.0170***	10.596	22.096***
	(0.117)	(5.837)	(10.022)	(4.247)
GDP per capita	0.0011***	0.0002	0.001**	0.000
	(0.000)	(0.000)	(0.001)	(0.000)
(GDP per capita)^2	-1.429e-08**	-3.789e-10	-0.000*	-0.000
	(6.93e-09)	(1.96e-09)	(0.000)	(0.000)
(GDP per capita)^3	5.442e-14	—	0.000*	—
	(8.21e-14)	—	(0.000)	—
Population Density	-0.0331***	-0.1359***	-0.103***	-0.111***
	(0.004)	(0.041)	(0.036)	(0.036)
Adjusted R^2	0.953	0.988	—	—
Turning Point	38,488	263,922	—	—

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2019 doi:10.20944/preprints201905.0058.v1

5.5 Residential Subsector

The remaining subsectors are all relatively small with the residential subsector responsible for only about 6% of total US GHG emissions from energy production. We find limited evidence at best for the EKC hypothesis in the residential subsector from our models, primarily because the presence of fixed effects erodes belief in the EKC hypothesis for this subsector. The models with fixed effects for year and state were yet again the most accurate within the data set, however, these models did not support (even at the 10% level) the EKC hypothesis. Indeed, only the baseline model with no cubic term for GDP per capita provided support for the EKC hypothesis in this subsector. Estimated turning points ranged from $38,204 to $81,954, a range similar to the other subsectors and to the energy production sector as a whole, but we eschew these in light of the lack of support for the EKC hypothesis in the residential subsector.

Table 10: Estimation Results for Residential Subsector Baseline and Year Fixed Effects Models

Variable	No Fixed Effects With Cubic Term	No Fixed Effects Without Cubic Term	Year Fixed Effects With Cubic Term	Year Fixed Effects Without Cubic Term
Constant	8.082e-07***	-10.3990***	0.2958***	-22.9455***
	(6.47e-08)	(2.805)	(0.060)	(3.531)
GDP per capita	-0.0001	0.0008***	-0.0005***	0.0016***
	(8.79e-05)	(0.000)	(0.000)	(0.000)
(GDP per capita)^2	1.542e-08***	-1.047e-08***	3.487e-08***	-1.792e-08***
	(4.3e-09)	(2e-09)	(5.51e-09)	(2.34e-09)
(GDP per capita)^3	-2.261e-13***	—	-4.244e-13***	—
	(5.09e-14)	—	(6.18e-14)	—
Population Density	0.0400***	0.0403***	0.0357***	0.0347***
	(0.003)	(0.003)	(0.003)	(0.003)
Adjusted R^2	0.168	0.164	0.193	0.179
Turning Point	—	38,204	—	44,643

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2019 doi:10.20944/preprints201905.0058.v1

Table 11: Estimation Results for Residential Subsector State Level Fixed Effects and Random Effects Models

Variable	State Fixed Effects With Cubic Term	State Fixed Effects Without Cubic Term	State Random Effects With Cubic Term	State Random Effects Without Cubic Term
Constant	0.3073***	25.9935***	1.888	3.714**
	(0.043)	(1.119)	(1.960)	(1.437)
GDP per capita	0.0001**	0.0001***	0.000**	0.000***
	(5.12e-05)	(3.86e-05)	(0.000)	(0.000)
(GDP per capita)^2	2.784e-09	-6.101e-10	-0.000*	-0.000**
	(2.53e-09)	(3.76e-10)	(0.000)	(0.000)
(GDP per capita)^3	-5.493e-14*	—	0.000	—
	(3e-14)	—	(0.000)	—
Population Density	0.0390***	0.0016	0.011	0.010
	(0.001)	(0.008)	(0.007)	(0.007)
Adjusted R^2	0.859	0.990	—	—
Turning Point	—	81,954	—	—

5.6 Commercial Subsector

The commercial subsector represents less than 4% of total GHG emissions from US energy production. The models tested for the commercial subsector actually lend quite a bit of support to the EKC hypothesis. The models which did not provide estimated turning points did so because the signs of β_1 and β_2 were reversed. However, in each such instance the presence of a significant cubic term with negative estimated coefficient β_3 leads to a 'backwards N' shaped EKC. Theoretically, this amounts to perhaps an overly emphatic EKC where emissions eventually continue to decrease quite rapidly with economic growth until production contributes non-positive net emissions. Idealistically, this can be interpreted as evidence that emissions mitigation measures and cleaner (perhaps even negative net emissions) technologies are gaining traction in the sector, though realistically this observed effect is an exaggeration due to the lack of data (recall that the data set spans only 22 years). Turning points under the EKC hypothesis were very consistent and ranged from $40,984 to $45,620.

Table 12: Estimation Results for Commercial Subsector Baseline and Year Fixed Effects Models

Variable	No Fixed Effects With Cubic Term	No Fixed Effects Without Cubic Term	Year Fixed Effects With Cubic Term	Year Fixed Effects Without Cubic Term
Constant	4.641e-07***	-6.0905***	0.1889***	-14.0608***
	(4.0e-08)	(1.753)	(0.038)	(2.205)
GDP per capita	-5.93e-05	0.0005***	-0.0003***	0.0010***
	(5.49e-05)	(9.52e-05)	(7.34e-05)	(0.000)
(GDP per capita)^2	93469e-09***	-6.01e-09***	2.1185e-08***	-1.096e-08***
	(2.69e-09)	(1.25e-09)	(3.44e-09)	(1.46e-09)
(GDP per capita)^3	-1.373e-13***	—	-2.636e-13***	—
	(3.18e-14)	—	(3.86e-14)	—
Population Density	0.0230***	0.0232***	0.0202***	0.0195***
	(0.002)	(0.002)	(0.002)	(0.002)
Adjusted R^2	0.153	0.148	0.178	0.165
Turning Point	—	40,984	—	45,620

Table 13: Estimation Results for Commercial Subsector State Level Fixed Effects and Random Effects Models

Variable	State Fixed Effects With Cubic Term	State Fixed Effects Without Cubic Term	State Random Effects With Cubic Term	State Random Effects Without Cubic Term
Constant	0.3191***	13.8771***	0.560	2.169**
	(0.023)	(0.903)	(1.449)	(0.987)
GDP per capita	4.175e-05	8.764e-05***	0.000**	0.000***
	(2.73e-05)	(3.11e-05)	(0.000)	(0.000)
(GDP per capita)^2	2.979e-09**	-9.73e-10***	-0.000**	-0.000***
	(1.35e-09)	(3.03e-10)	(0.000)	(0.000)
(GDP per capita)^3	-4.579e-14***	—	0.000	—
	(1.6e-14)	—	(0.000)	—
Population Density	0.0222***	0.0036	0.011**	0.010*
	(0.001)	(0.006)	(0.005)	(0.005)
Adjusted R^2	0.895	0.984	—	—
Turning Point	—	45,036	—	—

5.7 Fugitive Emissions

The final subsector is fugitive emissions. These are emissions due to various equipment malfunctions causing leakages. Because these malfunctions often occur prior to the emissions reaching their final, cleanest state, fugitive emissions often contribute excessively to air pollution and represent a major health concern (Hassim et al., 2010). Estimated results from the regression models indicated limited support for the EKC hypothesis for fugitive emissions. However, as legislation on emissions mitigation and equipment status are not uniformly enforced, state level effects should be included in any meaningful model for fugitive emissions. The models which included state level effects all indicated support for the EKC hypothesis for fugitive emissions. The favored model, the model with year and state fixed effects and no cubic term for GDP per capita, was especially supportive. Estimated turning points ranged from \$17,190 to \$58,185. Altogether, these results indicate that fugitive emissions should have peaked or should be peaking in the very near future, and that enforcing requirements on the status of emissions producing equipment does indeed help to reduce fugitive emissions.

Table 14: Estimation Results for Fugitive Emissions Baseline and Year Fixed Effects Models

Variable	No Fixed Effects With Cubic Term	No Fixed Effects Without Cubic Term	Year Fixed Effects With Cubic Term	Year Fixed Effects Without Cubic Term
Constant	-2.905e-07***	7.5058***	0.0282	7.7681***
	(4.89e-08)	(2.110)	(0.046)	(2.704)
GDP per capita	0.0004***	-0.0002*	0.0003***	-0.0002
	(6.65e-05)	(0.000)	(9.01e-05)	(0.000)
(GDP per capita)^2	-1.05e-08***	3.069e-09**	-8.726e-09**	3.211e-09*
	(3.25e-09)	(1.5e-09)	(4.22e-09)	(1.79e-09)
(GDP per capita)^3	1.02e-13***	—	8.3e-14*	—
	(3.85e-14)	—	(4.74e-14)	—
Population Density	-0.0157***	-0.0158***	-0.0160***	-0.0156***
	(0.002)	(0.002)	(0.002)	(0.002)
Adjusted R^2	0.035	0.040	0.037	0.023
Turning Point	19,048	—	17,190	—

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2019 doi:10.20944/preprints201905.0058.v1

Table 15: Estimation Results for Fugitive Emissions State Level Fixed Effects and Random Effects Models

Variable	State Fixed Effects With Cubic Term	State Fixed Effects Without Cubic Term	State Random Effects With Cubic Term	State Random Effects Without Cubic Term
Constant	0.6276***	-1.3194	-2.409	-1.602
	(0.018)	(1.886)	(2.664)	(1.563)
GDP per capita	0.0002***	0.0003***	0.000**	0.000***
	(2.2e-05)	(6.5e-05)	(0.000)	(0.000)
(GDP per capita)^2	-1.944e-09*	-2.578e-09***	-0.000	-0.000***
	(1.09e-09)	(6.34e-10)	(0.000)	(0.000)
(GDP per capita)^3	4.147e-15	—	0.000	—
	(1.29e-14)	—	(0.000)	—
Population Density	-0.0161***	-0.0157	-0.017**	-0.018**
	(0.001)	(0.013)	(0.008)	(0.008)
Adjusted R^2	0.947	0.945	—	—
Turning Point	51,440	58,185	—	—

6 Forecasting GHG Emissions

Using the results from our series of EKC regression models, we forecast GHG emissions from the US energy production sector using the forecasting method used by Selden and Song (1994). We then complement these forecasts using aggregated subsector emissions forecasts following the same methodology. This dual forecasting approach allows us to achieve several novel results. First, it allows for us to forecast GHG emissions at the subsector level. Additionally, this allows for state level forecasts of emissions at the subsector level. But most importantly, we can compare aggregated forecasted subsector emissions against forecasted emissions from the energy sector as a whole. By doing so, we can potentially improve upon forecasts of emissions from the energy production sector.

In order to justify this forecasting approach, we first present a comparison of the in-sample accuracy of the models for the energy sector as a whole versus the accuracy of aggregated subsector models. We find that, in accordance with the 'sum of the parts is more than the whole' approach espoused by Ferreira and Santa-Clara (2011), the aggregated subsector models are more accurate than the energy sector models. The following table contains the comparisons between the Adjusted R^2 values for the independent energy sector models and the aggregated energy subsector models within the energy sector data.

Table 16: Comparison of Adjusted R^2 Values

	Adjusted R^2		
Model	Independent	Aggregate	Difference
Baseline	0.0355	0.0339	1.61e-3
Baseline Without Cubic Term	0.0309	0.0309	2.11e-12
Year Fixed Effects	0.0382	0.0373	9.27e-4
Year Fixed Effects Without Cubic Term	0.0165	0.0165	2.39e-12
Year and State Fixed Effects	0.8893	0.8878	1.45e-3
Year and State Fixed Effects Without Cubic Term	0.9952	0.9952	2.00e-11

Since the EKC hypothesis asserts that GHG emissions are a function of economic growth, we can forecast GHG emissions using the estimated parameters from our regression models and forecasted economic growth. Because economic growth in the United States is relatively consistent, we can forecast GDP per capita on a state by state basis with the following model specified by Selden and Song (1994) in which $y_{i,t}$ represents GDP per capita in state s in year t, and $\theta_{s,t}$ denotes the error term.

$$\ln\left(\frac{y_{s,t}}{y_{s,t-1}}\right) = \gamma_0 + \gamma_1 \ln\left(y_{s,t-1}\right) + \gamma_2 \ln\left(y_{s,t-1}^2\right) + \theta_{s,t} \tag{4}$$

Because we used population density as a covariate in our regression models, and because pop-

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2019 doi:10.20944/preprints201905.0058.v1

ulation density can be easily forecasted, we include forecasted population density in our forecasts using a linear trend on a state by state basis. We then use our forecasted GDP per capita and population densities along with the estimated parameters from each of our models to create a raw forecast of emissions for each model and (sub)sector combination over a 100 year horizon (from 2012 through 2111).

These raw emissions are then transformed using a Tobit functional form. This approach flattens forecasts, creating more gradual increases and decreases. Selden and Song (1994) preferred this method for asymptotic properties compared to the standard EKC and because it does not affect the turning points estimated by the EKC model. The Tobit transformation used here is given by the following equation in which $\hat{m}_{s,t}$ denotes the Tobit transformed forecasted emissions for state s in year t, $m_{s,t}$ denotes the raw forecasted emissions for state s in year t, Φ denotes the standard normal cumulative distribution function, ϕ denotes the standard normal probability distribution function, and σ denotes the standard deviation of the error terms in the original regression models from which the forecasts are based.

$$\hat{m}_{s,t} = \Phi\left(\frac{m_{s,t}}{\sigma}\right) * m_{s,t} + \sigma * \phi(m_{s,t}) \tag{5}$$

Once the Tobit transformations are complete, we aggregate the state level forecasts for each model into a single national level forecast of GHG emissions.

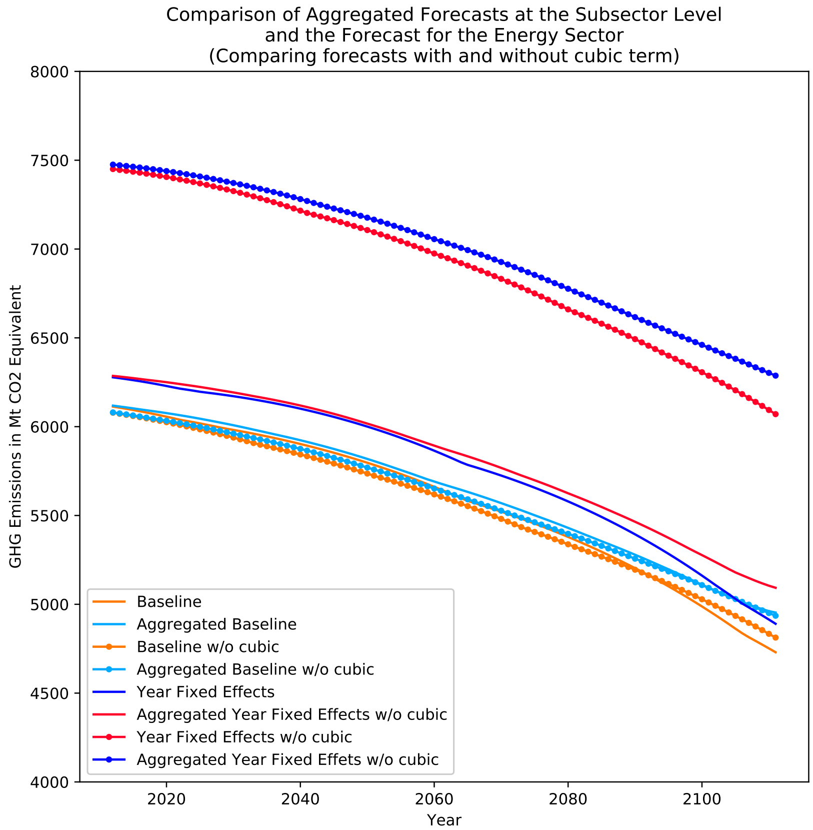
The forecasted models are presented below and are grouped by model. For legibility purposes, the forecasts are separated into two plots. Before discussing the results, we mention that the first plot contains all models which predict that emissions have peaked and the second plot contains models which indicated continued growth in emissions. This distinction is made only for ease of interpreting the plots. In fact, this variance in outcome is indicative of the importance of including some sort of spatiotemporal controls into EKC models.

Throughout all forecasts we see that the aggregated forecasts, which we just demonstrated to be more accurate within the data set, universally predict greater future emissions than the forecasts

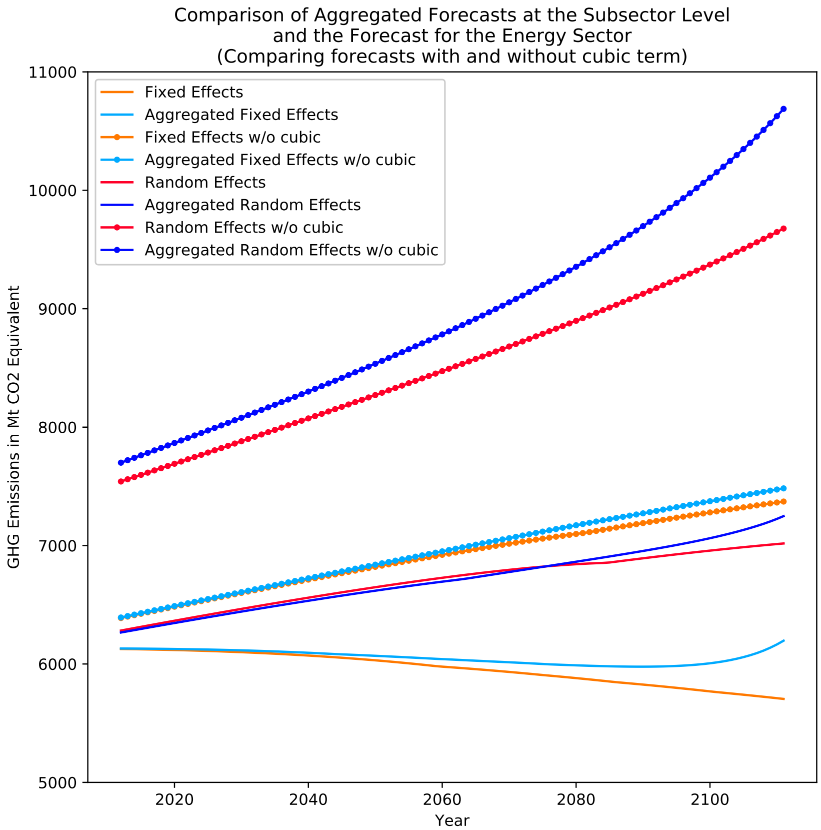
 doi:10.20944/preprints201905.0058.v1

using energy sector level data. This indicates that forecasts based on the EKC are relatively inaccurate; in this case we see that forecasts based on the standard EKC consistently forecast below aggregate forecast levels. Given that the aggregated subsector level models outperform the energy sector level models for in-sample accuracy, we conclude that the 'sum of the part is greater than the whole' concept applies to emissions estimation and forecasting at the sector and subsector level.

Additionally, we see that, even with most of the models supporting the EKC hypothesis, volumetric emissions are forecasted to increase in roughly half of the models, even when aggregating subsector level forecasts. These forecasting results are consistent with those from Selden and Song (1994), indicating that even if economic growth eventually leads to reduced GHG emissions, we should not rely solely on economic growth as a means to reduce existing levels of greenhouse gases in a meaningful time frame.



Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2019 doi:10.20944/preprints201905.0058.v1



7 Conclusion

In this paper we tested the EKC hypothesis for GHG emissions from the US energy production sector and its subsectors using several variants of the standard EKC regression model which models GHG emissions as a function of economic growth. We found that certain models did indeed lend support to the EKC hypothesis, but that these models, whether at the sector or subsector level, were more likely to support the EKC hypothesis if they did not contain a cubic term for GDP per capita. Additionally, the models performed most accurately when fixed effects for year and state were included. It is certainly possible for the EKC hypothesis to be true for several subsectors of a sector of the economy while not being true for the sector as a whole.

After testing the EKC hypothesis against a battery of models in the energy production sector of the US economy and its six subsectors, we forecasted future emissions for the energy production sector both independently and as the aggregate of forecasts of its subsectors. We found that the aggregated forecasts were higher than the independent forecasts. This result was validated with in-sample testing of the accuracy of these two approaches.

To extend this research, we offer two immediate opportunities. First, we only tested the EKC hypothesis against fairly standard econometric models. Given the novelty, depth, and intricacy of this study, we forewent the use of additional models. Testing the EKC hypothesis using the data set used in this study with greater attention paid to spatiotemporal effects would certainly help improve upon this work. Additionally, this approach of subsector level emissions based forecasting can be applied elsewhere, beyond the United States. In-sample accuracy testing may not offer the same results, however, so any subsector level forecasting should pay special attention to how it is validated.

 doi:10.20944/preprints201905.0058.v1

References

Al-Mulali, U., Saboori, B., and Ozturk, I. (2015). Investigating the environmental kuznets curve hypothesis in vietnam. *Energy Policy*, 76:123–131.

Aldy, J. E. (2005). An environmental kuznets curve analysis of us state-level carbon dioxide emissions. *The Journal of Environment & Development*, 14(1):48–72.

Aldy, J. E. (2006). Per capita carbon dioxide emissions: convergence or divergence? *Environmental and Resource Economics*, 33(4):533–555.

Aldy, J. E. (2007). Divergence in state-level per capita carbon dioxide emissions. *Land Economics*, 83(3):353–369.

Apergis, N., Christou, C., and Gupta, R. (2017). Are there environmental kuznets curves for us state-level co2 emissions? *Renewable and Sustainable Energy Reviews*, 69:551–558.

Auffhammer, M. and Carson, R. T. (2008). Forecasting the path of china's co2 emissions using province-level information. *Journal of Environmental Economics and Management*, 55(3):229–247.

Auffhammer, M. and Steinhauser, R. (2012). Forecasting the path of us co2 emissions using state-level information. *Review of Economics and Statistics*, 94(1):172–185.

Azomahou, T., Laisney, F., and Van, P. N. (2006). Economic development and co2 emissions: a nonparametric panel approach. *Journal of Public Economics*, 90(6-7):1347–1363.

Bruvoll, A. and Medin, H. (2003). Factors behind the environmental kuznets curve. a decomposition of the changes in air pollution. *Environmental and Resource Economics*, 24(1):27–48.

Burnett, J. and Bergstrom, J. C. (2010). Us state-level carbon dioxide emissions: a spatial-temporal econometric approach of the environmental kuznets curve. Technical report.

Burnett, J. W. (2016). Club convergence and clustering of us energy-related co2 emissions. *Resource and Energy Economics*, 46:62–84.

Burnett, J. W., Bergstrom, J. C., and Dorfman, J. H. (2013a). A spatial panel data approach to estimating us state-level energy emissions. *Energy Economics*, 40:396–404.

Burnett, J. W., Bergstrom, J. C., and Wetzstein, M. E. (2013b). Carbon dioxide emissions and economic growth in the us. *Journal of Policy Modeling*, 35(6):1014–1028.

Canas, A., Ferrao, P., and Conceicao, P. (2003). A new environmental kuznets curve? relationship between direct material input and income per capita: evidence from industrialised countries. *Ecological Economics*, 46(2):217–229.

Caron, J. and Fally, T. (2018). Per capita income, consumption patterns, and co2 emissions. Technical report, National Bureau of Economic Research.

Chow, G. C. and Li, J. (2014). Environmental kuznets curve: conclusive econometric evidence for co2. *Pacific Economic Review*, 19(1):1–7.

Criado, C. O. and Grether, J.-M. (2011). Convergence in per capita co2 emissions: A robust distributional approach. *Resource and Energy Economics*, 33(3):637–665.

Criado, C. O., Valente, S., and Stengos, T. (2011). Growth and pollution convergence: Theory and evidence. *Journal of Environmental Economics and Management*, 62(2):199–214.

Dinda, S. (2004). Environmental kuznets curve hypothesis: a survey. *Ecological Economics*, 49(4):431–455.

Edelenbosch, O., Kermeli, K., Crijns-Graus, W., Worrell, E., Bibas, R., Fais, B., Fujimori, S., Kyle, P., Sano, F., and van Vuuren, D. (2017). Comparing projections of industrial energy demand and greenhouse gas emissions in long-term energy models. *Energy*, 122:701–710.

Fankhauser, S. and Tol, R. S. (2005). On climate change and economic growth. *Resource and Energy Economics*, 27(1):1–17.

Ferreira, M. A. and Santa-Clara, P. (2011). Forecasting stock market returns: The sum of the parts is more than the whole. *Journal of Financial Economics*, 100(3):514–537.

Fox, J., Axsen, J., and Jaccard, M. (2017). Picking winners: Modelling the costs of technology-specific climate policy in the us passenger vehicle sector. *Ecological Economics*, 137:133–147.

Friedl, B. and Getzner, M. (2003). Determinants of co2 emissions in a small open economy. *Ecological Economics*, 45(1):133–148.

Galeotti, M., Lanza, A., and Pauli, F. (2006). Reassessing the environmental kuznets curve for co2 emissions: A robustness exercise. *Ecological Economics*, 57(1):152–163.

Gill, A. R., Viswanathan, K. K., and Hassan, S. (2018). The environmental kuznets curve (ekc) and the environmental problem of the day. *Renewable and Sustainable Energy Reviews*, 81:1636–1642.

Harbaugh, W. T., Levinson, A., and Wilson, D. M. (2002). Reexamining the empirical evidence for an environmental kuznets curve. *Review of Economics and Statistics*, 84(3):541–551.

Hassim, M. H., Pérez, A. L., and Hurme, M. (2010). Estimation of chemical concentration due to fugitive emissions during chemical process design. *Process Safety and Environmental Protection*, 88(3):173–184.

Hüttler, W., Schandl, H., and Weisz, H. (1998). Are industrial economies on the path of dematerialization? material flow accounts for austria 1960-1996: indicators and international comparison. In *ConAccount workshop Ecologizing Societal Metabolism*, page 23.

Kagawa, J. (2002). Health effects of diesel exhaust emissions—a mixture of air pollutants of worldwide concern. *Toxicology*, 181:349–353.

Kampa, M. and Castanas, E. (2008). Human health effects of air pollution. *Environmental Pollution*, 151(2):362–367.

Konan, D. E. and Chan, H. L. (2010). Greenhouse gas emissions in hawai'i: Household and visitor expenditure analysis. *Energy Economics*, 32(1):210–219.

Kunnas, J. and Myllyntaus, T. (2007). The environmental kuznets curve hypothesis and air pollution in finland. *Scandinavian Economic History Review*, 55(2):101–127.

List, J. A. and Gallet, C. A. (1999). The environmental kuznets curve: does one size fit all? *Ecological Economics*, 31(3):409–423.

Maddison, D. (2006). Environmental kuznets curves: A spatial econometric approach. *Journal of Environmental Economics and Management*, 51(2):218–230.

Magnani, E. (2000). The environmental kuznets curve, environmental protection policy and income distribution. *Ecological Economics*, 32(3):431–443.

Marrero, G. A. (2010). Greenhouse gases emissions, growth and the energy mix in europe. *Energy Economics*, 32(6):1356–1363.

Mazzanti, M., Montini, A., and Zoboli, R. (2008). Environmental kuznets curves for air pollutant emissions in italy: evidence from environmental accounts (namea) panel data. *Economic Systems Research*, 20(3):277–301.

Millimet, D. L., List, J. A., and Stengos, T. (2003). The environmental kuznets curve: Real progress or misspecified models? *Review of Economics and Statistics*, 85(4):1038–1047.

Ordás Criado, C. (2008). Temporal and spatial homogeneity in air pollutants panel ekc estimations. *Environmental and Resource Economics*, 40(2):265–283.

Ravishankara, A., Daniel, J. S., and Portmann, R. W. (2009). Nitrous oxide (n2o): the dominant ozone-depleting substance emitted in the 21st century. *Science*, 326(5949):123–125.

Ringel, M., Laidi, R., and Djenouri, D. (2019). Multiple benefits through smart home energy management solutions—a simulation-based case study of a single-family-house in algeria and germany. *Energies*, 12(8):1537.

Rupasingha, A., Goetz, S. J., Debertin, D. L., and Pagoulatos, A. (2004). The environmental kuznets curve for us counties: A spatial econometric analysis with extensions. *Papers in Regional Science*, 83(2):407–424.

Selden, T. M. and Song, D. (1994). Environmental quality and development: is there a kuznets curve for air pollution emissions? *Journal of Environmental Economics and Management*, 27(2):147–162.

Shittu, W. O., Musibau, H., and Hassan, S. (2018). Revisiting the environmental kuznets curve in malaysia: the interactive roles of deforestation and urbanisation. *International Journal of Green Economics*, 12(3-4):272–293.

Soytas, U. and Sari, R. (2009). Energy consumption, economic growth, and carbon emissions: challenges faced by an eu candidate member. *Ecological Economics*, 68(6):1667–1675.

Soytas, U., Sari, R., and Ewing, B. T. (2007). Energy consumption, income, and carbon emissions in the united states. *Ecological Economics*, 62(3-4):482–489.

Stern, D. I. (2004). The rise and fall of the environmental kuznets curve. *World Development*, 32(8):1419–1439.

Stern, D. I. (2017). The environmental kuznets curve after 25 years. *Journal of Bioeconomics*, 19(1):7–28.

Stern, D. I. and Common, M. S. (2001). Is there an environmental kuznets curve for sulfur? *Journal of Environmental Economics and Management*, 41(2):162–178.

Suri, V. and Chapman, D. (1998). Economic growth, trade and energy: implications for the environmental kuznets curve. *Ecological Economics*, 25(2):195–208.

Tol, R. S., Pacala, S. W., and Socolow, R. (2006). Understanding long-term energy use and carbon dioxide emissions in the usa.

West, J. J., Smith, S. J., Silva, R. A., Naik, V., Zhang, Y., Adelman, Z., Fry, M. M., Anenberg, S., Horowitz, L. W., and Lamarque, J.-F. (2013). Co-benefits of mitigating global greenhouse gas emissions for future air quality and human health. *Nature Climate Change*, 3(10):885.

WRI CAIT Climate Data Explorer (2014). Climate analysis indicators tool: Wri's climate data explorer. *Washington, DC: World Resources Institute*. Available online at: `http://cait.wri.org`. Accessed on April 2, 2019.

Yang, C., McCollum, D., McCarthy, R., and Leighty, W. (2009). Meeting an 80% reduction in greenhouse gas emissions from transportation by 2050: A case study in california. *Transportation Research Part D: Transport and Environment*, 14(3):147–156.

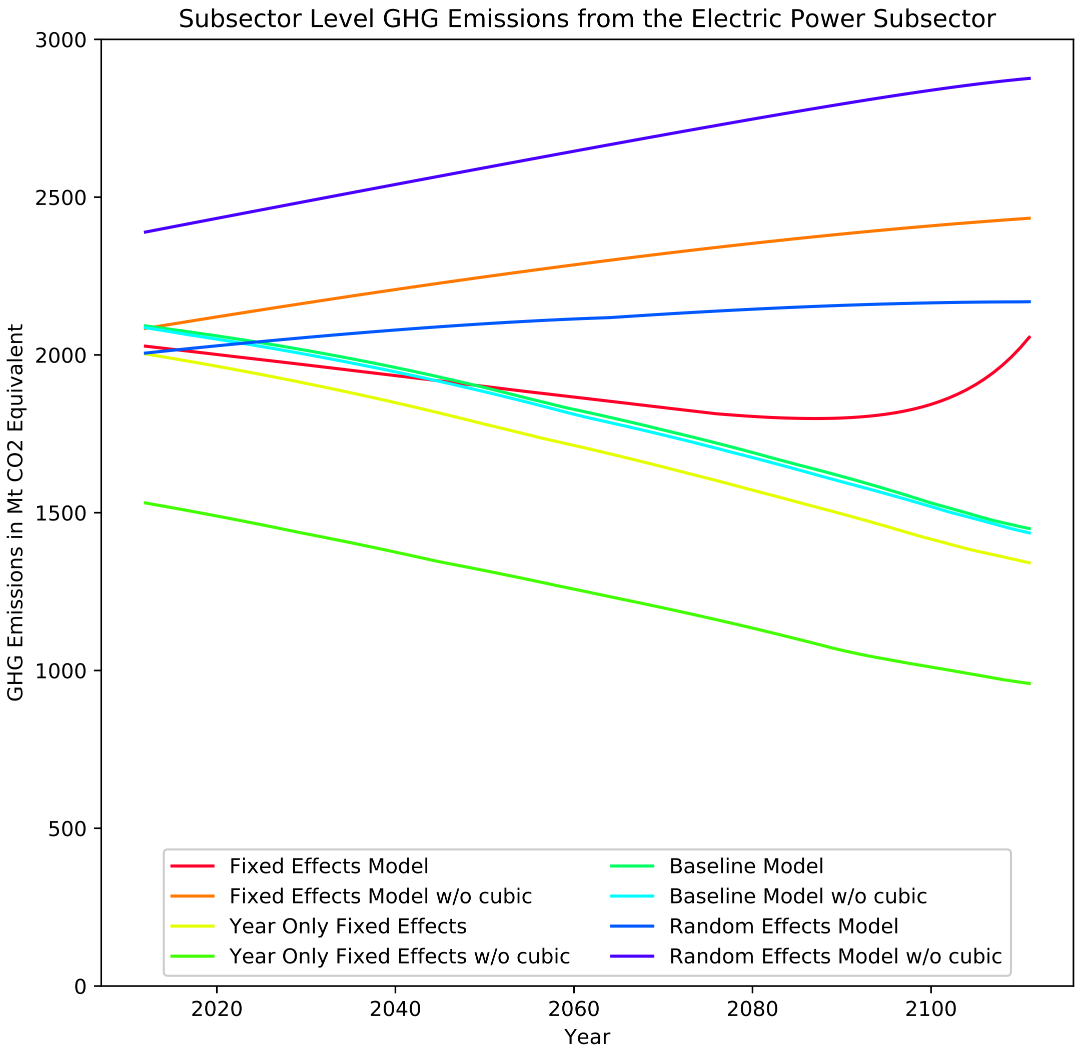
Yates, D. N. and Strzepek, K. M. (1998). An assessment of integrated climate change impacts on the agricultural economy of egypt. *Climatic Change*, 38(3):261–287.

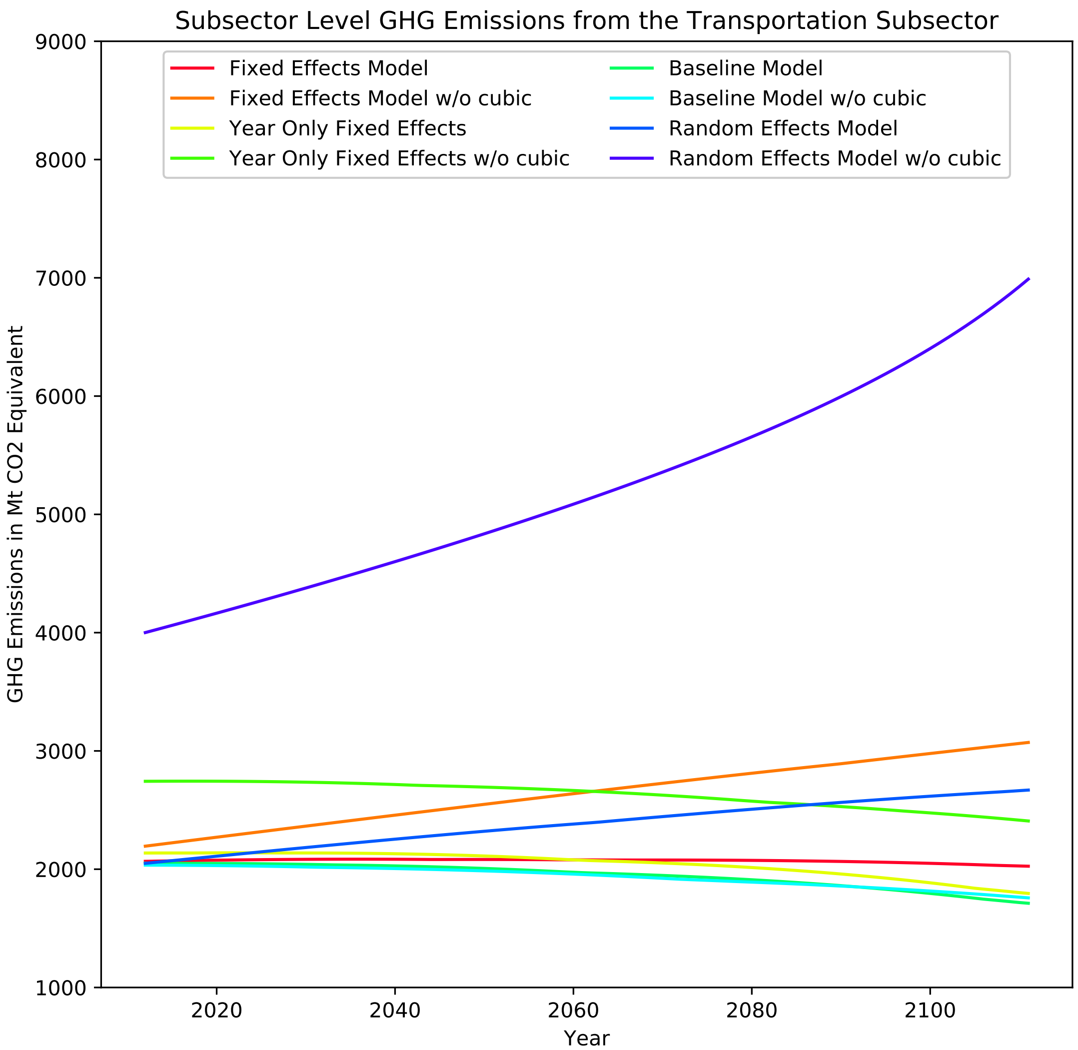
Zhou, Y. and Gurney, K. R. (2011). Spatial relationships of sector-specific fossil fuel co2 emissions in the united states. *Global Biogeochemical Cycles*, 25(3).

Zoundi, Z. (2017). Co2 emissions, renewable energy and the environmental kuznets curve, a panel cointegration approach. *Renewable and Sustainable Energy Reviews*, 72:1067–1075.

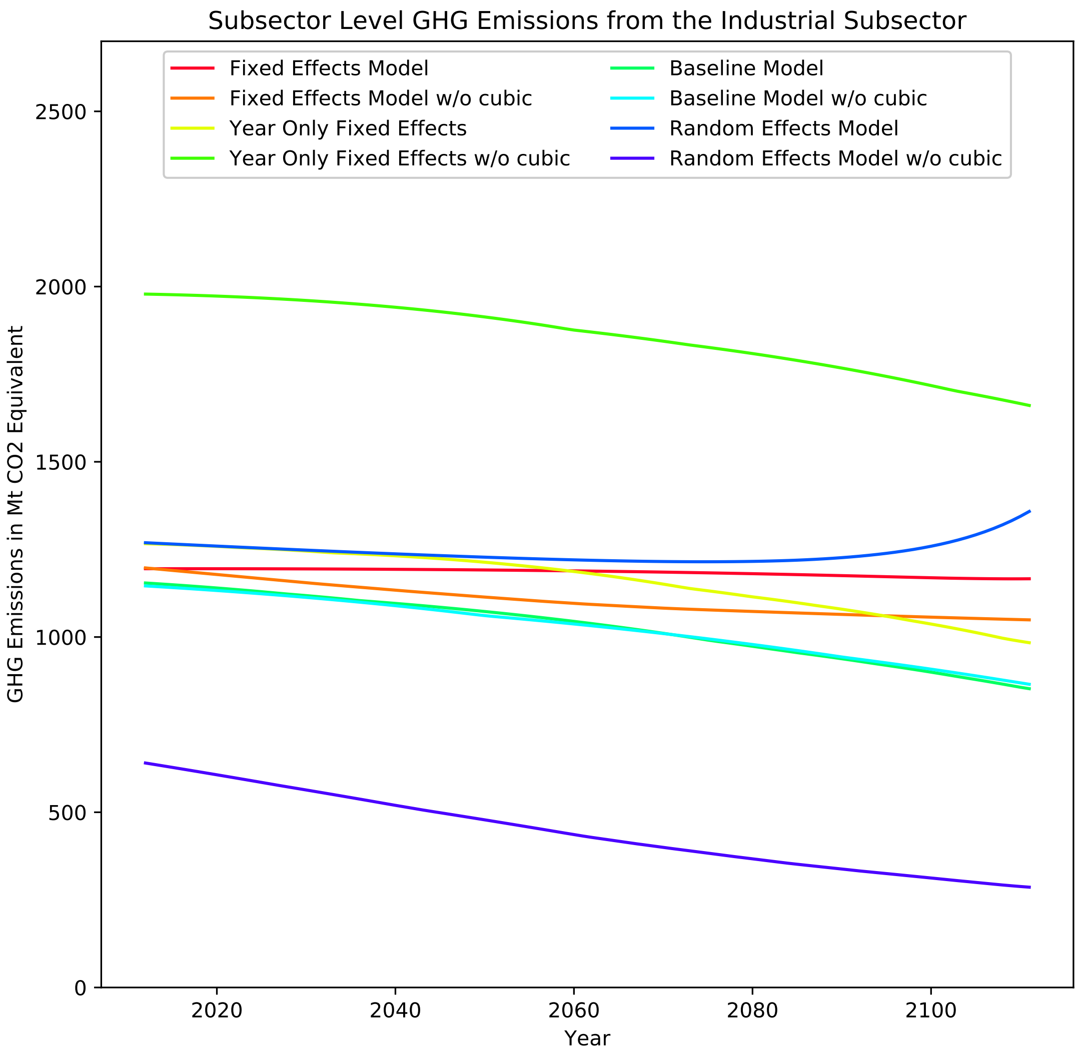
Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2019 doi:10.20944/preprints201905.0058.v1

A Subsector Level Forecasts

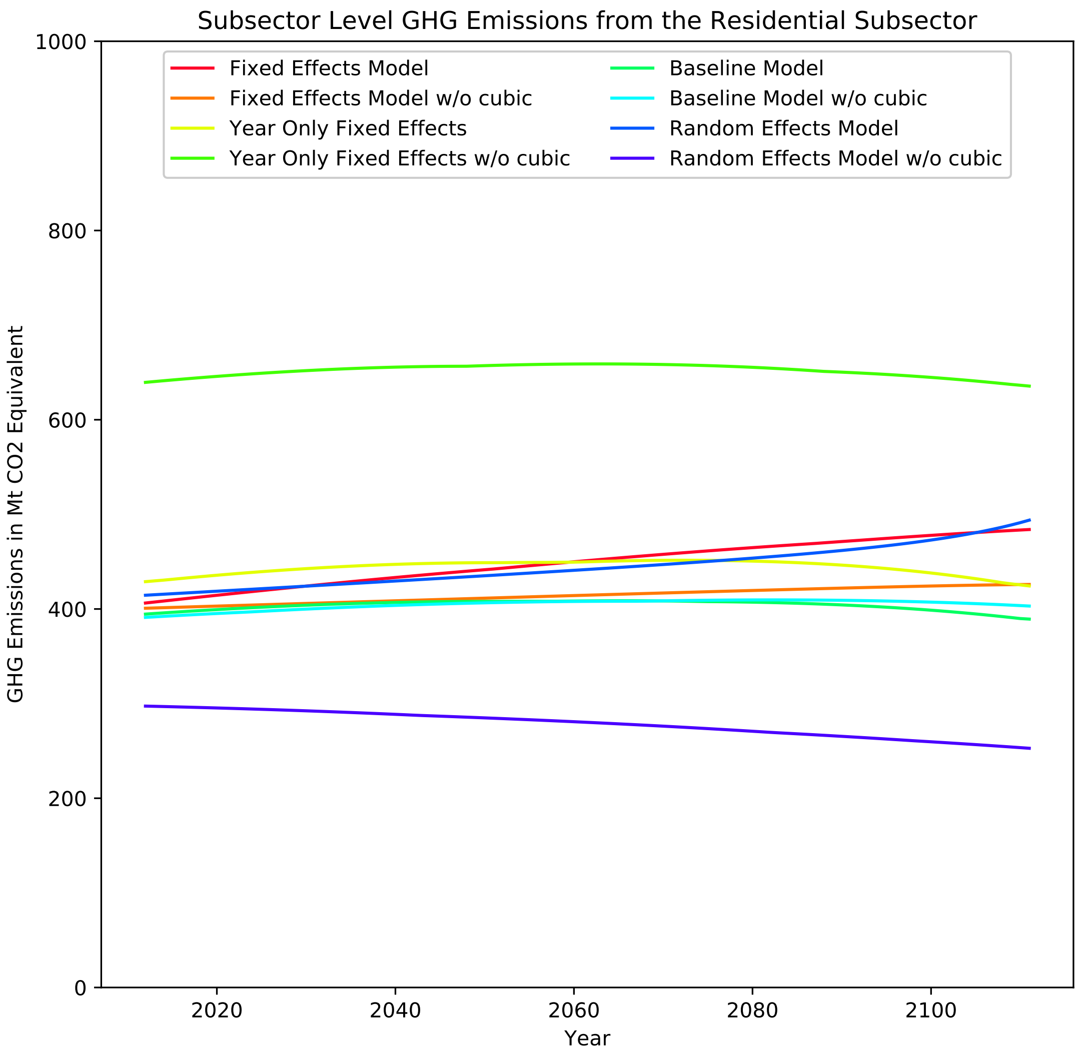



Subsector Level GHG Emissions from the Transportation Subsector
Fixed Effects Model
Fixed Effects Model w/o cubic
Year Only Fixed Effects
Year Only Fixed Effects w/o cubic
Baseline Model
Baseline Model w/o cubic
Random Effects Model
Random Effects Model w/o cubic
GHG Emissions in Mt CO2 Equivalent
Year
9000
8000
7000
6000
5000
4000
3000
2000
1000
2020
2040
2060
2080
2100

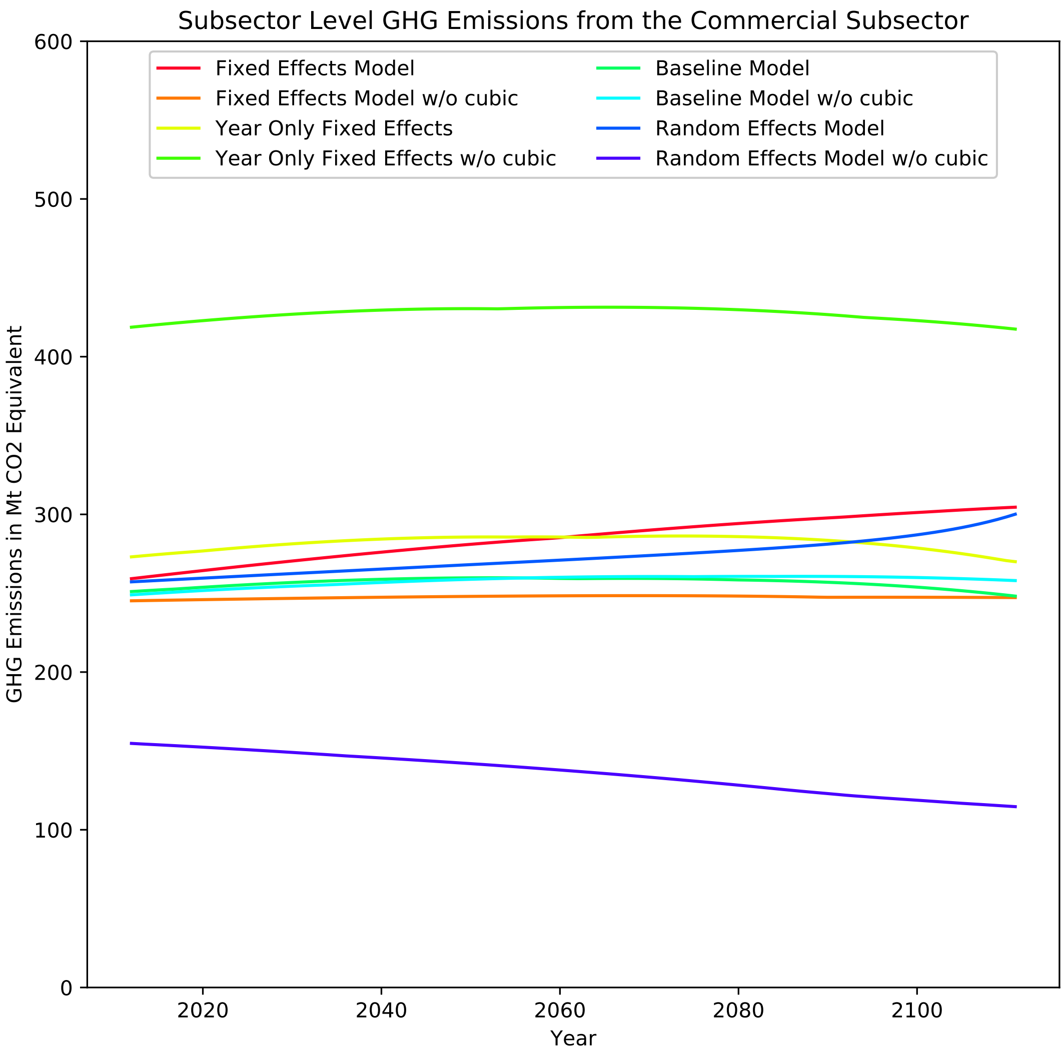
Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2019 doi:10.20944/preprints201905.0058.v1



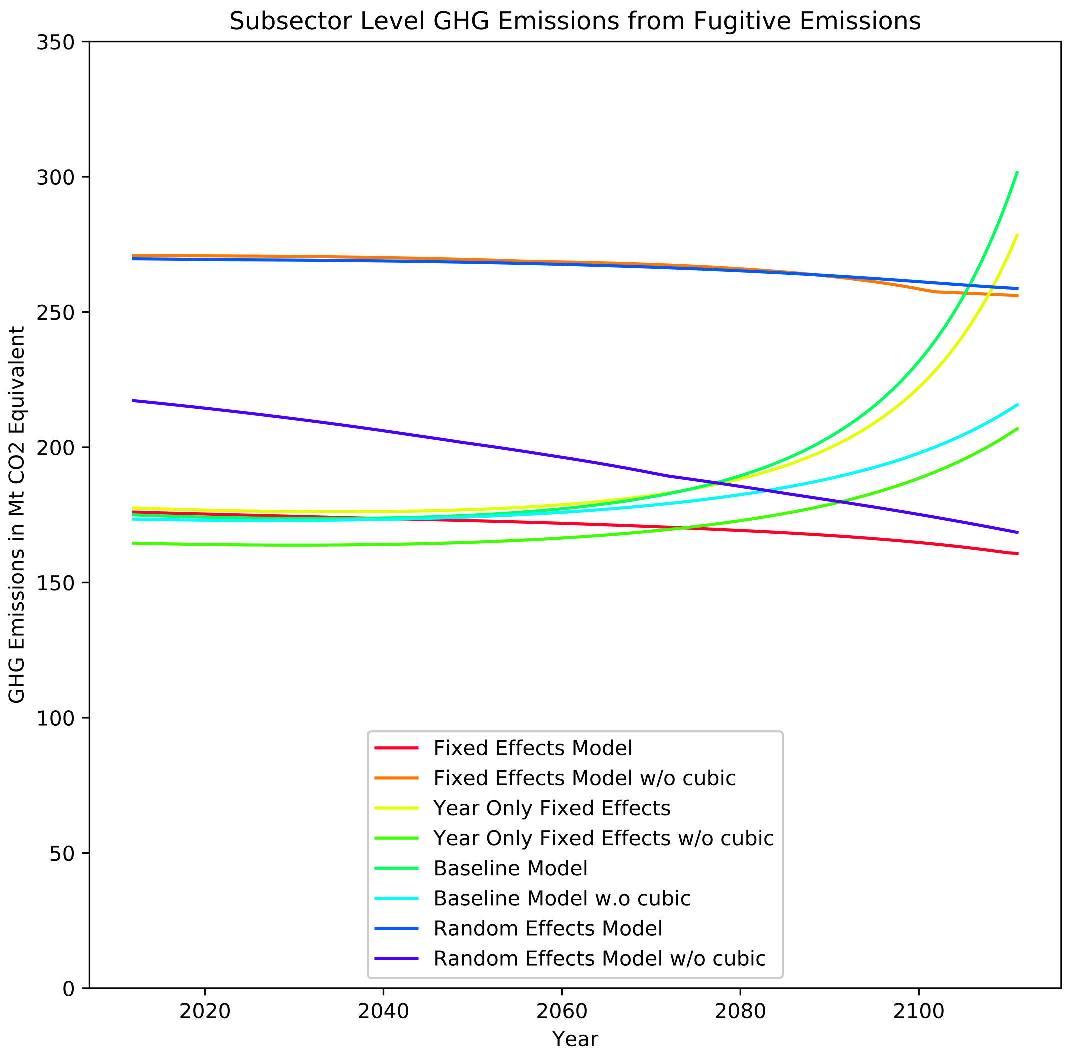
 doi:10.20944/preprints201905.0058.v1



Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2019 doi:10.20944/preprints201905.0058.v1



Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2019 doi:10.20944/preprints201905.0058.v1



B Code and Data Source Availability

The python script and csv files containing data from WRI CAIT Climate Data Explorer (2014) are available online at:

`https://github.com/cat-astrophic/energy_subsector_emissions`

Please cite the data source if using the data used in this paper, but feel free to use, edit, and redistribute the script state_level_stats.py freely!