A Spatio-Temporal Econometric Approach to Estimating U.S. State-Level Energy Emissions

J. Wesley Burnett[1], John C. Bergstrom[2], and Jeffrey H. Dorfman[3]

[1]Assistant Professor, Division of Agricultural and Resource Economics, West Virginia University, PO Box 6108, Morgantown, WV 26506, 304.293.5639, jwburnett@mail.wvu.edu
[2]Professor, Department of Agricultural and Applied Economics, University of Georgia, 208 Conner Hall, Athens, GA, 30602, 706.542.0739, jberg@uga.edu
[3]Professor, Department of Agricultural and Applied Economics, University of Georgia, 312 Conner Hall, Athens, GA, 30602, 706.542.0754, jdorfman@uga.edu

August 26, 2011

Keywords: Pollution Economics, Energy Economics, Spatial Econometrics, Dynamic Panel Data, Carbon Dioxide Emissions, Global Climate Change
JEL Codes:

Abstract

We take advantage of an unusually long panel data set to allow for the estimation of a complex specification of the relationship between state-level, per-capita energy emissions and per-capita economic growth. In particular, we propose a newly developed spatio-temporal panel data estimation procedure to account for linkages between states due to factors such as interstate energy commerce and the time series properties of the data. When we fail to account for spatial dependence within the data we find strong evidence of an inverted-U shaped relationship between energy emissions and state-level GDP. However, the evidence of this relationship breaks down when we specify the model with spatial dependence while variables such as interstate electricity commerce and percentage of electricity produced from renewables remain significant. Our results suggest that reductions in per-capita energy emissions is taking place because some states are switching to less carbon intensive energy inputs and because some states (arguably with laxer environmental regulations) are exporting electricity to states with stronger environmental regulations.

1 Introduction

Economists, ecologists, private industries and government decision-makers have long been interested in the relationship between economic growth and environmental quality. These relationships are often the subject of intense public policy debates such as the recent Copenhagen Treaty signed by the current presidential administration at the 2009 U.N. Climate Change Conference. Under this treaty the administration has proposed to cut greenhouse gas emissions in the U. S. by 17% by 2020 and 42% by 2030. In the U. S. many opponents to this legislation claim that carbon pollution abatement policies will hinder economic growth. Supporters, on the other hand, claim that such policies are absolutely necessary to prevent irreversible global warming caused by anthropogenic emissions of greenhouse gases.

From an ecological or environmental perspective, the assumption is often made that economic growth is bad for the environment. But, what story does the empirical data tell us? One's intuition may lead to the belief that pollution will continue unabated as a country's economy grows through time. An examination of the empirical relationship between economic growth and emissions, however, often reveals different results as evidenced by the environmental Kuznets curve (EKC) hypothesis. The EKC hypothesis describes the time path that pollution follows through a country's economic development. This hypothesis claims that environmental degradation follows an inverted U-shaped relationship as a country's economy develops over time.

Since the conception of the hypothesis, researchers have examined a wide variety of pollutants seeking evidence of the EKC. Separate studies have experimented with different econometric approaches, including: different orders of polynomials, fixed and random effects, semi-parametric and non- parametric techniques, splines, and different covariates specifications (Levinson, 2008). Past studies have also examined different jurisdictions and time periods. Certain generalizations seem to emerge across these different approaches—i.e., it seems that pollutions levels at least approximately improve for some pollutants as income per capita grows.

Despite a sizable literature, issues regarding the spatial and temporal dependence within the data have not been thoroughly addressed leaving a gap in the literature. There has been more attention as of late to the temporal dependence within the data,[1] however very little attention has been paid to spatial dependence.

In this paper we contribute to the literature by controlling for both the spatial and temporal aspects within the data. Given a long data set, similar to that used by Aldy (2005), we seek to find empirical confirmation of the environmental Kuznets curve in the U.S. More precisely, we will test this hypothesis by examining the relationship between state-level, per-capita income and per-capita CO_2 emissions. Consistent with Aldy (2005) and Carson (2010), we argue that interstate energy commerce affects this CO_2-income relationship, however we posit that space and intrastate pollution policies largely dictate this commercial process. In other words, differing intrastate environmental standards encourage some states to import electricity (because high environmental standards make it costly to generate electricity from dirty sources) and indirectly encourage other

[1] For example, see Stern and Common (2001), Perman and Stern (2003), Egli (2004), and Aldy (2005)

states to export electricity (arguably from states with laxer environmental standards).[2] Of course, interstate commerce is much more complicated than simply observing the importation and exportation of electricity–states engage in other energy trading activities including the trade of coal, natural gas, and oil.

Space is important in this relationship because it is costly to transport energy or energy resources across large distances. Although the US electrical grid is capable of transmitting electricity across several state lines there is no national grid system–the current regulatory environment has divided the system into three power grids which are not simultaneously linked together (DOE, 2010).[3] Thus, electricity transmission is somewhat limited by space. Additionally, some energy sources such as coal are heavy and therefore difficult (costly) to transport across vast distances. These transportation costs add to the wholesale price of the energy resource thereby making that resource non-competitive with other energy sources in some areas. The cost of shipping coal by rail can sometimes be as costly as actually mining the coal itself. Therefore, the basic argument is that interstate energy commerce becomes more expensive the greater the distance between any two states.

To see how space matters consider the state of California. California has the highest populous in the Nation and a demand for energy second only to Texas (EIA, 2008). Yet, despite its population and demand California has one of the lowest per capita energy production rates in the US. How is this possible? Simple, California is the largest electricity importer in the Nation. It imports hyrdoelectricity form the Pacific Northeast and coal-fired electricity from the Desert Southwest. Using California as our example, our argument is simply that economic growth stimulates energy demand (so as state-level GDP expands we would expect to see increases in energy demand). However, economic growth in California is driving energy production in its *neighboring* states (and consequently potential environmental damages as a result of coal-fired electricity generation).

It would be extremely difficult to model all of US interstate energy commerce and its effect on the environment, so a parsimonious way of modeling this relationship is to incorporate space into the empirical model. Thus, modeling the spatial-temporal relationship is a simple way of controlling for energy commerce between states.

Our examination of carbon dioxide emissions and GDP takes place within the 48 contiguous states of the US from 1963-2001. Carbon dioxide accounted for 84% of U.S. greenhouse gas emissions in 2005 and is one of the largest hypothesized contributors to climate change (Brown et al., 2008). The emissions estimates in our data are based on the combustion of fossil fuels which is the main source of CO_2 emissions in the U.S. According to a U.S. Environmental Protection Agency report, fossil fuel combustion produced 94.1% of the CO_2 emitted in the U.S. in 2008 (U.S. Environmental Protection Agency, 2008). We seek to properly model and estimate the CO_2-income relationship across the 48 contiguous states and verify the inverted-U shaped relationship within the conventional EKC hypothesis.

This paper uses a dataset similar to Aldy (2005), but this paper contributes to the literature

[2]This effect of differing environmental standards on interstate commerce seems to imply a pollution haven hypothesis whereby some states emit more CO_2 emissions than others.

[3]These three power grids include the Eastern Interconnect, the Texas Interconnect, and Western Interconnect.

by offering a newly proposed spatio-temporal panel data model which incorporates a spatial error process to naively model energy trade between states.

This model not only goes beyond Aldy's (2005) approach but it is widely applicable to several other future EKC studies including studies not related to energy.[4] In principle, our model is much simpler to estimate than most spatial panel models which often use complicated maximum likelihood estimation schemes. Our model is primarily an iterated, OLS estimation scheme so we believe that it is fairly intuitive and easily accessible for a large audience.

Another subtle contribution of this paper, which is not distinguished in Aldy's (2005) paper, is that we are potentially modeling both sides of the inverted-U shaped relationship purported by the EKC hypothesis because we have emissions that data back to the early '60s.[5] Many EKC papers are based on datasets starting in the '70s (but mostly in the '80s), after the founding of the US EPA. According to the EPA (2010), the US has experienced decreasing trends in the criteria pollutants (carbon monoxide, lead, ground-level ozone, nitrogen dioxide, particulate matter, and sulfur dioxide) since the '80s. So this begs the question: is the US on the right side of the curve for most pollutants? Are we on the right side of the curve for CO_2 emissions? We posit that it is possible that the enactment (and enforcement) of the US Clean Air Act (CAA) and the foundation of EPA may have biased the shape of the curve.[6] By using data that dates back to before the stronger enforcement error of the CAA we argue that we are getting a truer picture of the full EKC curve.

We find, after modeling several different specifications of the spatio-temporal model, fairly robust evidence to support the EKC hypothesis for CO_2 emissions, but it is possible that this evidence may break down in the face of significant spatial autocorrelation. If the spatial model is correctly specified and the spatial autocorrelation is indeed significant (something we will assert with skepticism as we find mixed support for the spatial dependence within the data), then our findings suggest that no EKC relationship exists between CO_2 emissions and state-level income; instead, these findings imply that certain states may be developing into pollution havens. A projection of our different models finds a consistent estimate of the EKC curve peaking around $36,000 per-capita income.

The map in Figure 1 shows a heat map of contiguous U.S. states based upon each state's per-capita carbon dioxide emissions in 2001 (expressed in teragrams of carbon emitted). As the map shows, Wyoming, North Dakota, and West Virginia have the highest per-capita emissions in the Nation based upon this dataset. This is not coincidental given that each of these states is rich in fossil fuels including coal and natural gas. The high per-capita CO_2 emissions in each of the states is driven by relatively small populations and high dependence on coal-fired electricity which is often exported to other surrounding states.

[4]This spatio-temporal estimator is applicable to any polutant which displays spatial dependence.

[5]This is the reason for using CO_2 estimates instead of actual concentration levels

[6]In other words, did we experience the peak and downtown just because of the CAA or was the inverted U-shaped relationship driven by economic development and growth? Or perhaps the enactment of the CAA was a natural outcropping of economic growth?

[Figure 1 will go about here.]

Consider Wyoming specifically. Wyoming is the largest producer of coal in the Nation (EIA, 2010); it also has the fifth highest per-capita income in the Nation in 2010 (BEA, 2011). Since the state has such a large abundance of coal over 90% of its electricity is generated by coal-fired plants, which (coupled with lax environmental standards) explains its high CO_2 emissions. This relationship between high per-capita income and high per-capita emissions seems counter to the EKC hypothesized relationship. But upon closer glance Wyoming has low in-state electricity demand, so the electricity is being exported to its neighbors–once again space matters.

The rest of this paper is structured as follows. Section 2 will offer a brief review of the literature. In Section 3 we will set up the spatial fixed panel data model. Section 4 will provide a description of the data used to estimate this model and present the empirical estimation procedures and results. In Section 5 we will present the empirical findings. Finally, in Section 6 we will discuss potential policy implications and offer suggestions for further research.

2 Literature Review

Rupasingha et al. (2004) were some of the first authors to offer a spatial econometric approach to testing the EKC hypothesis. Specifically, the authors examined the relationship between per capita income and toxic pollutants at the US county-level. With a quadratic specification the authors find the conventional inverted U-shaped relationship; however, with a cubic specification they find that toxic pollution first decreases but then increases again as income continues to grow over time (this is sometimes referred to in the literature as a N-shaped relationship). These authors' findings are interesting but the pollution-income relationship is only examined in a cross-sectional context the authors could not control for fixed effect within states.

Maddison (2006) examined the emissions of sulfur dioxide, nitrogen oxides, volatile organic compounds and carbon monoxide across 135 nations. He analyzed the conventional EKC with spatially augmented weighted values of the dependent (the pollution emissions) and independent variables (national income and other covariates) to account for potential spatial dependence within the data. The author found that national per capita emissions of sulfur dioxide and nitrogen oxides are heavily influenced by the per capita emissions of neighboring countries. Despite Maddison's (2004) contribution his analysis is limited to only two years. He took first differences of the data to control for the independent effects (fixed effects) within each country and then analyzed the data as one large cross-section.

Aldy (2005) found evidence for the EKC hypothesis with CO_2 emissions in the contiguous US states using a similar dataset to this paper. Aldy made an important distinction between consumption-based emissions and production-based emissions. He argued that through interstate trade, a state's emissions intensity from production may differ from its intensity from consumption. To account for this distinction he modified the data for states that are net exporters of electricity by deducting the state's average electricity carbon intensity (as a proxy for exported electricity) from its total emissions for a given year. Carson (2010) points out that this distinction is important because it helps control for net electricity importing states that consume energy without experiencing externalities associated with their production. We agree with Aldy and Carson, however, we believe that Aldy's procedure is not sufficient enough to capture the full weight of interstate energy trading.[7]

3 Methodological Approach

To control for spatial dependence, temporal dependence, and state-level independent effects we propose a fixed effects estimation procedure as follows. First we specify the traditional panel data approach and then we contribute to the literature by expanding it to a dynamic panel approach (i.e., include a lag term of the dependent variable). The dynamic panel approach is a parsimonious way of accounting for persistent effects within the explanatory variable. We argue that the underlying economic activity, which drives energy demand and consumption, is driving the persistence within the CO_2 estimates. Specifically, the standard model we consider is

$$y_t = X_t\beta + \mu_i + \eta_t + u_t, \tag{3.1}$$

where y_t denotes a (*N x 1*) vector of U.S. state-level per capita carbon dioxide emissions. X_t is an (*N x K*) matrix of the explanatory variables including per capita GDP and per capita GDP squared. The coefficient μ_i denotes the individual effect for each U.S. state and η_t denotes the time effect. In the present analysis we treat the individual effect as fixed meaning that we assume that this variable is correlated with the explanatory variables and approximately "fixed" over time for each state within the sample. This fixed effect may be thought of as state infrastructure, political structure, etc. We could estimate μ_i directly by adding a dummy variable for each cross-section and estimating the equation via ordinary least squares (OLS); this is sometimes referred to as the least squares dummy variable estimator (LSDV).[8] If we allow for the fixed effects term to enter into the error term and we estimate (3.1) without controlling for it, then the estimates will result in omitted variable bias. To control for fixed effects without including a fixed effects term in (3.1), we could demean the data (fixed effects or within estimator) or take the first difference of the data (first-difference estimator). Econometric theory tells us that the LSDV coefficient estimates are asymptotically equivalent to the fixed effects estimates.

Alternatively, if we think that the dynamic model is a more appropriate specification then we could specify the model as

$$y_t = \rho y_{t-1} + X_t\beta + \mu_i + \eta_t + u_t, \tag{3.2}$$

[7] In other words, a simple transformation of the data to account for electricity trade is somewhat ad hoc since there is no theory to determine how to control for interstate energy trade.

[8] In order to conduct the LSDV estimator we must eliminate the constant term in the matrix X_t or add a constraint to μ_i such as $\sum_i \mu_i = 0$, otherwise the individual effect (dummy variable) would be indistinguishable from the intercept term.

where all the variables are the same as (3.1) but we have included a lagged dependent variable on the RHS. The coefficient ρ denotes a scalar on the time lagged variable of CO_2 emissions. (3.2). Again, we can estimate (3.2) by any standard procedure including LSDV, fixed effects, or first-difference estimation.

The choice between a standard fixed effects versus the dynamic fixed effects specification depends largely on the degree of stationarity of the dependent variable (carbon dioxide emissions). We will see later in the analysis that if CO_2 is first-difference stationary then the dynamic panel specification could potentially over-difference the CO_2 series which could cause further bias. If the CO_2 series is second-difference stationary then perhaps the dynamic specification is correct. Panel unit root tests are conducted in the Appendix in attempt to determine which specification better fits the data.[9]

To complicate the model a little further, we now assume that emissions and income in one state are affected by energy trade with other states so that there is a spatial relationship. Specifically, we assume a spatial autoregressive error process as follows where the error term u_t is defined as,

$$u_t = \lambda W u_t + \epsilon_t \tag{3.3}$$
$$E(\epsilon_t) = 0 \tag{3.4}$$
$$E(\epsilon_t \epsilon_t^{'}) = \sigma_t^2 I_N. \tag{3.5}$$

I_N is an identity matrix of size *N*. λ is the coefficient on the spatial autocorrelation term. *W* denotes an (*N x N*) non-negative spatial weight matrix consisting of zeros along the diagonal and elements w_{ij} elsewhere. w_{ij} is measure of the *a priori* strength of the interaction between location i (the row of the W matrix) and location j (the column) (Anselin, Le Gallo, and Jayet, 2008). In the simplest case the weights matrix is binary with $w_{ij} = 1$ when i and j are neighbors and $w_{ij} = 0$ otherwise.[10] In the spatial econometrics literature the weights are generally standardized such that the elements in each row sum to one (row standardization). Using the spatial autogressive error specification is a way of naively modeling the underlying economic process (including trade) which drives energy consumption.[11] An ideal model may consist of an engineering model of the US electrical grid system, but, such a model would require a copious amount of data. By specifying it in this manner we are essentially saying that we do not know exactly how that process works because it is so complicated, but this process does have an effect on energy consumption which is reflected in the error term (this in turn will affect our CO_2 emission estimates). With (3.3) - (3.5) above we can assume that ϵ_{it} is a white noise process or we can make the stronger assumption that the error terms are i.i.d.for all *i* and *t* with mean zero and variance σ_ϵ^2.

According to Anselin, Le Gallo, and Jayet (2008), the specification of the spatial weights matrix is of great import in applied spatial econometrics. This specification is particularly important in our

[9]The panel unit root results are somewhat inconclusive so for completeness we present both estimation specifications; presenting both specifications also allows for a sensitivity analysis.

[10]Hence the zeros along the diagonal since a state cannot be a neighbor to itself.

[11]This specification is often referred to as "indirect" approach to modeling the spatial relationship.

context because it represents the basic structure of energy commerce in our model. Initially we will specify a spatial weighting matrix based upon the inverse of the distance between state centroids–this implies that the greater the distance between the states the less likely energy commerce has taken place between the states.[12] This assumption will be relaxed later to consider a first-order contiguity specification whereby states are more likely to trade with their immediate neighbors. Finally, we assume the weights remain constant over time. Alternative specifications may allow the scalar parameter to vary over time or allow the weights to vary and the parameter to remain constant. Such specifications would only complicate the analysis, so to keep the empirical approach tractable we assume constant weights across time.

To see how the assumption of spatial autoregressive errors affects (3.1) we can rewrite (3.3) as,

$$(I_N - \lambda W)\, u_t = \epsilon_t \tag{3.6}$$

$$u_t = (I_N - \lambda W)^{-1}\, \epsilon_t. \tag{3.7}$$

We can now plug (3.7) into (3.1) and rewrite some terms to derive,

$$y_t = X_t\beta + \mu_i + (I_N - \lambda W)^{-1}\, \epsilon_t. \tag{3.8}$$

Intuitively, if we estimate λ then what we are doing with (3.8) is controlling for interstate energy commerce by removing it from the error term. From an econometrics perspective the issue with estimating (3.8) is a problem of efficiency. If one estimates equation (3.1) or (3.2) without taking the spatial effects into account, then a standard OLS regression will yield downwardly biased standard errors. In other words, naive standard errors overestimate the amount of precision in the parameter estimate and biases t-statistics upward. This problem of inefficiency is generally accounted for through instrumentation (IV or GMM) or by specifying a complete distribution model (maximum likelihood).

Elhorst (2005) developed an unconditional maximum likelihood estimator for a dynamic, spatial panel data in which he takes the first difference of the data to eliminate the fixed effects term. The drawback with Elhorst's procedure is that it requires assumptions about the initial conditionals for y_0 and X_0. Bond (2002) criticizes this procedure by stating that the distributions of both y and X for $t = 2, 3, \ldots, T$ could depend in a "non-negligible" way on what is assumed about the initial condition (especially if T is short).

Elhorst (2003) also offered a fixed effects (FE) estimator for panel data but Anselin, et al. (2008) argue that it is biased.[13]

Elhorst's (2005) unconditional ML estimator is attractive but complicated, and since Elhorst's (2003) FE estimator is biased we propose the following iterative Spatial Fixed Effects (SFE) estimator which is consistent, asymptotically normal and robust to heteroskedasticity and serial corre-

[12] Of course, we know that is not entirely accurate because Southern states purchase low-sulfur coal from the Powder River Basin in Wyoming and Montana, but in general we believe this to be a fairly accurate representation of interstate energy commerce

[13] See the Appendix below for an explanation for why Elhorst's FE estimator is biased.

lation. In principal our SFE estimator is also far simpler to compute than Elhorst's unconditional ML and Elhorst's FE estimator. To begin we rewrite (3.1) and (3.3) by stacking the observations as successive cross-section over time.

$$y = Z\gamma + (I_T \otimes \mu) + u \tag{3.9}$$
$$u = \lambda W u + \epsilon, \tag{3.10}$$

where $Z = (y_{-1}\ X)$ and $\gamma = (\rho\ \beta)'$ and the weighting matrix $W = (I_T \otimes W_N)$. For the standard fixed effects estimation the vector Z would not contain the lagged dependent variable and the vector γ would not contain ρ. I_T is an identity matrix of dimension (T) and the subscript N on W indicates that it is of dimensions (N x N). Following Anselin et al. (2008) we define the "within" transformation operator as[14]

$$Q = I_{NT} - \left(\frac{\imath_T \imath_T'}{T} \otimes I_N \right), \tag{3.11}$$

where $\imath_T$ is a vector of ones of length T. The within transformation means that each cross-section is demeaned "within" its own section; e.g., all the CO_2 emissions in Wisconsin will be demeaned based upon the mean of emissions from Wisconsin. We can multiply the within transformation operator through (3.9) and (3.10) to eliminate the fixed effects term, μ, as follows

$$Qy = QZ\gamma + Qu \tag{3.12}$$
$$Qu = \lambda W Q u + Q\epsilon. \tag{3.13}$$

Ord (1975) states that if the spatial autocorrelation parameter (λ) in (3.13) is unknown, even a constrained least squares procedure produces inconsistent estimators. He defines an iterative procedure that we extend here to a panel data model:

1. From (3.12), apply GLS of Qy on $QZ\gamma$ to derive residuals $Q\tilde{u}$
2. Estimate $\tilde{\lambda}$ from $Q\tilde{u} = \lambda W Q\tilde{u} + Q\epsilon$ by using a Newton-Raphson algorithm to derive $\tilde{\lambda}$
3. Construct new variables $\ddot{y} = \left(I_{NT} - \tilde{\lambda} W\right) Qy$ and $\ddot{Z} = \left(I_{NT} - \tilde{\lambda} W\right) QZ$
4. Apply OLS for $\ddot{y}$ on $\ddot{Z}$ to yield $\hat{\gamma}$; i.e., $\hat{\gamma} = \left(\ddot{Z}'\ddot{Z}\right)^{-1} \ddot{Z}'\ddot{y}$
5. Construct the new residuals as $\hat{\ddot{u}} = \ddot{y} - \ddot{Z}\hat{\gamma}$ and return to Step 2 to recalculate $\hat{\lambda}$. Repeat Steps 2 through 5 until the convergence criteria for λ and the objective function are met.

[14]The reader should note that the demeaning operator is slightly different from the operator within the traditional panel data models because the data is organized differently.

6. Construct the robust covariance for $\hat{\gamma}$ by calculating $\hat{A} = \frac{N}{N-K}\left(\ddot{Z}'\hat{\hat{u}}\hat{\hat{u}}'\ddot{Z}\right)$.[15]

Notice that we use the generalized least squares (GLS) procedure because Q^{-1} does not exist since it is idempotent. Instead we use the pseudo-inverse (or Moore-Penrose Inverse) as defined in Hsiao (2003). The convergence criteria for the parameters and objective function for the loop process (Steps 2 - 5) are offered in Gill, Murray, and Wright (1982).[16]

3.1 Newton-Raphson Method

The Newton-Raphson method to derive $\tilde{\lambda}$ from Step 2 above is derived from Ord (1975) as follows: The ML estimator is the value of λ which minimizes

$$f(\lambda) = -\frac{2}{n}\sum_{i=1}^{n} ln(1-\lambda\kappa_i) + ln(s^2) \tag{3.14}$$

where

$$s^2 \equiv s^2(\lambda) = \tilde{u}'\tilde{u} - 2\lambda\tilde{u}'\tilde{u}_W + \lambda^2(\tilde{u}_W)'\tilde{u}_W, \tag{3.15}$$

and $\tilde{u}_W = W \cdot \tilde{u}$. The derivatives of (3.14) are

$$f_\lambda(\lambda) = \frac{2}{n}\sum_{i=1}^{n} \kappa_i/(1-\lambda\kappa_i) + 2\left(\lambda(\tilde{u}_W)'\tilde{u}_W - \tilde{u}'\tilde{u}_W\right)/s^2 \tag{3.16}$$

and

$$\begin{aligned} f_{\lambda\lambda}(\lambda) \quad &= \tfrac{2}{n}\textstyle\sum_{i=1}^{n}(\kappa_i)^2/(1-\lambda\kappa_i)^2 + 2(\tilde{u}_W)'\tilde{u}_W/s^2 \\ &\quad -4\left(\lambda(\tilde{u}_W)'\tilde{u} - \tilde{u}'\tilde{u}_W\right)^2/s^4. \end{aligned} \tag{3.17}$$

λ, is determined iteratively from the expression

$$\lambda_{i+1} = \lambda_i - f_\lambda(\lambda_i)/f_{\lambda\lambda}(\lambda_i) \tag{3.18}$$

where the starting value is $\lambda_0 = \tilde{u}'\tilde{u}_W/\tilde{u}'\tilde{u}$.

Provided that the variables are stationary and the orthogonality condition $E(\ddot{Z}'\hat{\hat{u}}) = \mathbf{0}$ holds, then this estimator is consistent and asymptotically normally (as *N* and *T* approach ∞). In the case of the dynamic specification, this iterated procedure relies upon the assumption that if T is sufficienty large then the dynamic panel bias becomes insignificant. We will refer to this procedures dynamic counterpart simply as the Dynamic SFE.

[15] $\hat{A}$ is often referred to as the meat of the sandwich from the “sandwich estimator." The term $\frac{N}{N-K}$ is a degrees of freedom correction since $\hat{\hat{u}}$ is biased.

[16] Our estimation results yielded stable coefficient estimates and a stable objective function.

Conveniently this same iterated procedure can be performed with a panel first-difference (FD) estimator. We refer to this procedure as the iterated spatial first-difference (SFD) estimator. The algorithm is almost identical to the SFE procedure (for the specific algorithm please refer to the Appendix). Unlike the SFE, the variables do not have to be stationary, but if they are first-order stationary then the SFD estimator is consistent (provided the orthogonality condition holds) and asymptotically normal.

Since the spatial error coefficient, λ, is numerically approximated via a Newton-Raphson algorithm we do not estimate its standard error directly. Without the standard error estimate we cannot calculate the usual t-statistic to determine if the coefficient is statistically significant. Its significance may imply that our spatial error process is a correct specification. However, we can somewhat get around this problem by estimating the asymptotic standard error for the spatial error parameter. We derive the estimate for the asymptotic standard error from "Appendix B" of Ord (1975). Additionally, we use a Lagrange Multiplier (LM) test to determine if there is spatial dependence within the data. The LM test (Burridge, 1980) is as follows

$$LM = \frac{\left(Q\tilde{u}' \left(I_T \otimes W_N\right) Q\tilde{u} / \left(Q\tilde{u}' Q\tilde{u}/NT\right)\right)^2}{tr\left(\left(I_T \otimes W_N^2\right) + \left(I_T \otimes W_N' W_N\right)\right)}, \tag{3.19}$$

where $Q\tilde{u}$ are the derived residuals from Step 1 of the algorithm above. tr indicates the trace of the elements in the denominator. Burridge's (1980) LM test is distributed as a chi-square with one degree of freedom.

4 Data Description

The CO_2 emissions data were obtained from the Carbon Dioxide Information Analysis Center (CDIAC) within the U.S. Department of Energy (Blasing et al., 2004). CDIAC estimates the emissions by multiplying state-level coal, petroleum, and natural gas consumption by their respective thermal conversion factors. Therefore, this dataset represents estimates of CO_2 emissions and not actual emissions, which is somewhat problematic as actual emissions data would be more desirable. The reason for using this particular dataset, however is that it offers emissions estimates dating back to 1960, well before the establishment of the Environmental Protection Agency (EPA) and stronger enforcement of the U.S. Clean Air Act. CO_2 emissions are offered in per capita terms to control for changes in energy consumption based upon changes in the state population.

The GDP data was obtained from the Bureau of Economic Analysis (BEA) within the U.S. Department of Commerce (Bureau of Economic Analysis, 2010). The BEA offers annual state-level GDP estimates from 1963 to the near present. The estimates are based on per capita nominal GDP by state. The estimates were converted to real dollars by using the BEA's implicit price deflater for GDP. Following the traditional EKC hypothesis, GDP is expected to have an inverted U-shaped relationship with CO_2 emissions; in other words, a quadratic polynomial of GDP will be specified in which the expected sign on the GDP term is positive while the expected sign on the squared term is negative. To test this specification the polynomial can be extended to higher powers to determine if the leading term is statistically significant. For example, if a cubed GDP term is positive and statistically significant, the implication is that the CO_2-income relationship is N-shaped—i.e., a relationship that is characterized by an initial increase in pollution, followed by a decrease, and then an increase once again as income continues to grow through time. We hypothesize that the quadratic relationship is the correct specification.

Cooling Degree Days (CDD) and Heating Degree Days (HDD) data were obtained from the National Climate Data Center within the National Oceanic and Atmospheric Administration (National Climate Data Center, 2010). The data are offered in a state population-weighted format consistent with the rest of the data in the study. CDD (or HDD) is a unit of measure to relate the day's temperature to the energy demand of cooling (or heating) at a residence or place of business—it is calculated by subtracting 65 degrees Fahrenheit from the day's average temperature (Swanson, 2005). Residential energy consumption has been found to be highly correlated with CDD and HDD (Diaz and Quayle, 1980). Since the CO_2 emissions are estimated from energy consumption, the CDD and HDD data as quantitative indices should capture much of the year-to-year variation in energy consumption. CDD and HDD are expected to be positively related to CO_2 emissions as cooler (or hotter) days would induce households or businesses to demand higher amounts of energy for heating (or cooling) a residence or place of business.

The energy production data were obtained from the Energy Information Administration within the U.S. Department of Energy (Energy Information Administration, 2008). The energy production data represent state-level annual production of coal, crude oil, natural gas, and renewable energies. The data are represented in physical units: short tons, barrels, and cubic feet, respectively. The production data were converted to a population-weighted format by dividing today physical units by the state's annual population estimate. Due to data limitations, the natural gas and renewable energy production measures were dropped from the analysis. The coal and oil production measures are left in levels as several of the measures contain zeros; i.e., not all states produce coal or oil. State-level production is expected to be positively related to CO_2 emissions as an increased supply in energy may make consumption of the energy more readily available for the state. For example, if a state produces coal then it is expected that that state will keep some in reserve to use in energy production within the state.

Annual state population data were obtained from the U.S. Bureau of Census (Population Estimates). These population estimates represent the total number of people of all ages within a particular state.

5 Empirical Estimation and Results

Following the traditional EKC hypothesis with the quadratic specification, we define the spatio-temporal CO_2-income relationship as

$$\begin{aligned} ln(y_{it}) = {} & \rho y_{it-1} + \beta_1 ln(GDP_{it}) + \beta_2 ln(GDP_{it})^2 \\ & + \beta_3 ln(CDD_{it}) + \beta_4 ln(HDD_{it}) + \beta_5 coal_{it} + \beta_6 oil_{it} + \mu_i + \eta_t + u_{it} \\ & i = 1, \ldots, N; \quad t = 1, \ldots, T \end{aligned} \tag{5.1}$$

where y_{it} is real per capita CO_2 emissions in a U.S. state and y_{t-1} denotes its lagged value. GDP_{it} is real per capita state-level GDP, and GDP_{it}^2 is the square of the same term. CDD_{it} is per capita cooling degree days, whereas HDD_{it} is per capita heating degree days. $coal_{it}$ and oil_{it} denote per capita state-level annual production of coal and crude oil, respectively. We assume fixed state-specific effects, μ_i. Time effects are denoted by η_t. The state-specific effects capture heterogeneous elements within each state that may affect CO_2 emission levels. All variables with the exception of the intercept terms and energy production are expressed in natural logarithms out of the convention and because this specification fits the data well.[17] The observations in (5.1) are available from 1963-2001 so that $T = 39$. The observations in (5.1) constitute the 48 contiguous states in the U.S. so that $N = 48$.

For ease of exposition the variables are stacked as successive cross-sections over time for $t = 1, \ldots, T$. Next we place the explanatory variables into an *(N x K)* matrix X_t and place their corresponding coefficients into a *(K x 1)* matrix β and rewrite (5.1) as

$$y_t = \rho y_{t-1} + X_t\beta + \mu_i + \eta_t + u_t. \tag{5.2}$$

Following (3.2) - (3.4) above we define the error term as

$$u_t = \lambda W u_t + \epsilon_t \tag{5.3}$$

$$E(\epsilon_t) = 0 \tag{5.4}$$

$$E(\epsilon_t \epsilon_t^{'}) = \sigma_t^2 I_N. \tag{5.5}$$

Given the assumption of the error term in (5.3) - (5.5) we can rewrite (5.2) as

$$y_t = \rho y_{t-1} + X_t\beta + \mu_i + \eta_t + (I_N - \lambda W)^{-1} \epsilon_t. \tag{5.6}$$

The regression equation in (5.6) has a spatial autoregressive process incorporated in the error term with a spatial weight matrix specified as the inverse distance from the state centroids (we will also consider a normalized spatial weight matrix specified as a normalized binary contiguity matrix).

For a sensitivity analysis we compare the iterated SFE and SFD estimators (and their dynamic counterparts) to the other estimation schemes discussed in the Methodological Approach; i.e., LSDV, Elhorst's FE, and Elhorst's unconditional maximum likelihood (UML). Table 1 reports the estimation results based on the complete sample of 1872 observations (or 1824 observations

[17] Production is not expressed in natural logs because several states had zero values for coal or oil produced which is undefined when converted to natural logs.

in terms of the SFD estimator). The second column indicates the least squares dummy variable (LSDV) estimates.[18] The LSDV estimation procedure yields asymptotically biased standard error estimates (if the spatial dependence is significant) because it does not account for the spatial effects within the data. Nevertheless, the LSDV estimates are used as a baseline of comparison against the other estimation schemes. Columns three and four report the results for the spatial fixed effects estimator (Elhorst, 2003) and the fixed effects, unconditional maximum likelihood estimator (Elhorst, 2005) respectively. The fifth through eighth columns represent the spatial fixed effects and spatial first-difference estimators outlined in this paper. Unlike the other estimation schemes the LSDV estimator does not account for spatial error autocorrelation which explains the absence of an estimate for λ (the spatial autocorrelation parameter) in its column.

[Table 1 will go about here.]

The LSDV estimates imply the usual inverted-U shaped relationship of the EKC hypothesis and both indicators of income are highly statistically significant. The positive sign on the HDD is consistent with expectations as an increase in HDD is expected to increase the cooling of buildings which in turn requires the additional combustion of fossil fuels which in turn raises CO_2 emissions. The negative sign on oil production is not necessarily consistent with expectations as an increase in oil production within a state may be expected to elevate CO_2 emissions as the burning of that fossil fuel would be more readily available within that particular state for the production of energy. It could be however that the states that are producing higher levels of oil are exporting a significant portion of their oil to other states or abroad. As expected coal production is positive and statistically significant at the one percent level.

Looking across the different estimation schemes, the estimates for the lag of CO_2 emissions are highly statistically significant for two of estimation schemes. If there is a high degree of persistence within CO_2 emissions then the dynamic, spatial first-difference estimator may inherently yield a difference stationary process for CO_2 emissions. To see this rewrite (5.2) as

$$\Delta y_t = (\rho - 1)\, y_{t-1} + X_t \beta + \mu_i + \eta_t + u_t \tag{5.7}$$

so if the true value of ρ is close to one (persistence) then the lagged emissions variables is approximately equal to zero, and taking the difference yields a difference stationary process. If this is the case then SFD procedure may over-difference the CO_2 series which is biasing the estimates–this may explain the lack statistical significance of the lagged CO_2 emissions term.

The income terms, our main variables of interest, are statistically significant across all the estimation schemes and are consistent with the purported inverted U-shaped relationship espoused by the EKC hypothesis. Elhorst's FE estimation yields the highest estimates of the income terms. Elhorst's UML estimator yields the lowest estimates of the income terms, yet the estimated coefficients are still significant. The spatial fixed effects estimated coefficients in general are very

[18] The LSDV estimates are equivalent to the parameter estimates from Step 1 of the iterated SFE as predicted by econometric theory.

similar to the LSDV estimates, although only the income terms are statistically significant with the SFE estimator. In general the spatial first difference estimation schemes yields lower estimated coefficients than their fixed effects counterparts. According to Wooldridge (2002) the choice between a standard FE vs. a FD estimator depends on the assumptions of the idiosyncratic error term, ϵ. He claims that FE is more appropriate when ϵ_{it} are serially uncorrelated while the FD is more appropriate when ϵ_{it} follows a random walk. Since the underlying demand for energy is driving CO_2 emissions we have reason to believe that it is non-stationary; in other words, we expect a degree of persistence along the time dimension of CO_2 emissions due to the underlying economic activity. Taking the first difference of a non-stationary series such as a unit root will often yield a stationary series. Given stationarity the normal *F* and *t*-stats are applicable. If both GDP and CO_2 are both difference stationary processes then the SFD may be the better estimation scheme because it controls for the non-stationarity within both series. Our panel unit root tests in Table A.1 in the Appendix seem to indicate that both GDP and CO_2 are stationary after we take the first difference and include a trend term. The Dynamic SFD procedure may over-difference the CO_2 series which yields biased estimates–refer to (5.7).

Coal production was found to be statistically significant across three estimation procedures. All signs are positive which is consistent with expectations as we believe that an increase in coal production would increase its burning as a fossil fuel which in turn would increase CO_2 emissions. The effect is relatively small, but it is fairly similar across the estimators; e.g., in the case of the LSDV estimate the interpretation is that a 100% increase in coal production yields a 0.15% increase in CO_2 emissions.

One glaring problem with Table 1 is that the spatial autocorrelation coefficient (λ) is not statistically significant across any of these estimation procedures. The lack of significance does not necessarily indicate that the spatial specification is incorrect but perhaps that the inverse-distance weighting matrix is the incorrect specification. So alternatively we will examine the estimation results with the first-order contiguity spatial weighting matrix. These estimates are reported in Table 2.

As indicated in Table 2 the income terms are significant across estimators except with the spatial fixed effects and spatial first-difference estimators. The estimation results in Table 2 are fairly similar to the results in Tables 1, including magnitudes and signs of estimated coefficients. It is interesting to note that the CO_2-income relationship becomes insignificant with the SFE and SFD procedures when the spatial autocorrelation coefficient is highly statistically significant. If the contiguity specification is correct then these results imply that the EKC hypothesis does not hold when spatial effects are taken into account. Although, these results are somewhat suspect because the first-order contiguity matrix implies that energy commerce primarily takes place between neighboring states which is arguably too simple of an explanation. However, if one were to accept this spatial specification then these results imply that the EKC relationship does not hold when we control for spatial dependence within the data. If the EKC does not hold then CO_2 emissions are improving because they are being displaced by some states. Emissions can be displaced by cleaner forms of energy, cleaner technologies, or some states are becoming pollution havens–the true explanation probably lies somewhere in between. Given our observations of per-capita emissions in

West Virginia, Wyoming, and North Dakota this pollution haven hypothesis is not completely implausible. The estimated spatial autocorrelation coefficient is significant but negative for Elhorst's fixed effects estimator. A negative sign on the spatial autocorrelation coefficient implies that the CO_2-income relationship is more dissimilar across neighboring states which is not consistent with expectations.

[Table 2 will go about here.]

The spatial autocorrelation coefficient (λ) was found to be statistically significant with three procedures in Tables 1 and 2. You will recall that we used a Newton-Raphson algorithm to numerically approximate this coefficient, so we had to use the asymptotic standard error as an approximate estimate of the second moment of the spatial autocorrelation coefficient. Since we had to approximate the second moment we can further test the significance of spatial dependence by using a LM test (Burridge, 1980) outlined in the Methodological Approach section of the paper. Our results for Burridge's (1980) LM test statistic are listed in Table 3 below. The results for this test imply that only the dynamic spatial fixed effects procedure yields statistically significant results.

[Table 3 will go about here.]

Next, we follow up with Aldy's (2005) analysis where he found statistically significant results for both a quadratic and cubic specification of income. If we find statistically significant results for a cubic specification then it may imply that the CO_2-income relationship is following an N-shaped path; i.e., pollution initially rises, tapers off some, and then rises again.[19] This N-shaped relationship implies that CO_2 emissions are not decreasing but increasing with income over time—this refutes the inverted U-shaped relationship.[20]

The results for the cubic specification estimates are presented in Table D.1 in the Appendix. The lagged CO_2 estimate remains significant with this parametric specification (with the exception of the Dynamic SFD); however, the income terms become insignificant for almost all of the estimation schemes. For the SFE and SFD estimators the GDP terms become completely insignificant. Thus, there does not appear to be ample evidence to support growing CO_2 emissions over time with income which seems to offer support in favor of the traditional EKC inverted-U shaped relationship.

Lastly, we project the economic growth-pollution relationship based upon our estimation results derived from Table 1 above. LSDV denotes least squares dummy variable, FE denotes spatial fixed effects, FD denotes spatial first differencing, DFE denotes dynamic spatial fixed effects, and DFD denotes dynamic spatial first differencing. Interestingly, all the different estimation schemes seem to yield estimated results that peak around 10.5 for the natural log of per-capita income (i.e., approximately $36,000). These similar peaks suggest the different estimation approaches are clustered around the true parameter values. As outlined above, the LSDV estimates are larger

[19] This would imply that the sign on GDP is positive, negative for GDP^2, and positive for GDP^3.

[20] Although, Carson (2010) argues that a significant cubic specification may come from the fact that the economic growth-pollution relationship has a relatively flat right tail which the cubic specification may fit better than the quadratic specification.

than the other estimators probably because of multicolinearity and it fails to control for spatial and temporal dependence. Asymptotically, the SFE estimates should be similar to the LSDV estimates, but the SFE estimates are transformed to control for spatial dependence so its estimates are slightly smaller than the LSDV estimates. The Dynamic SFE controls for temporal dependence in the dependent variable and spatial dependence for all the variables so its estimates are smaller than the SFE estimates.

The SFD estimates control for spatial and temporal dependence, and therefore should have the smallest estimates which the figure below seems to corroborate (assuming the data are first-difference stationary). The Dynamic SFD estimates are the smallest because the dependent variable is differenced twice (and these estimates control for spatial dependence), but we question these results because the Dynamic SFD scheme may be over-differencing the data.[21] Therefore, based upon the results and Figure 2, it would seem that in this case the SFD estimator probably offers the best explanation of the economic growth-pollution relationship.

[Figure 2 will go about here.]

6 Implications and Conclusions

Issues associated with spatial dependence have largely been ignored in the Environmental Kuznets Curve literature. In this paper, we introduced a spatial panel data model approach to account for spatial dependence that is expected to be found with CO_2 emissions and state-level income. Compared to past estimators this procedure is in principal much easier to implement and yields consistent and asymptotically normal estimates. Additionally, this estimation procedure contributes to the spatial econometrics literature by offering standard errors that are robust to heteroskedasticity and serial correlation. Based on the empirical results, we believe that we have relatively compelling evidence that this spatial panel approach tells a consistent story in support of the EKC hypothesis with expected signs, magnitudes, significant levels, and possible spatial dependence.

Based upon our empirical results we find evidence that is consistent with the traditional EKC hypothesized inverted-U shaped relationship between CO_2 emissions and income. This inverted-U shaped relationship is found even after controlling for spatial dependence (and arguably interstate energy trade) within the data, which seems to offer further support to the purported EKC hypothesis.[22] We found little statistical evidence of spatial autocorrelation within the data. What evidence we did find for spatial dependence implies that CO_2 pollution emissions are not necessarily a local issue.[23] In other words, a neighboring state's demand for energy may be driving pollution emissions locally. Our findings have implications for potential regional initiatives to alleviate carbon

[21] This result could possibly be correct if CO_2 is second-order difference stationary.

[22] Additionally, it could be argued that the SFD scheme potentially controls for temporal non-stationarity if indeed the first-difference procedure yields difference-stationary processes.

[23] We are not making an argument about CO_2 concentrations in the atmosphere, but rather CO_2 emissions because our data are estimates of emissions based upon the combustion of fossil fuels.

dioxide emissions. Such initiatives are already being undertaken such as the Western Climate Initiative which is a regional agreement among American governors and Canadian premiers to target greenhouse gas reductions (Institute for Energy Research, 2010).

Our results have implications for national and state carbon dioxide policies and especially for policies related to CO_2 abatement and global climate change. Specifically, our estimates for the turning points of emissions at approximately $36,000 per-capita income can be used by policy makers to determine which states are still on the left hand side of the EKC curve; i.e., states where the marginal utility of income still outweighs the marginal disutility of CO_2 emissions– these findings may give policy makers insight into why some states have higher emissions than other. Further, if the true data generating process for the CO_2-income relationship is characterized by our spatial model then this signals to policy makers that CO_2 emissions are a regional problem. In terms of practical application, policy makers can use this research to examine their own state's current energy infrastructure and energy trade balance to help mitigate the CO_2 emissions of their trading partners. For example, California could set up a regional trading plan with the Desert States since California's consumption of electricity is driving coal-fired electricity generation in its neighboring states. However, if the EKC hypothesis is correct then the lagging state's CO_2 emissions may improve over time simply through "free" gains from income growth; i.e., eventually the state's marginal disutility of pollution will outweigh its marginal utility of income and then we would expect to see a decline in emissions. So instead policy makers may decide that economic growth policies are better measures to alleviate carbon dioxide emissions.

We found little statistical evidence to support spatial dependence within the data, but the lack of evidence does not necessarily undermine the results of our spatial-temporal panel procedures.

Finally, the model may be greatly improved by specifying a spatial heterogeneous parameter model as opposed to the homogeneous model we have specified in this analysis. By homogeneity we are implicitly assuming that each state has the same CO_2-income relationship (including the shape) on average across time. As our analysis is restricted to the contiguous 48 states this may not be so problematic, but should the analysis be extended to an international study then the homogeneity assumption may prove more problematic. Of course, implementing a heterogeneous dynamic panel model is already problematic because of the incidental parameters problem, so extending the dynamic panel to include heterogeneous spatial dependence may prove to be very difficult. A clustering estimation scheme may be more appropriate for considerations of spatial heterogeneity.

References

[1] Aldy, J. (2005). "Environmental Kuznets Curve Analysis of U.S. State-Level Carbon Dioxide Emissions." *J. of Environment and Development*, *14* (1), 48-72.

[2] Anselin, L., Le Gallo, J., & Jayet, H. (2008). . Spatial Panel Econometrics. In L. Matyas, & P. Sevestre (Eds.), *The Econometrics of Panel Data* (pp. 625-660). Berlin: Springer-Verlag.

[3] Blasing, T., Broniak, C., & Marland, G. (2004). “Estimates of Annual Fossil-Fuel CO2 Emitted for Each State in the USA and the District of Columbia for Each Year from 1960 through 2001.” Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory. Oak Ridge, TN: US Department of Energy.

[4] Bond, S. (2002). "Dynamic Panel Data Models: A Guide to Micro Data Methods and Practice." Institute for Fiscal Studies, Department of Economics, UCL: Working Paper CWP09/02.

[5] Brown, M., Southworth, F., & Sarzynski, A. (2008). *Shrinking the Carbon Footprint of Metropolitan America.* Brookings Institute: Metropolitan Policy Program.

[6] Burridge, P. (1980). "On the Cliff-Ord test for spatial autocorrelation." *J. of the Royal Statistical Society B , 42*, 107-108.

[7] Carson, R.T. (2010). "The Environmental Kuznets Curve: Seeking Empirical Regularity and Theoretical Structure." *Rev Environmental Economics and Policy, 4 (1)*, 3-23.

[8] Diaz, H., & Quayle, R. (1980). "Heating Degree Day Data Applied to Residential Heating Energy Consumption." *J. of Applied Meterology, 3*, 241-246.

[9] Egli, H. (2004). "The Environmental Kuznets Curve - Evidence from Time Series Data for Germany." Economic Working Paper Series, Swiss Federal Institute of Technology Zurich, Center of Economic Reserach, Working Paper 03/28.

[10] Elhorst, J. (2003). "Specification and Estimation of Spatial Panel Data Models." *International Regional Science Review, 26*, 244-268.

[11] Elhorst, J. (2005). "Unconditional Maximum Likelihood Estimation of Linear and Log-Linear Dynamic Models for Spatial Panels." *Geographical Analysis , 37*, 85-106.

[12] Gill, P.E., Murray, W. and. Wright, M.H. (1982). *Practical Optimization*. Bingley, U.K.: Emerald Publishing Group.

[13] Grossman, G., and Krueger, A. B. (1991). "Environmental Impacts of North American Free Trade Agreement." *NBER Working Paper Series*, Working Paper No. 3914.

[14] Grossman, G., and Krueger, A. (1995). "Economic Growth and the Environment." *Quarterly Journal of Economics, 110* (2), 353-377.

[15] Hsiao, C. (2003). *Analysis of Panel Data (2nd ed.)*. Cambridge: Cambridge University Press.

[16] Hsiao, C., Pesaran, M., and Tahmiscioglu, A. (2002). "Maximum Likelihood Estimation of Fixed Effects Dynamic Panel Data Models Covering Short Time Periods." *J. of Econometrics, 109*, 107-150.

[17] Im, K., Pesaran, M., and Shin, Y. (2003). "Testing for Unit Roots in Heterogeneous Panels." *J. of Econometrics, 115*, 53-74.

[18] Institute for Energy (2010). Retrieved September 2010, from the Institute for Energy: http://www.instituteforenergyresearch.org.

[19] Levinson, A. (2008). "Environmental Kuznets Curve." In S. Durlauf, and L. Blume (Eds.), *New Palgrave Dictionary of Economics.* London: Palgrave Macmillan.

[20] Maddison, D. (2006). "Environmental Kuznets Curves: A Spatial Econometric Approach." *J. of Environmental Economics and Management, 51*, 218-230.

[21] Nickell, S. (1981). "Biases in Dynamic Models with Fixed Effects." *Econometrica*, 1417-1426.

[22] Ord, K. (1975). "Estimation Methods for Models of Spatial Interaction." *J. of the American Statistical Association*, 120-126.

[23] Perman, R., and Stern, D. (2003). "Evidence From Panel Unit Root and Cointegration Tests That the Environmental Kuznets Curve Does Not Exist." *The Australian J. of Agricultural and Resource Economics, 47* (3), 325-347.

[24] Rupasingha, A., Goetz, S., Debertin, D., and Pagoulatos, A. (2004). "The Environmental Kuznets Curve for US Counties: A Spatial Econometric Analysis with Extensions." *Papers in Regional Science, 83*, 407-424.

[25] Stern, D., and Common, M. (2001). "Is There an Environmental Kuznets Curve for Sulfur?" *J. of Environmental Economics and Management, 41* (2), 162-178.

[26] Swanson, B. (2005, September 7). "Answers archive: Heating and cooling degree days." Retrieved January 2010, from USA Today: http://www.usatoday.com.

[27] U.S. Bureau of Economic Analysis. (2010). *Gross Domestic Product by State*. Washington, DC: U.S. Department of Commerce.

[28] U.S. Bureau of Economic Analysis. (2011). *Gross Domestic Product by State*. Washington, DC: U.S. Department of Commerce.

[29] U.S. Census Bureau. (n.d.) "Population Estimates." Retrieved February 2010, from U.S. Census Bureau. http://www.census.gov/popest/states/states.html.

[30] U.S. Department of Energy (2010). *Electricity Transmission Fact Sheet*. Retrieved February 2010, from U.S. Department of Energy: http://www.eia.gov/.

[31] U.S. Energy Information Administration. (2008). *State Energy Data System*. Retrieved January 2010, from U.S. Energy Information Administration: http://www.eia.doe.gov.

[32] U.S. Energy Information Administration (2010) Independent Statistics and Analysis. Retrieved January 2010, from U.S. Department of Energy: http://www.eia.doe.gov.

[33] U.S. Environmental Protection Agency (2008). "Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2008."

[34] U.S. National Climate Data Center. (2010). "Heating and Cooling Degree Day Data." Washington, DC: National Oceanic and Atmospheric Administration.

Appendices

A Panel Unit Root Test

To determine which estimation procedure (fixed effects vs. dynamic fixed effects) is most appropriate for our dataset we need to determine if the main variables (CO_2 and GDP) are characterized by a unit root or higher order of non-stationarity. If we determine that both variables are non-stationary then we need to transform them, otherwise, the estimates may yield incorrect F and t stats. [24] According to Barbieri (2006) there are basically two types of panel unit root tests: the first type assumes cross-sectional independence, whereas the second type assumes cross-sectional dependence. Consistent with our spatial spillovers argument we have reason to believe that the data would be characterized by cross-sectional dependence. In other words, we believe that the economic activity within one state affects the economic activity of its neighbor or trading partner; i.e., there is dependence across the cross sections. In order to determine if there is cross-sectional independence we employ the parametric testing procedure proposed by Pesaran (2004). Using Pesaran's test (not shown) we reject the null hypothesis (at a one percent significance level) which offers evidence against the null hypothesis of cross-sectional independence. Thus, it would seem that the second type of panel unit root test would be more appropriate. Nevertheless, we offer the test results for both types of panel unit root tests in Table 1 below.

[Table A.1 will go about here.]

The test in second column of Table 1 is based upon the work of Im, Pesaran, Shin (2003) (Ipshin test) which assumes cross-sectional independence. The test in the third column is based upon Pesaran (2005) which assumes cross-sectional dependence; the "CADF" indicates a cross-sectional augmented Dickey-Fuller test. The null hypothesis of both tests assumes that the series is non-stationary (i.e., is characterized by a unit root).[25] The Ipshin test implies that CO_2 and GDP are stationary when just the lag term is included, but are non-stationary when the trend is

[24] Or we need to use a cointegrating model if the variables are cointegrated.

[25] The results of both tests were robust to tests with higher lags of variables, so we specify two lags out of convention that both variables are offered annually.

included. The Pesaran CADF test implies that CO_2 is non-stationary for both specifications, and GDP is non-stationary if two lags and a trend are included. Thus, there seems to be good evidence that CO_2 emissions are characterized by a unit root. The results for GDP are less certain, but intuition would lead one to believe that GDP is characterized by a unit root as well; in other words, preceding state-level GDP should have a significant effect on current state-level GDP. Lastly, taking the difference of the variables (represented by the sixth and seventh rows) implies stationarity with the exception of GDP. The Pesaran test implies stationarity for GDP when two differences were taken (not shown).

If CO_2 is characterized by a unit root and the rest of the variables are stationary then the dynamic spatial fixed effects estimator may be the most appropriate estimation scheme because the dynamic specification controls for the non-stationarity within the dependent variable. If CO_2 and GDP are both first-order difference stationary then the spatial first-difference estimator may be the most appropriate estimation scheme because differencing the data renders both variables stationary; there is marginal evidence in Table A.1 that both are first-order difference stationary. Since the test results of the Ipshin and Pesaran CADF are not conclusive we will estimate all four types of our proposed estimator; i.e., spatial fixed effects, spatial first differences, dynamic spatial fixed effects, and dynamic first differences. Offering the results for all four types presents a sensitivity analysis across our proposed estimation schemes.

B Bias of Elhorst (2003) Fixed Effects Estimator

To demonstrate how Elhorst's FE estimator is biased we restate equations (3.12) - (3.13) from above:

$$Qy = QZ\gamma + Qu \tag{B.1}$$
$$Qu = \lambda WQu + Q\epsilon. \tag{B.2}$$

where we used the demeaning operator to remove the fixed effect. We can now rewrite (B.1) and (B.2) as

$$\left(I_{NT} - \lambda W\right) Qy = \left(I_{NT} - \lambda W\right) Z\gamma + Q\epsilon. \tag{B.3}$$

Elhorst incorrectly assumes that

$$E(\epsilon\epsilon^{'}) = \sigma^2 I_{NT}, \tag{B.4}$$

when in actual fact

$$E(\epsilon\epsilon^{'}) = \sigma^2 Q, \tag{B.5}$$

where Q is an idempotent matrix such that $Q^{'}Q = Q$. Elhorst then constructs a maximum likelihood estimation scheme based upon assumption (B.4). If one tries to correct for assumption in

(B.4) by replacing it with (B.5) then the MLE procedure is no longer appropriate because it requires the inverse of Q to be calculated. Since Q is idempotent its inverse does *not* exist.[26] Therefore, Elhorst FE estimator is biased.

C Iterative Spatial First Difference Estimation Algorithm

Instead of using the demeaning operator to get rid of the fixed effects, one can alternatively first difference the data for (3.6) and (3.7) to obtain:

$$\Delta y = \Delta Z\gamma + \Delta u \tag{C.1}$$
$$\Delta u = \lambda W \Delta u + \Delta\epsilon \tag{C.2}$$

Based on first differencing the data, the new algorithm is as follows:

1. Compute the OLS residuals from (C.1) to derive $\Delta\tilde{u}$
2. Estimate λ from $\Delta\tilde{u} = \lambda W \Delta\tilde{u} + \Delta\epsilon$ by using the Newton-Raphson Method to derive $\tilde{\lambda}$
3. Construct new variables $\ddot{y} = \left(I_{NT} - \tilde{\lambda}W\right)\Delta y$ and $\ddot{Z} = \left(I_{NT} - \tilde{\lambda}W\right)\Delta Z$
4. Apply OLS of $\ddot{y}$ on $\ddot{Z}$ to yield $\tilde{\gamma}$
5. Construct the new residuals $\hat{\ddot{u}} = \ddot{y} - \ddot{Z}\tilde{\gamma}$ and return to Step 2 to calculate $\hat{\lambda}$
6. Construct the robust covariance for $\hat{\gamma}$ by calculating $\hat{A} = \frac{N}{N-K}\left(\ddot{Z}'\hat{\ddot{u}}\hat{\ddot{u}}'\ddot{Z}\right)$

The Newton-Raphson method to derive $\tilde{\lambda}$ is the same as listed in the Methodological Section.

D Alternative Specification Results

[Table D.1 will go about here.]

[26]The identity matrix is the only idempotent matrix that has an inverse that exists.

E Tables and Figures

Figure 1: 2001 Per−Capita CO2 Emissions (teragrams of carbon emitted)

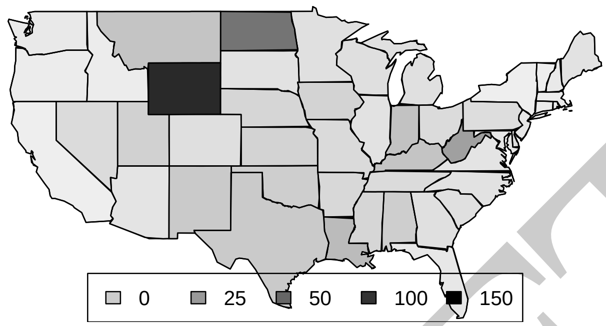

Table 1. Estimation Results for the Economic Growth-CO_2 Emissions Relationship (Quadratic Specification) with Inverse Distance-Based Spatial Weighting Matrix

	Model Types						
Explanatory Variables	**LSDV**	**Elhorst FE**	**Elhorst UML**	**SFE**	**SFD**	**Dynamic SFE**	**Dynamic SFD**
$CO_{2,t-1}$	N/A	N/A	0.9295*** (126.0493)	N/A	N/A	0.0877*** (3.2292)	0.0036 (0.7676)
GDP	12.6768*** (18.4822)	15.4993*** (11.0228)	1.3792** (2.2303)	12.6877*** (4.5095)	5.5944** (2.5124)	10.6052*** (5.6853)	5.2660* (1.9483)
GDP^2	-0.6196*** (-18.4207)	-0.7754*** (-11.1887)	-0.0652** (-2.1484)	-0.6201*** (-4.5379)	-0.2651** (-2.1167)	-0.5214*** (-5.7346)	-0.2492* (-1.8530)
CDD	-0.0179 (-0.8664)	0.3235*** (22.9872)	0.0112 (0.7139)	-0.0179 (-0.2195)	0.0128 (1.0132)	0.0108 (0.2137)	0.0134 (0.9988)
HDD	0.0692 (1.3415)	0.3493*** (16.5007)	0.049527 (1.2695)	0.0692 (0.4266)	0.1006*** (3.3496)	0.1369 (1.3964)	0.1026*** (3.2373)
Coal	0.0015*** (10.1853)	0.0037*** (26.4352)	0.0002 (1.5129)	0.0015 (1.5779)	0.0011 (1.2951)	0.0013** (2.2208)	0.0011 (1.2212)
Oil	-0.0008*** (-3.1082)	0.0036*** (25.4736)	0.0002 (0.9688)	-0.0008 (-0.6949)	0.0005 (0.5209)	-0.0007 (-0.9866)	0.0005 (0.5099)
λ	N/A	0.48 (0.9246)	0.01 (0.092)	0.2667 (0.0162)	0.3872 (0.0179)	0.4300 (0.0248)	0.3592 (0.0169)
R^2	0.9402	0.6028	0.9796	0.6028	0.6019	0.7347	0.7347
Adjusted R^2	0.9371	0.5894	0.9789	0.5932	0.5934	0.7291	0.7289
Robust SE	No	No	No	Yes	Yes	Yes	Yes

Notes: The numbers in the parentheses denote t-statistics. The superscripts "***", "**", "*" denote a significance level 0.01, 0.05, and 0.10 respectively. LSDV denotes the least squares dummy variable estimate. Parentheses for spatial autocorrelation coefficient (λ) represent estimates of the asymptotic standard errors.

Table 2. Estimation Results for the Economic Growth-CO_2 Emissions Relationship (Quadratic Specification) with First-Order Contiguity Spatial Weight Matrix

	Model Types						
Explanatory Variables	**LSDV**	**Elhorst FE**	**Elhorst UML**	**SFE**	**SFD**	**Dynamic SFE**	**Dynamic SFD**
$CO_{2,t-1}$	N/A	N/A	0.9294*** (124.5168)	N/A	N/A	0.1065*** (3.4544)	0.0034 (0.8414)
GDP	12.6768*** (18.4822)	15.0124*** (10.624)	1.3812** (2.2361)	6.0542 (1.6440)	1.9783 (0.7438)	9.4698*** (4.8915)	6.2262** (2.3872)
GDP^2	-0.6196*** (-18.4207)	-0.7514*** (-10.7837)	-0.0652** (-2.1541)	-0.2971 (-1.6664)	-0.0864 (-0.6569)	-0.4661*** (-4.9558)	-0.2956** (-2.2749)
CDD	-0.0179 (-0.8664)	0.3211*** (22.534)	0.0108 (0.6773)	-0.0055 (-0.0394)	0.0049 (0.2380)	0.0141 (0.2633)	0.0153 (1.2861)
HDD	0.0692 (1.3415)	0.3511*** (16.4965)	0.0469 (1.19)	-0.1024 (-0.4190)	0.0848* (1.8957)	0.1202 (1.1649)	0.1129*** (3.8957)
Coal	0.0015*** (10.1853)	0.0037*** (26.3352)	0.0002 (1.5069)	0.0010 (1.3111)	0.0007 (1.0497)	0.0012** (2.1647)	0.0012 (1.3184)
Oil	-0.0008*** (-3.1082)	0.0037*** (25.3857)	0.0002 (0.9659)	-0.0015 (-1.3035)	-0.0001 (0.0682)	-0.0006 (-0.8888)	0.0006 (0.5698)
λ	N/A	-0.974*** (-6.5299)	0.0484 (0.8870)	0.3782*** (5.8595)	0.3760*** (5.5658)	0.1225 (0.4136)	-0.2144 (N/A)
R^2	0.9402	0.6027	N/A	0.6028	0.6019	0.7347	0.7347
Adjusted R^2	0.9371	0.5893	N/A	0.5932	0.5934	0.7291	0.7289
Robust SE	No	No	No	Yes	Yes	Yes	Yes

Notes: The numbers in the parentheses denote t-statistics. The superscripts "***", "**", "*" denote a significance level 0.01, 0.05, and 0.10 respectively. LSDV denotes the least squares dummy variable estimate. Parentheses for spatial autocorrelation coefficient (λ) represent estimates of the asymptotic standard errors.

Table 3. Burridge LM Test Results

	Model Types			
Explanatory Variables	**SFE**	**SFD**	**Dynamic SFE**	**Dynamic SFD**
	0.1699 (0.8890)	0.0374 (0.8466)	4.2956 (0.0225)	1.0098 (0.2396)
The top number is the chi-squared statistic. The bottom number is the p-value of the statistic.				

Figure 2. Projected Plots Based Upon Estimation Results

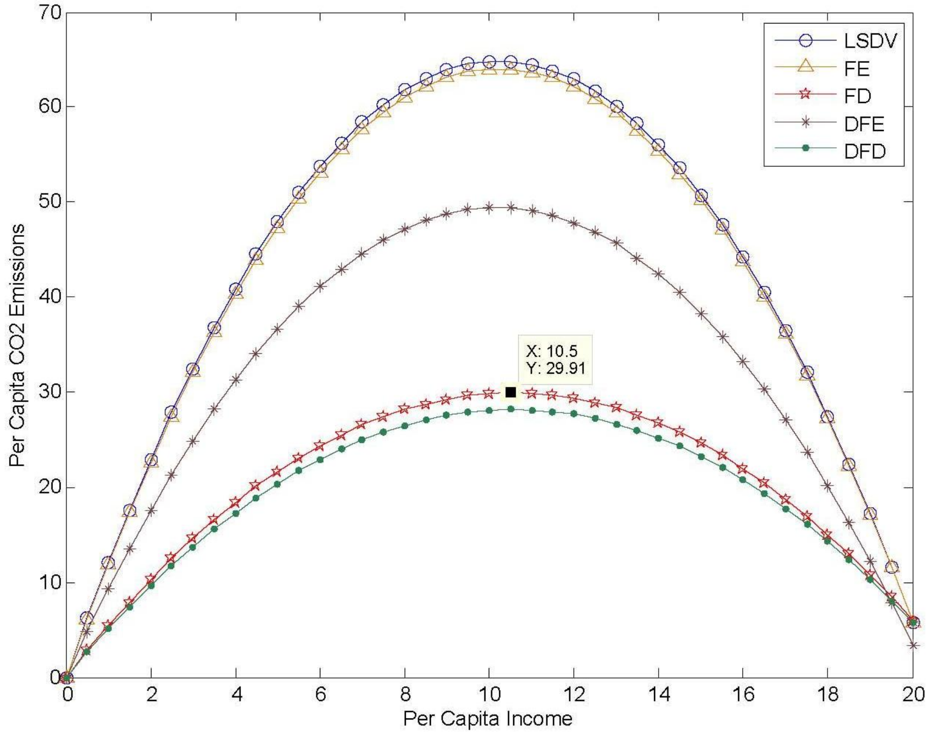

Table A.1. Panel Unit-Root Tests

Variable	Im, Pesaran, Shin Test	Pesaran CADF Test
CO_2 with two lags	-1.671 (0.070)	-1.744 (0.577)
CO_2 with two lags and trend	-1.969 (0.893)	-2.045 (0.990)
GDP with two lags	-1.764 (0.014)	-2.010 (0.038)
GDP with two lags and trend	-2.144 (0.420)	-2.456 (0.181)
D.CO_2 with two lags and trend	-3.775 (0.000)	-3.618 (0.000)
D.GDP with two lags and trend	-2.944 (0.000)	-2.399 (0.320)
Notes: The top number represents the t-bar value and the bottom number represents the p-value. The *D.** indicates a difference of variable was taken.		

Table D.1. Estimation Results for the Economic Growth-CO_2 Emissions Relationship (Cubic Specification, Inverse Distance)

	Model Types						
Explanatory Variables	**LSDV**	**Elhorst FE**	**Elhorst UMLE**	**SFE**	**SFD**	**Dynamic SFE**	**Dynamic SFD**
$CO_{2,t-1}$	N/A	N/A	0.9311*** (336.4948)	N/A	N/A	0.0880*** (3.0466)	0.0038 (0.7696)
GDP	11.2151 (0.5618)	120.2322*** (2.8865)	-3.0028*** (-11.5689)	12.1311 (0.1788)	0.6772 (0.0115)	-4.5107 (-0.0954)	-4.4811* (-0.0724)
GDP^2	-0.4748 (-0.2403)	-11.1048*** (-2.6956)	0.0349*** (16.5305)	-0.5650 (-0.0845)	0.2232 (0.0384)	0.9759 (0.2098)	0.7180 (0.1171)
GDP^3	-0.0048 (-0.0733)	0.3392** (2.4995)	-0.0140* (-11.2059)	-0.0018 (-0.0083)	-0.0161 (-0.0841)	-0.0494 (-0.3239)	-0.0320 (-0.1581)
CDD	-0.0180 (-0.8677)	0.3269*** (24.4464)	0.0108 (0.8546)	0.0691 (0.3694)	0.0128 (0.9623)	0.0105 (0.1956)	0.0135 (0.9455)
HDD	0.0691 (1.3401)	0.341*** (16.6541)	0.0457 (1.4645)	0.0015 (1.3692)	0.1006*** (3.1795)	0.1367 (1.3171)	0.1027*** (3.0595)
Coal	0.0015*** (10.1484)	0.0038*** (27.6229)	0.0002*** (1.9110)	-0.0008 (-0.6014)	0.0011 (1.2287)	0.0013** (2.1029)	0.0011 (1.1558)
Oil	-0.0008*** (-3.0934)	0.0037*** (26.6272)	0.0002 (1.2901)	-0.0038 (-0.0268)	0.0005 (0.4976)	-0.0007 (-0.9077)	0.0005 (0.4886)
λ	N/A	-0.074*** (-3.0762)	0.0728*** (0.8001)	0.2653 (0.0161)	0.3872 (0.0179)	0.3733 (0.0220)	0.3579 (0.0168)
R^2	0.9402	0.6041	N/A	0.6044	0.6038	0.7348	0.7348
Adjusted R^2	0.9371	N/A	N/A	0.5947	0.5951	0.7270	0.7288
Robust SE	No	No	No	Yes	Yes	Yes	Yes

Notes: The numbers in the parentheses denote t-statistics. The superscripts "***", "**", "*" denote a significance level 0.01, 0.05, and 0.10 respectively. LSDV denotes the least squares dummy variable estimate.