Paper submitted to the VII world conference of the Spatial Econometrics Association 2013

Forecasting U.S. State-Level Carbon Dioxide Emissions

J. Wesley Burnett* and X. Zhao†

Due to criticisms of potential identification issues within spatial panel data models, this study contributes to the literature by comparing forecasts of U.S. state-level carbon dioxide emissions against empirical reality using panel data models with and without spatial effects. From a policy standpoint, understanding how to predict emissions is important for designing climate change mitigation policies. From a statistical standpoint, it is important to test spatial econometrics models to see if they are a valid strategy to describe the underlying data. We find, unlike past studies, that non-spatial OLS estimators perform best in one to two-year-ahead forecasts, but the spatial panel data models perform better over a two to five year out-of-sample horizon.

Keywords: Dynamic, spatial panel data, Forecasting, Carbon dioxide emissions, United States

JEL Codes: C33, C53, Q50

* Assistant Professor, West Virginia University, wesley.burnett@mail.wvu.edu
† PhD Candidate, West Virginia University, xueting.zhao@mail.wvu.edu

1 Introduction

Understanding the spatial and temporal distribution of carbon dioxide (CO_2) emissions can aid policy in helping to develop proper regulation frameworks to mitigate harmful anthropogenic greenhouse gas (GHG) emissions, which arguably are the cause of global climate change. While the geographic distribution of CO_2 emissions does not affect the global climatic impact, the distribution of the sources of the emissions will be important for policy formulation at the international, national, and ultimately at the local level. Assuming that the United States, adopts an international multi-lateral agreement to mitigate emissions, such as the Kyoto Protocol, then it must begin to look inward to determine the major sources of emissions and how to reduce these emissions. As no viable technologies yet exist to mitigate CO_2 (i.e., in significant quantities at economical costs), the alternative is increase the price (through taxation or some market-oriented policy) of such sources to account for the externalities or social costs associated with climate change. Most sources of CO_2 come from energy-related activities, principally through coal-fired electricity generation and transportation (U.S. Energy Information Administration, 2012). Energy consumption is arguably an important component of economic growth, so increasing the cost of energy arguably comes at the expense of future potential economic growth. Such measures may be politically unpalatable, as they are in the United States. Therefore, understanding the subnational-level sources of emissions and spatial interactions of these sources across regions will be important for formulating national-level policies to mitigate GHG emissions.

One of the reasons perhaps that the U.S. has been slow to adopt a national mitigation scheme is due to uncertainty in state-level abatement costs. For example, if a state is currently a high emitter then arguably its marginal cost of reducing a unit of CO2 should be relatively low, whereas a low-emitting state arguably would have higher relative marginal costs for reducing another unit. Current policy regimes often ignore location and dispersion characteristics of the sources of emissions, and emissions are penalized at a single permit price (Fowlie & Muller, 2013). Fowlie & Muller (2013) argue that in the presence of uncertainty in abatement costs, differentiated policies may improve welfare. We will explore these differentiated policies in the context of clubs of states whose emissions are converging through time.

Global climate change is an international problem in scope, yet domestic or regional policies can be implemented to mitigate CO2 emissions. In the U.S., the federal government has not been able to successfully formulate a national climate change policy that includes some mechanism to reduce CO2 emissions, but various states have implemented programs. Renewable Portfolio Standards (RPSs), for example, have been adopted by thirty-three states and the District of Columbia as of 2009 (US Environmental Protection Agency, 2009). RPSs are goals or requirements for electric utilities or other retail electric providers to supply a specified minimum percentage of customer base load with electricity from various renewable energy sources. The goal of such programs is to not only develop sustainable forms of energy but also to reduce harmful greenhouse gases (GHGs) including CO2. Additionally, some groups of states have adopted regional programs such as the Regional Greenhouse Gas Initiative (RGGI).[1] According to Regional Greenhouse Gas Initiative (2012), it is the first market-based regulatory program in the U.S. to reduce GHG emissions with the explicit goal of reducing regional CO2 emissions from the power sector by ten percent by 2018. An understanding of the distribution of emissions can aid further states and regional initiatives in setting emission reduction goals and renegotiating emission obligations.

Spatial panel data models are a promising means to examine the spatial and temporal distribution of CO_2 emissions. There has been tremendous growth in the spatial econometric models literature over the past two to three decades. Spatial econometrics is an applied field of econometrics that deals with sample data that is collected with reference to location measured as points in space. What distinguishes spatial econometrics from traditional econometrics is that the locational data may be characterized by spatial dependence or spatial heterogeneity (LeSage and Pace, 2009). The idea of spatial dependence, or technically spatial autocorrelation, is similar to the concept of temporal autocorrelation found within the times series literature. As in time series, if this autocorrelation is present and unaccounted for then it could lead to biased or worse inconsistent regression estimates. Traditional econometrics had largely ignored spatial autocorrelation until the development of spatial econometrics. Recent advances in spatial econometrics have led to the development of longitudinal or panel data models that control for spatial autocorrelation. Longitudinal data are simply cross-section observations collected over time. These models offer the dual benefit of potentially controlling for province-level unobserved or heterogeneous fixed effects and spatial dependence.

[1] The state participants in RGGI include Connecticut, Delaware, Maine, Maryland, Massachusetts, New Hampshire, New York, Rhode Island, and Vermont.

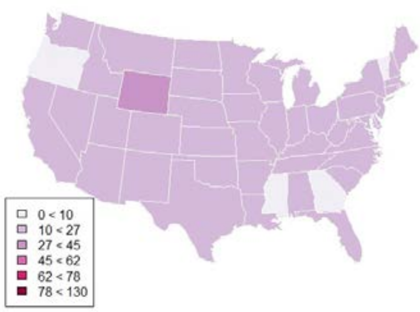
Despite the advances in this literature, spatial econometric models have come under criticism recently for problems associated with identification and for a lack of appeal to theoretical foundations (Partridge *et al.*, 2012). According to these criticisms, the problem of identification is similar to Manski's (1993) "reflection problem," where group average characteristics (neighboring province carbon dioxide emissions and structural characteristics) affect individual outcomes (local carbon dioxide emissions) but the parameters in the model are not identifiable. That is, it is hard to separate out the effects of what causes emissions locally versus what causes in neighboring provinces. We agree that these criticisms cast a bit of shadow on spatial econometric approaches, but rather than appealing to causality based upon correct model specification and/or the correct interpretation of parameter estimates, we instead appeal to an alternative validation strategy that is less dependent on prior theory. That is, we take these models as a black box and test them against empirical reality (Freedman, 1991). Against this background, we compare forecasts of province-level carbon dioxide emissions against empirical reality using panel data models with and without spatial effects.

This study contributes to the literature by offering an assessment of how the spatial panel data models perform in forecasting against non-spatial panel data models in a root mean square error context. We compare the performance of several predictors for province-level CO2 emissions for one through five-year-ahead forecasts. Based on forecast performance we find, unlike Baltagi and Li (2006), that a non-spatial OLS estimator performs best for forecasts one to two years ahead. In contrast, our results suggest that spatial panel estimators yield a better forecasting performance for two to five-year-ahead forecasts. These findings may suggest that models that take into account spatial autocorrelation (and heterogeneity) provide the best "within sample" fit to the data and good medium- to long-fun "out of sample" prediction ability.

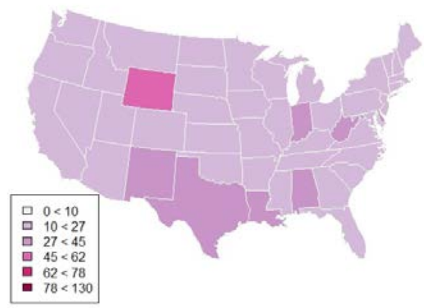
It is difficult to compare total carbon dioxide emissions across States because of the variation in their sizes, so we analyze state-level, per-capita emissions. Per-capita measures normalize emissions across States to offer a more compatible apples-to-apples comparison. Further, per-capita emissions offer a truer picture of how wasteful regions are. For example, China is the largest aggregate emitter of CO2 emissions but the U.S. is the largest emitter per capita (International Energy Agency, 2011). From a policy sense, an analysis of per-capita emissions offers a more equitable measure for negotiating multilateral agreements. The structural and non-structural factors we examine are climate, population density, income per capita, the percentage of electricity from coal, and the percentage of electricity used in the industrial sector.

Chloropleth maps of state-level (aggregate) emissions for 1960-2009 (by decade) are offered in Figures 1 and 2. Each map's scale is based on the bins of per-capita emissions in the year 2009. These maps show a general increase in per-capita, state-level emissions, but a gradual easing of intensities in some states starting after 2000.

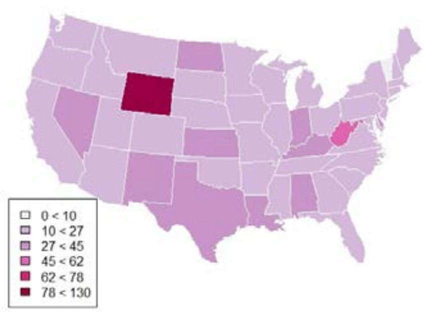
Figure 1. Per-Capita CO2 Emissions, 1960-1980



(a) 1960

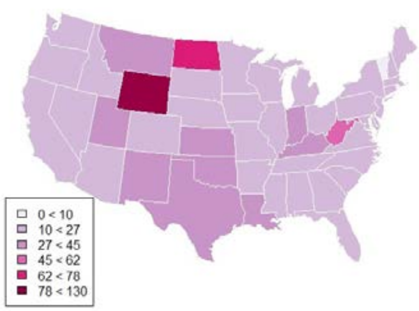


(b) 1970

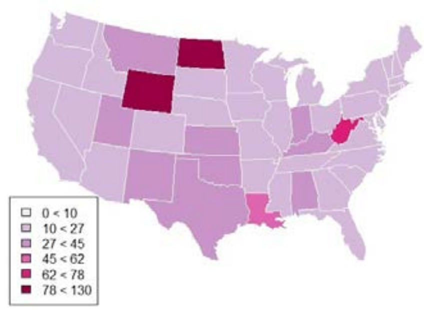


(c) 1980

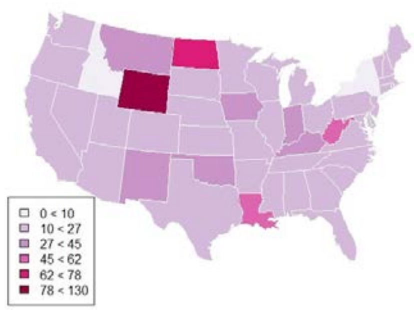
Figure 2. Per-Capita CO2 Emissions, 1990-2009



(a) 1990



(b) 2000



(c) 2009

2 Background

Until fairly recently, few papers had examined the forecasting ability of spatial econometric models – exceptions include Baltagi and Li (2006), Baltagi *et al.* (2012), Elhorst (2010), Kelejian and Prucha (2007), and Kelejian and Robison (2000), among others. However, there does seem to be a growing interest in the alternative validation strategy of prediction.

A similar line of literature includes estimating dynamic, spatial panel data models. These models push the econometrics to consider not only spatial dependence but also temporal dependence; hence, they are often called spatio-temporal panel data models. Giacomini and Granger (2004) arguably offered the seminal paper in this literature. Similar to this study, a few papers have used this methodology to examine the sub-national forecasts of carbon dioxide emissions (Auffhammer and Carson, 2008; Auffhammer and Steinhauser, 2007; Auffhammer and Steinhauser, 2012).

Unlike these studies, however, our focus is on forecasting standard spatial panel data models. That is, we do not consider temporal dependence in addition to spatial dependence, but rather focus on forecasting models that consider only the latter. Principally, we do not consider the dynamic models because as Nickell (1981) pointed out, including a temporal lag of the dependent variable in a panel data context may lead to biased coefficient estimates – this is sometimes referred to as dynamic panel data bias. Judson and Owen (1999), using Monte Carlo techniques, showed that a least-squares dummy variable model (applied to a dynamic panel data) reduces the bias of the coefficient estimates when there are sufficient time observations of the cross-sections. However, in our context, we have only thirty-five years of observations to regress within sample to compare five year-ahead forecasts against empirical observations out of sample. The relatively small number of time observations within sample (thirty-five years) raises questions about whether we have sufficient time series observations to estimate a dynamic panel data model without introducing bias. Therefore, we abstract away from the dynamic models and instead focus on the standard panel data models, both with and without spatial effects.

3 Methodological Approach

To control for state-level independent effects we propose the following fixed-effects model:

$$y = X\beta + (\iota_t \otimes I_N)\mu + (I_T \otimes \iota_N)\eta + u, \tag{1}$$

where y denotes a *(NT x 1)* vector of U.S. state-level per capita carbon dioxide emissions. X is an *(NT x K)* matrix of the explanatory variables including energy prices, per capita GDP, and climate variables. All the terms are represented in natural logs so that estimated coefficients can be interpreted as elasticities. The coefficient μ denotes the individual effect (or heterogeneity) for each U.S. state and η denotes the time effect. The time fixed effects control for shocks that occur to all states simultaneously through time; an example of such shocks include the oil crises in the 1970s and the amendment to the clean air act in the early 1990s. ι_N denotes a column vector of ones of length N, or T if the subscript "T" is indicated. I_N denotes an identity matrix of dimensions *(N x N)*, or dimensions *(T x T)* if the subscript "T" is indicated. 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. If we allow for the fixed effects term to enter into the error term and we estimate (1) without controlling for it, then the estimates will result in omitted variable bias (i.e., if the fixed effect is correlated with the explanatory variables). To control for fixed effects we can either (1) estimate μ and η directly in the model by creating dummy variables for these parameters as in a least squares dummy variable (LSDV) model; or (2) we could demean the data as in a fixed effects (FE) or within estimator. The FE estimation method is usually preferred to the LSDV model due to the incidental parameter problem of having to estimate parameters for the covariates, the fixed effects, and time-period effects terms.

The FE estimation method is carried out by demeaning each of the variables. Demeaning the variables along the temporal dimension eliminates the heterogeneous fixed effects. This is carried out by transforming the data as follows

$$z_{it} = z_{it} - \frac{1}{T}\sum_{t=1}^{T} z_{it}, \tag{2}$$

where z_{it} denotes any of the variables (dependent or independent) within the study. Additionally, demeaning the variable along the cross-sectional dimension eliminates the time fixed effects, which is transformed as follows

$$z_{it} = z_{it} - \frac{1}{N}\sum_{i=1}^{N} z_{it}. \tag{3}$$

Finally, both the heterogeneous and time fixed effect can be eliminated by transforming the data as follows

$$z_{it} = z_{it} - \frac{1}{T}\sum_{t=1}^{T} z_{it} - \frac{1}{N}\sum_{i=1}^{N} z_{it} + \frac{1}{T \cdot N}\sum_{t=1}^{T}\sum_{i=1}^{N} z_{it}. \tag{4}$$

An alternative assumption in (1) is that state-level individual effects, μ, are not fixed but rather are unobserved "random" variables which follow a probability distribution (usually normal) with finite parameters – this is referred to as the random effects (RE) estimator. The key difference between these estimators is in the assumption of orthogonality of μ. If μ is assumed to be uncorrelated with the explanatory variables then RE is the appropriate estimator; conversely, if μ is assumed to be correlated with the explanatory variables then FE is the appropriate estimator. The Hausman specification test statistic can be used to determine which estimator provides a better fit of the data (Hausman, 1978; Lee & Yu, 2010b). This test statistic is asymptotically distributed as chi-squared (χ^2) with K degrees of freedom, where K is the number of explanatory variables. Another way to test the random effects estimator against the fixed effects estimator is to estimate the "phi" parameter as outlined in Baltagi (2005). If this parameter equals zero, the random effects estimator converges to its fixed effects counterpart (Elhorst, 2010).

In this study we will assume that the state-level individual effects are correlated with the explanatory variables, and therefore our primary focus will be on the FE estimation procedure. However, for the sake of completeness we will estimate both the FE and RE models and then use the diagnostic tests discussed in the previous paragraph to determine which model provides a better fit to the data.

Given the assumption of fixed effects, there are many different avenues to estimate the parameters in (3.1). One, we can assume homogeneity of parameters across states, pool the data, and estimate a single demand equation by ordinary least squares (i.e., omitting the fixed effects and time effects terms). OLS is consistent if the disturbances are orthogonal to the right-hand side (RHS) variables. Two, allow for a limited degree of heterogeneity in time invariant unobservables by adopting a fixed effects estimator – this approach still assumes that all coefficients are identical across states. Three, allow all the coefficients to vary across states. Under this assumption, one could potentially estimate the equations state by state which could result in imprecise estimated coefficients due to the short time series for any given state.

Baltagi & Griffin (1997) explore a large number of estimators, including an instrumental variables estimator, and compare the plausibility of the estimators for a dynamic demand model for gasoline in 18 OECD countries. They found that the pooled estimators yield the most plausible estimates. Therefore, we proceed by assuming homogeneity of the parameters across states, pool the data, and estimate a single demand equation.

We specify a fixed effects model to control for possible endogenous characteristics of the individual states within the study – these are characteristics that do not change (or change very little) over time such as unobservable geographic characteristics. The time-period effects control for time-specific shocks that may affect per capita energy consumption in all states such as oil shocks, recessions, and federal policies applicable to all states (Aroonruengsawat *et al.*, 2012). An example of such a federal policy is the second amendment to the Clean Air Act of 1990.

The contribution of this paper is to not only consider state-level unobserved heterogeneity and time-period effects, but to extend the model in (1) to consider how economic distance may affect this relationship. We introduce spatial effects into the model by using a standard (pre-specified and non-negative) spatial weighting matrix, W_N, as an $(N \times N)$ positive matrix where the rows and columns correspond to the cross-sectional observations (contiguous 48 states). An element of the weighting matrix, w_{ij} , expresses the prior strength of interaction between state i and state j. Since we are dealing with a spatial panel, the weights are extended to the entire panel as

$$W_{NT} = I_T \otimes W_{N,} \tag{4}$$

where I_T again denotes the identity matrix of dimension T and $\otimes$ denotes the kronecker product.

3.1 Spatial, Panel Data Models

We extend the non-spatial model, equation (1), above to incorporate spatial autocorrelation. The three potential types of spatial panel specifications we consider are the: spatial lag, spatial error, and spatial Durbin models. The spatial lag or spatial autoregressive model (SAR) can be expressed in matrix form as

$$y = \delta(I_T \otimes W_N)y + X\beta + (\iota_T \otimes I_N)\mu + (I_T \otimes \iota_N) + u, \tag{5}$$

where δ is the spatial autoregressive coefficient.[2] We have suppressed all subscripts for cross-sections (i) and time (t) for each of exhibition.

The spatial error model (SEM), on the other hand, can be expressed as

$$\begin{aligned} y &= X\beta + (\iota_T \otimes I_N)\mu + (I_T \otimes \iota_N) + u \\ u &= \rho(I_T \otimes W_N)u + \varepsilon, \end{aligned} \tag{6}$$

[2] If we assume fixed effects then we denote the model as a spatial autoregressive model with fixed effects (SAR FE). If we assume random effects then we denote this model as SAR RE.

where u reflects the spatially auto-correlated error term and ρ denotes the (scalar) spatial autocorrelation coefficient on the error term.

Finally, the (unrestricted) spatial Durbin model (SDM) is specified as

$$y = \delta(I_T \otimes W_N)y + X\beta + \gamma(I_T \otimes W_N)X + (\iota_T \otimes I_N)\mu + (I_T \otimes \iota_N) + u, \tag{7}$$

where the parameters are the same as before but the parameter γ now indicates a spatial autocorrelation coefficient on the explanatory variables.

Hereafter, each of the spatial data models will be analyzed by the fixed effect estimation method outlined above, unless stated otherwise. In other words, the spatial data models will be estimated by transforming the data to deviations in means to eliminate the heterogeneous fixed effects, the time fixed effects, or both.

The spatial Durbin model can be used to determine if the model can be simplified to a spatial lag model or a spatial error model because the models nests dependence in both the disturbances and the dependent variable (LeSage & Pace, 2009). The two null hypothesis tests for determining the correct spatial model are: $H_0 : \gamma = 0$ and $H_0 : \gamma + \delta \cdot \beta = 0$ (LeSage & Pace, 2009; Elhorst, 2009). The first hypothesis determines if the spatial Durbin can be simplified to the spatial lag model whereas the second hypothesis determines if it can be simplified to the spatial error model. The second hypothesis stems from the fact that the spatial Durbin model is the reduced form of the spatial error model. If both hypotheses are rejected then the spatial Durbin model provides the best fit for the data.

To further test if a spatial effects model outperforms a model without any spatial interaction effects, one may use Lagrange Multiplier (LM) tests for a spatially lagged dependent variable and for spatial error autocorrelation – these tests contain robust counterparts as well (Debarsy & Ertur, 2010).[3] If a spatial lag model and a spatial error model are estimated separately then likelihood ratio (LR) tests can be conducted to determine which model provides the best fit for the data. The LR tests can also be complemented with Wald tests (Elhorst, 2010).

[3] The LM test statistics only require estimation of a non-spatial model associated with the null hypothesis that the spatial autocorrelation coefficient is equal to zero (LeSage & Pace, 2009).

3.2 Forecasting

Once we have determined which model provides the best fit to the data within the entire sample, we will next proceed by determining which model provides the best to the data out of sample. In order to carry out the forecasts we shorten the within sample observations by omitting the last five years of observations (2005-2009). We now define these final five years of observations as the out-of-sample observations. First we run the regressions on the now shorter within-sample, and then forecast the various models against empirical reality to see which model provides the best fit to the data out-of-sample. To evaluate which model provides the best out-of-sample fit we need a metric to compare the models.

Three common metrics used to evaluate forecast accuracy are the: ME (mean error), MAE (mean absolute error), and RMSE (root mean square error), which are defined as

$$ME = \sum_{t=1}^{T}\sum_{i=1}^{N}\frac{1}{N \cdot T}\left[F(t) - A(t)\right] \tag{8}$$

$$MAE = \sum_{t=1}^{T}\sum_{i=1}^{N}\frac{1}{N \cdot T}\left|F(t) - A(t)\right| \tag{9}$$

$$RMSE = \left\{\sum_{t=1}^{T}\sum_{i=1}^{N}\frac{1}{N \cdot T}\left[F(t) - A(t)\right]^{2}\right\}^{1/2}. \tag{10}$$

The symbol *T* denotes the total number of time periods and *N* denotes the total number of states within each cross-section. The symbol *F(t)* denotes the forecasted value and *A(t)* denotes the actual empirical observation. The difference between the forecasted value and the empirical observation denotes the forecast error. Therefore, the smaller the forecasted error, the better the model predicts future values. The errors in the RMSE metric are squared before averaging, so the RMSE gives a relatively higher weight to large errors. Based upon this construction, the RMSE arguably offers the most severe penalty for forecasting error. Thus, we will concentrate on the RMSE as our metric in the current study.

There are two principal types of forecasts. The first type of forecasts compares predicted values to actual observations one year at a time. These one-step-ahead forecasts are conducted by evaluating the regression for the entire initial within-sample observations (in our case 35 years) and then forecasting one year in advance. In the next iteration a regression is conducted on the within-sample which has been expanded by an additional year (36 years) and then forecasted a year in advance. The process is repeated until the one-step-ahead forecasts are available for comparison against the entire initial out-of-sample observations (five years). The second type of forecasts compares the predicted values to actual observations over the entire out-of-sample period. The five-year-ahead forecasts are conducted by regressing the model on the entire initial within-sample (35 years) designation, and then forecasting over the entire out-of-sample period (five years) using the empirical observations of the independent variables within the out-of-sample period. The one-step-ahead forecasts provide a metric for evaluating the short-run predictive ability of the model. The five-year-ahead forecasts, on the other hand, provide a metric for evaluating the medium- or long-run predictive ability of the model. To compare each of the models we will evaluate both types of forecasts.

4 Data

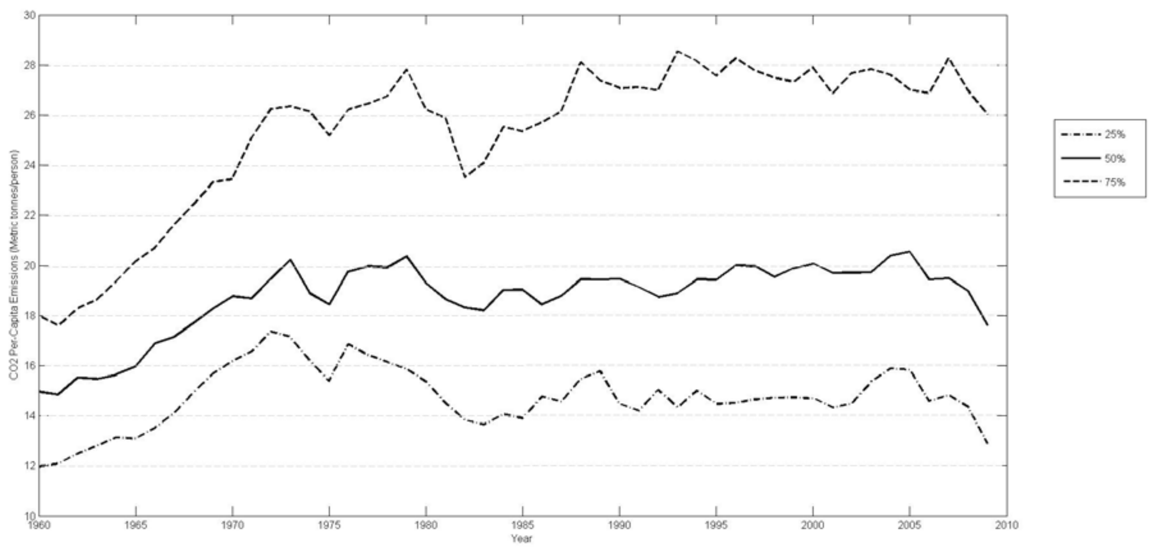
In this paper we analyze the relationship between energy consumption, economic activity, and pollution emissions while controlling for potential spatial effects within the data. The pollution variable, carbon dioxide (CO2), examined in this paper is estimated by the Department of Energy (DOE) based upon the conversion of fossil fuels to their final energy use; e.g., the conversion of coal into electrical energy in a power plant generates emission gases as a byproduct of the combustion process. In other words, CO2 emissions are estimated based upon a state's observed energy use. Therefore, CO2 emissions are not to be confused with actual CO2 pollution emissions that are emitted from the end of a smokestack or tailpipe.[4]

The energy-related carbon dioxide data for this analysis were obtained from the Carbon Dioxide Information Analysis Center (CDIAC) within the U.S. Department of Energy (Blasing *et al.*, 2004; US Energy Information Administration, 2012). CDIAC estimates the emissions by multiplying state-level coal, petroleum, and natural gas consumption by their respective thermal conversion factors. Therefore, the data is based on estimates of CO2 emissions and not actual atmospheric emissions. Despite this deficiency, this measure of emissions is one of the more common used in the literature as it is difficult to measure atmospheric emissions of carbon dioxide. The energy emission estimates are extended by using more recent calculations of energy-related carbon dioxide emission (2000-2009) offered by the US Energy Information Administration (2012). The U.S. Energy Information Administration (EIA) calculates emissions identically to the CDIAC, however, we visually inspected the data to ensure that new emission estimates are consistent with the previous estimates. The estimates are offered in units of a million metric tonnes for the forty-eight contiguous states excluding the District of Columbia.

[4] The CO2 data should not be confused either with atmospheric CO2 pollution, which following emission enters the upper atmosphere and is more global in scope.

A plot of the quartiles of aggregate U.S. per-capita CO2 emissions for the period 1960-2009 is offered in Figure 3. Figure 3 demonstrates that aggregate emissions rose through the 1960s and 1970s, and then fell relatively sharply in the latter 1970s due to the second oil crisis. Emissions rose slowly over the next two decades and then appear to level off starting in 2000. The sharp drop in 2008 is a reflection of the global recession.

Figure 3. U.S. Aggregate Per-Capita CO2 Emissions, 1960-2009



Itkonen (2012) offers the following simple explanation of how the energy emissions are estimated. The CDIAC and EIA define carbon dioxide emissions as a linear function of fossil fuel combustion and cement manufacturing. The amount of CO2 emissions is determined by the chemical composition of the fuel source. Emission estimates are calculated by multiplying the amount of fuel usage by a constant thermal conversion factor as determined by the chemical properties of the fuel. Therefore, CO2 emissions are a linear combination of the usage of oil, E_t^{oil}, solid fuels such as coal, E_t^{coal}, natural gas, E_t^{gas}, and emissions from cement manufacturing, S_t. Formally, this is expressed as

$$CO_{2,t} \equiv \alpha_{oil} \cdot E_t^{oil} + \alpha_{coal} \cdot E_t^{coal} + \alpha_{gas} \cdot E_t^{gas} + \alpha_{flare} \cdot E_t^{flare} + S_t, \qquad (11)$$

where α_{oil}, α_{coal}, α_{gas}, $\alpha_{flare} > 0$ are the related thermal conversion factors.

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.

To model climatic influences on energy demand we use Cooling Degree Days (CDD) and Heat ing Degree Days (HDD), which were obtained from the National Climate Data Center within the National Oceanic and Atmospheric Administration (National Climate Data Center, 2010). 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 CO2 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 CO2 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.

Energy prices were obtained from the EIA (Energy Information Administration, 2008). The energy prices represent state-level annual average prices of coal, natural gas, and oil. The prices were converted to real values by again using the BEA's implicit price deflator – this ensures that the index used to convert nominal to real values is consistent with that of state-level GDP.

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.

The descriptive statistics for the variables are offered in Table 1.

Table 1. Descriptive Statistics

Variables	Max	Min	Mean	Median	Std Dev
CO2	131.11	7.71	23.72	19.46	16.49
Coal price	7.86	0	2.19	1.99	0.9425
Elec price	49.43	7.63	25.22	24.11	7.36
Nat gas price	15.87	1.18	6.80	6.54	2.63
Oil price	24.62	5.43	11.73	10.56	3.54
GDP	64,576	13,483	30,050	28,522	8793
CDD	3875	80	1085	867	779.38
HDD	10,745	400	5243	5381	2049.2

Note: CO2 and GDP represent per-capita values. CO2 is measured in metric tonnes. GDP and the price data are measured in real USD. Prices are measured as USD per unit of btu.

5 Empirical Estimation and Results

Consistent with the outline in the Methodological Approach Section, we begin by first determining which panel data model (spatial versus non-spatial) seems to provide the best fit to the entire sample of data. As the spatial models have recently been criticized for being fundamentally unidentifiable, we proceed by comparing the forecasting ability of the different panel data models. That is, we use the same models to compare one-step-ahead and five-year-ahead forecasts to determine which model provides the best fit to the data out-of-sample.

5.1 Entire Sample Testing and Diagnostics

The empirical model we use for the current study is specified as follows

$$\begin{aligned}\ln(y_{it}) = {} & \beta_0 + \beta_1 \ln(p_{it}^{c}) + \beta_2 \ln(p_{it}^{ng}) + \beta_3 \ln(p_{it}^{o}) + \beta_4 \ln(p_{it}^{e}) \\ & + \beta_5 \ln(GDP_{it}) + \beta_6 \ln(GDP_{it})^2 + \beta_7 \ln(CDD_{it}) \\ & + \beta_8 \ln(HDD_{it}) + \mu_i + \eta_t + u_{it} \\ & i = 1,\ldots,N; \quad t = 1,\ldots,T\end{aligned} \tag{12}$$

where y_{it} is real per-capita energy emissions in state i at time t. GDP_{it} denotes real per-capita state-level GDP. CDD_{it} denotes cooling degree days, whereas HDD_{it} denotes heating degree days. The variables p_{it}^{c}, p_{it}^{ng}, p_{it}^{o} and p_{it}^{e} denote state-level prices of coal, natural gas, crude oil, and electricity respectively. The scalar parameters on the price yield measures of price elasticities of demand. Consistent with economic theory (the law of inverse demand) we predict all these parameters to be negative. We assume fixed state-specific effects, μ_i, and time-period effects are denoted by η_t. The observations in (12) are available in the 48 contiguous states from 1970–2009 so that T = 40 and N = 48.[5]

Equation (12) is a reduced-form model for energy demand (Ryan & Ploure, 2009) with a simple extension of adding the quadratic polynomial expression of GDP, and adding the climatological variables (CDD and HDD). Without the squared term of GDP, the model is very similar in nature to that of Aroonruengsawat *et al*. (2012). There is a potential problem of price endogeneity in (12) due to an increasing block price structure in electricity – this may lead to an upward-bias in the estimate of demand response (Hanemann, 1984). A common way to deal with this problem is to estimate marginal prices instead (Berndt, 1996), but such data is unfortunately unavailable for our sample, so following Baltagi *et al*. (2002) and Maddala *et al*. (1997) we will assume price exogeneity. We also abstract away from issues related inter-fuel substitution which have implications for price interaction effects.

In order to determine which type of model (spatial vs. non-spatial) best fits the data, we begin our investigation by testing several different model specifications. This testing procedure is a mixture between a specific-to-general approach and general-to-specific approach (Elhorst, 2010). The procedure begins by testing the non-spatial model against the spatial lag and spatial error models. If the non-spatial models are rejected, the spatial Durbin model is tested to determine if it can be simplified to either the spatial lag or spatial error model – this step seeks corroborating evidence from the first step.

The estimation results for the non-spatial panel data models are reported in Table 2. Columns one through four represent estimation results given the specification of: pooled OLS (no fixed or time-period effects), fixed effects only (no time-period effects), time-period effects only (no fixed effects), and both fixed effects and time-period effects, respectively. To investigate the null hypothesis that the fixed effects and time-period effects are jointly insignificant, we performed a likelihood ratio test. The LR test for the joint insignificance of the fixed effects was rejected at the one percent level (4347.0321, 48 degrees of freedom, $p < 0.01$). Likewise, the LR test for the joint insignificance of the time-period effects was also rejected

[5]The data set was limited to the years 1970–2009 because those are the only years for which the EIA has data on average annual, state-level energy prices.

(182.4867, 40 degrees of freedom, $p < 0.01$). These results justify the extension of the model with fixed effects and time-period effects.

Table 2. Estimation Results without Spatial Interaction Effects

Dependent variable: CO_2	Pooled OLS	Fixed Effects Only	Time-Period Effects Only	Fixed and Time Fixed Effects
Coal price	-0.2840***	0.0095	-0.4247***	0.0202
	(-9.3775)	(0.7603)	(-14.0733)	(1.4637)
Electricity price	-0.3198***	-0.3506***	-0.2339***	-0.2600***
	(-8.3752)	(-16.7574)	(-6.4689)	(-11.2087)
Gas price	-0.2695***	0.0529***	-0.4332***	0.0347*
	(-7.3393)	(3.8774)	(-10.3009)	(1.6526)
Oil price	0.2288***	-0.0269*	-1.9628***	-0.3376***
	(5.3271)	(-1.7390)	(-16.6450)	(-5.5978)
CDD	0.2389***	0.0283*	0.2291***	0.0008
	(11.6404)	(1.7703)	(12.5020)	(0.0435)
HDD	0.2906***	0.1281***	0.3889***	0.1654***
	(9.9685)	(3.4137)	(14.3919)	(3.6262)
GDP	11.5203***	8.1988***	0.8412	7.5689***
	(5.9809)	(12.4428)	(0.4202)	(10.4054)
GDP^2	-0.5579***	-0.4009***	-0.0444	-0.3799***
	(-5.9772)	(-12.5228)	(-0.4585)	(-10.8477)
σ^2	0.1493	0.0132	0.1157	0.0120
R^2	0.3319	0.2099	0.4780	0.1997
FE R^2		0.9408	0.4820	0.9462
Log-like	893.8655	1432.8	649.5129	1524.0
LM spatial lag	5.6619**	241.9989***	0.08413	138.6038***
LM spatial error	21.8779***	206.9125***	3.6114*	125.0031***
Robust LM spatial lag	28.0865***	35.5800***	15.3873***	13.7217***
Robust LM spatial error	44.3025***	0.4936	18.9146***	0.1210

Note: All variables are in natural logarithms. The symbol * denotes $p < 0.1$; ** denotes $p < 0.05$; and *** denotes $p < 0.01$. The test statistic for the LM test is based on a chi-squared distribution with one degree of freedom.

Recall that if the state-level fixed effects term is correlated with the explanatory variables but it is not controlled for within the model then OLS estimates will result in omitted variable bias (OVB). The pooled OLS estimates (column one of Table 2) for all the coefficients in the model are all highly statistically significant ($p < 0.001$) which arguably results from the OVB. Given the joint significance of the fixed and time-period effects from the LR test we focus on the estimation results in column four of Table 3. The estimated coefficient on electricity and oil prices conform with expectations as the coefficients are negative and indicative of inelastic response of emissions to price changes. The coefficient on gas prices do not conform with expectations but this coefficient in only marginally significant. The coefficient on heating degree days is positive (consistent with expectations) and highly statistically significant. The coefficients on the polynomial of income terms are also high statistically significant and conform with the EKC hypothesis of the inverted U-shaped relationship between emissions and income.

All of the non-spatial models may suffer from misspecification if spatial dependence exists within the data. To test for the presence of spatial dependence we begin by conducting the classical LM tests (Anselin *et al.* , 2008; Burridge, 1980). These test results are listed in the bottom part of Table 2. For the classical LM tests (labeled "LM spatial lag" and "LM spatial error"), the hypothesis of no spatially lagged dependent variable and the hypothesis of no spatially auto-correlated error term is strongly rejected for each of the specifications with the exception of model including time-period effects alone (although the hypothesis of no spatially auto-correlated error is rejected at the ten percent significance level with this specification). Examining these tests' robust counterparts (Debarsy & Ertur, 2010), both hypotheses are rejected for the pooled OLS and time-period effects specification, but only the hypothesis of no spatially lagged dependent variable can be rejected when fixed effects are introduced into the specification (i.e., columns two and four). These results imply that a model specification with a spatially lagged dependent variable may be favored over a non-spatial model since we find consistent rejection of the hypothesis of no spatially lagged dependence. A model specification with spatially auto-correlated error terms is questionable as the LM tests offer mixed results.

To further test which spatial model specification is appropriate we estimate a spatial Durbin model and then conduct likelihood ratio tests (Burridge, 1981) as outlined in the Methodological Approach Section. The results for the spatial Durbin model are listed in column one of Table 3. These results reflect estimation of the SDM with the bias correction (Elhorst, 2010; Lee & Yu, 2010a) – results without the bias correction are almost identical. The LR and Wald test statistics are reported in Table 4, and as noted by LeSage & Pace (2009) yield similar results. According to these tests, the first hypothesis ($H_0 : \gamma = 0$) cannot be rejected, which suggests that the spatial lag model is the most appropriate specification. The test results indicate that the second hypothesis ($H_0 : \gamma + \delta \cdot \beta = 0$) is rejected at the five percent level, which implies that the spatial error model is not appropriate. These LR and Wald test results are consistent with our LM test results in Table 2, which indicate that the spatial lag model is the most appropriate specification for this relationship. Therefore, we report the results for the spatial lag model with fixed and time-period effects (SAR FE) in column two of Table 3. For completeness we also report the estimation results for the spatial autoregressive model with random and time-period effects (SAR RE) in column three.

Table 3. Estimation Results with Spatial Interaction Effects

Dependent variable: CO_2	SDM	SAR FE	SAR RE
δ	0.3589***	0.3586***	0.3940***
	(13.4472)	(14.1340)	(15.5956)
Coal price	0.0101	0.0120	0.0050
	(0.7209)	(0.8940)	(0.4179)
Electricity price	-0.1964***	-0.2040***	-0.2264***
	(-7.0207)	(-8.9620)	(-10.9728)
Gas price	-0.0324	-0.0111	0.0700***
	(-1.3058)	(-0.5409)	(5.5626)
Oil price	-0.2200***	-0.2418***	-0.0518***
	(-3.2676)	(-4.1076)	(-3.6709)
CDD	0.0365	0.0056	0.0015
	(1.0361)	(0.3075)	(0.0993)
HDD	0.1300	0.1290***	0.0506
	(1.4936)	(2.9060)	(1.4502)
GDP	7.0360***	7.0278***	0.4680***
	(8.9879)	(9.9384)	(6.4000)
GDP^2	-0.3540***	-0.3519***	-0.0254***
	(-9.4569)	(-10.3323)	(-6.4619)
ϕ			0.0416***
			(6.9339)
σ^2	0.0113	0.0113	0.0123
R^2	0.9516	0.9513	0.9445
Log-like	1594.5733	1589.4405	1585.7626

Note: All variables are in natural logarithms. The coefficients for the spatially-weighted dependent variables in the SDM have been omitted for ease of exhibition. The symbol * denotes $p < 0.1$; ** denotes $p < 0.05$; and *** denotes $p < 0.01$.

To test the assumption of whether fixed effects provides a better fit to the data than random effects we conduct a Hausman specification test. The results of the Hausman test are listed in Table 4 – this test suggests that the random effects model is rejected at a one percent significance level. In addition to the Hausman test, we estimate the “phi” parameter from Baltagi

(2005). If this parameter equals zero, the random effects model converges to its fixed effects counterpart (Elhorst, 2010) – this parameter is reported in Table 3. We find that $\phi = 0.0416$ and is significant at a one percent level, which corroborates the Hausman test and suggests that the fixed effects assumption is the appropriate specification given the data.

Table 4. Post Diagnostic Tests of Spatial Specification

Test	Chi-squared Statistic	P-value
LR		
Spatial lag	10.2658	0.2469
Spatial error	17.6215	0.0243
Wald Test		
Spatial lag	10.4900	0.2323
Spatial error	16.5544	0.0351
Hausman	190.0861	0.0000

Note: LR tests follow a chi-squared distribution with K degrees of freedom (Burridge, 1981).

The estimated coefficient on the spatial lagged term of energy emissions is positive, similar in magnitude, and highly statistically significant for both the SDM (column one in Table 3) and the SAR models (columns two and three of Table 3). Given our normalization of the spatial weighting matrix, the significant spatial autocorrelation coefficient suggests that energy emissions in neighboring states, on average, are exerting a positive effect on local energy emissions. Since energy emissions are estimated based on energy consumption across states, this coefficient implies that economic distance across states are affecting local energy emissions or vice versa. Economic distance could potentially represent energy commerce across states such as the exportation of electricity across state lines. This is consistent with Aldy (2005) and Carson (2010) who argued that interstate energy commerce affects the emissions-growth relationship. Aldy (2005) dealt with interstate electricity commerce by modifying the data for states that are net exporters of electricity. Our approach is somewhat different as we allow for spatial dependence to indirectly account for spillovers due to economic distance.

The estimated coefficients for the energy price terms are all between zero and one (in absolute terms) which indicates that energy emissions are price inelastic in the short run. This is consistent with expectations as in the short run consumers do not have the possibility to change their capital stock (which use these energy inputs) and can only change consumption behavior

(Bhattacharyya, 2011). Ergo, price elasticities are expected to be more inelastic in the short run than in the long run as consumers can change their capital stock in reaction to higher energy prices.[6] In accordance with the law of demand it is expected that all coefficients on the prices are negative – this expectation is violated in the case of coal prices across the different models; however, none of these estimates are statistically significant. According to the SAR FE model results, the estimated coefficients on electricity prices are highly significant across all models and imply that an increase in price by 10% would lead to an approximate decrease in energy emissions by 2.0%; this differs from the estimated elasticity of 2.6% reported in column four of Table 2. The estimated coefficients on oil prices are also highly significant across all the models. These coefficients imply that an increase in crude oil prices by 10% would lead to an approximate 2.4% decrease in energy emissions; again, this differs from the estimated elasticity of 3.4% found with the non-spatial model.

The estimated coefficients on the climate variables, cooling degree days and heating degree days, are fairly consistent across the three models. The estimated coefficient on cooling degree days is not significant for any of the models, but the positive sign is consistent with expectations as additional hotter days within a year are expected to increase the demand for cooling commercial spaces or private places of residence. The estimated coefficient on heating degree days is highly significant for the SAR FE model. The coefficient estimate from this model implies that an additional 10 % increase of cooler days within the year leads to an approximate 1.3% increase in energy emissions; this differs from the estimate 1.7% found with the non-spatial model (column four of Table 2). This is consistent with intuition as the additional heating of homes or office spaces requires more electricity, natural gas, or heating fuel – which in turn creates additional emissions. As with the energy prices, the estimated coefficients on the climate variables represent short-run values.

Finally, the estimated coefficients on the quadratic polynomial of GDP are all highly significant and consistent with the inverted U-shaped relationship espoused by the environmental Kuznets curve hypothesis. The signs and magnitudes on these terms are highly consistent across the three specifications. The estimated turning point for energy emissions is approximately $21,709 (USD) according to the results of the SAR FE model–this puts the turning point

[6] We do not explore long-run price elasticities in this particular study because such analyses require dynamic panel data models. The development of dynamic, spatial panel data models is still in its infancy so we abstract away from long-run elasticities.

somewhere below the median value of per-capita GDP in our sample.[7] The estimated turning point given the non-spatial model estimates is very similar at approximately $21,200 (USD). It is difficult estimate the year the turning point may have occurred because there is significant variation in per-capita GDP between states, but the U.S. national average of per-capita GDP (real dollars) reached $22K between 1988 and 1989.

5.2 Out-of-Sample Forecasting Ability

The analysis in the preceding section seemed to imply that the spatial autoregressive, panel data model provided the best fit to the data within-sample over the entire period of analysis. We now proceed by evaluating which model provides the best out-of-sample fit of the data by testing the forecasting ability of the different types of models. If the predictive ability of the spatial autoregressive model outperforms all the other models and corroborate the findings in the previous sub-section, then perhaps the findings here would lend some credibility to spatial panel data modeling approach.

An explanation for the one-step-ahead forecasts and five-year-ahead forecasts was presented above in the Methodological Approach Section. The results for the one-step-ahead forecasts are provided in Table 5, whereas the five-year-ahead forecasts are provided in Table 6.

Table 5. Root Mean Square Error Performance of One-Step-Ahead Forecasts

Models	2005	2006	2007	2008	2009
OLS	1.3133	1.5442	1.4620	1.6876	1.7081
FE	7.5787	9.0813	10.6852	11.5689	13.0098
SAR	7.3020	8.7498	10.2952	11.1466	12.5349
SDM	8.4778	10.1785	12.0032	13.0095	14.6796
SEM	7.0991	8.5066	10.0090	10.8368	12.1865

Note: All models with the exception of the pooled OLS control for both heterogeneous and time fixed effects.

[7] The turning point is calculated by $\exp(-\beta_1/(2\cdot\beta_2))$, where β_1 is the coefficient on the log of GDP and β_2 is the coefficient on the log of GDP, whole quantity squared.

Table 6. Root Mean Square Error Performance of One-Step-Ahead Forecasts

Models	2005	2006	2007	2008	2009	Average
OLS	1.3133	1.4334	1.4430	1.5079	1.5500	1.4496
FE	7.5787	0.2164	0.3120	0.3328	0.3591	1.7598
SAR	7.3020	0.2286	0.3460	0.3683	0.3890	1.7268
SDM	8.4778	0.2085	0.3188	0.3414	0.3796	1.9452
SEM	7.0991	0.2145	0.3146	0.3356	0.3598	1.6647

In the one-step-ahead forecasts, the pooled OLS model hands-down performs best of all the panel data models rather spatial or non-spatial. This can be observed by recalling that the smaller the root mean square error, the better the forecasting ability of the particular estimator. For short one-year-ahead forecasts, the forecasting error of the pooled OLS estimator is better than all the other estimators by orders of magnitude. This suggests that pooled OLS model performs better short-run forecasts of state-level CO2 emissions.

In contrast, the five-year-ahead forecasts paint a different picture. The pooled OLS estimator still provides the best one-year-ahead forecast and the lowest average forecast over the sample, but it does not perform as well as the other estimators in the last four years of the sample. The average performance over the entire out-of-sample period was skewed by the high over-prediction of the panel data models in the first year of observation. In the last four years of the forecasting horizon the spatial error model performs best on average, although the spatial Durbin model performs best in the second year.

6 Conclusions and Policy Implications

In response to criticisms that spatial, panel data models are fundamentally unidentifiable, this study sought to focus on a different statistical validation strategy – i.e., out-of-sample forecasting. In other words, we treat the spatial, panel data models as a black box and test these models forecasting ability against empirical reality.

In the first part of our empirical exploration we sought to evaluate the best statistical estimator based upon its fit to the entire sample of observations. This evaluation was developed based upon sound post-diagnostic tests to determine if the spatial model were more appropriate over the non-spatial models. Our evidence in the first step of empirical evaluation suggested that the spatial autoregressive, panel data model seemed to provide the best to the data within sample over the entire period of observation.

We carried our empirical exploration further to compare the forecasting ability of the different estimators. The pooled OLS estimator provided the best fit to the data in a short run, root mean square error context. This may suggest that simple OLS models may perform best in determining short-run, state-level CO2 emission levels. The long-run forecasts, however, suggested that the panel data estimation methods performed better than the OLS estimator. And more specifically, the spatial error model performed best for years two through five.

Whether spatial panel data model are unidentifiable is yet to be determined. However, in criticizing such methods it is important to consider the validation strategy of out-of-sample prediction performance of such models. The evaluation of the forecasting ability of spatial panel data methods is still in its infancy. This paper sought to expand our understanding of spatial panel data by testing the forecasting performance of such models. The results of this study suggest that spatial panel data models offer value in medium- to long-run forecasts, especially in determining state-level CO2 emissions. Such analyses are important for helping to determine the spatial and temporal distribution of CO2 emissions in the U.S.

The findings within this study are important for two reasons. From a policy standpoint, it is important to better understand the driving forces of carbon dioxide emissions. Understanding these forces will help better equip policy makers to design effective climate change mitigation policies. From a statistical standpoint, it is important to continue to test spatial econometric models to see how they perform against non-spatial models. With advances in spatial panel data models, this methodology can now be tested in terms of the model's forecasting ability.

Future studies should consider the forecasting ability of the spatial panel data models by also incorporating temporal dependence into the model. This of course raises the possibility of introducing dynamic panel data bias (Nickell, 1981), but with increasing panel datasets that contain sufficient observations through time this should become less of an issue in the future. Past studies that have introduced spatio-temporal panel data (e.g., Giacomini and Granger, 2004), have generally focused less on the proper spatial model specification than the standard spatial econometrics literature. As the spatial panel data literature continues to expand there will probably be significant convergence between these two literatures in the near future. Nevertheless, the spatial econometric models will need to be further tested against empirical reality in the future to help prove their validity.

References

Aldy, J. (2005). "Environmental Kuznets Curve Analysis of U.S. State-Level Carbon Dioxide Emissions." *J. of Environment and Development*, 14(2), 48–72.

Anselin, L., Gallo, J. Le, & Jayet, H. (2008). "Spatial Panel Econometrics." Pages 625–660 of: Matyas, L., & Sevestre, P. (eds), *The Econometrics of Panel Data*. Berlin: Springer-Verlag.

Aroonruengsawat, A., Auffhamer, M., & Sandstad, A.H. (2012). "The Impact of State Level Building Codes on Residential Electricity Consumption." *The Energy Journal*, 33, 31–52.

Auffhammer, M., & Carson, R.T. (2008). "Forecasting the Path of China's CO2 Emissions Using Province-Level Information." *J. of Environmental Economics and Management*, 55, 229–247.

Auffhammer, M., & Steinhauser, R. (2007). "The Future Trajectory of US CO2 Emissions: The Role of State vs. Aggregate Information." *J. of Regional Science*, 47(1), 47–61.

Baltagi, B.H. (2005). *Econometric analysis of panel data*. 3 edn. Chichester, U.K.: Wiley.

Baltagi, B.H., & Griffin, J.M. (1997). "Pooled estimators vs. their heterogeneous counterparts in the context of dynamic demand for gasoline." *Journal of Econometrics*, 77, 303–327.

Baltagi, B.H., Bresson, G., & Pirotte, A. (2002). "Comparison of forecast performance for homogenous, heterogeneous and shrinkage estimators: Some evidence from US electricity and natural gas consumption." *Economic Letters*, 76, 375–382.

Berndt, E.R. (1996). *The Practice of Econometrics: Classic and Contemporary*. New York: Addison Wesley.

Bhattacharyya, S.C. (2011). *Energy Economics: Concepts, Issues, Markets, and Governance*. London, UK: Springer-Verlag.

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." Tech. rept. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, TN.

Burridge, P. (1980). "On the Cliff-Ord test spatial autocorrelation." *Journal of the Royal Statistical Society B*, 42, 107–108.

Burridge, P. (1981). "Testing for a common factor in a spatially autoregression model." *Environment and Planning A*, 13, 795–800.

Carson, R.T. (2010). “The Environmental Kuznets Curve: Seeking Empirical Regularity and Theoretical Structure.” *Rev Environmental Economics and Policy*, 4(1), 3–23.

Conley, T.G., & Ligon, E. (2002). “Economic distance and cross-country spillovers”. *Journal of Economic Growth*, 7, 157–187.

Debarsy, N., & Ertur, C. (2010). “Testing for spatial autocorrelation in a fixed effects panel date model.” *Regional Science and Urban Economics*, 40, 453–470.

Diaz, H., & Quayle, R. (1980). “Heating Degree Day Data Applied to Residential Heating Energy Consumption.” *J. of Applied Meterology*, 3, 241–246.

Elhorst, J. (2009). “Spatial Panel Data Models.” Pages 377–407 of: Fischer, M.M., & Getis, A. (eds), *Handbook of Applied Spatial Analysis*. Berlin: Springer-Verlag.

Elhorst, J. (2010). “Matlab software for spatial panels.” Tech. rept. University of Groningen, Groningen, Netherlands. http://www.regroningen.nl/elhorst/software.shtml, Accessed February 2012.

Hanemann, W.M. (1984). “Discrete/continuous models of consumer demand.” *Econometrica*, 541–561.

Hausman, J.A. (1978). “Specification Tests in Econometrics.” *Econometrica*, 46(6), 1251–1271.

Itkonen, J.V.A. (2012). “Problems estimating the carbon Kuznets curve.” *Energy*, 39, 274–280.

Kim, C.W., Phipps, T.T., & Anselin, L. (2003). “Measuring the benefits of air quality improvement: a spatial hedonic approach.” *Journal of Environmental Economics and Management*, 45, 24–39.

Kindle, A.G., & Shawhan, D.L. (2011). “An Empirical Test for Inter-State Carbon-Dioxide Emissions Leakage Resulting from the Regional Greenhouse Gas Initiative.” Tech. rept. Rensselaer Polytechnic Institute. Accessed December 2011.

Lee, L.F., & Yu, J. (2010a). “Estimation of spatial autoregressive panel data models with fixed effects.” *Journal of Econometrics*, 154, 165–185.

Lee, L.F., & Yu, J. (2010b). “Some recent developments in spatial panel data models.” *Regional Science and Urban Economics*, 40, 255–271.

LeSage, J.P., & Pace, R.K. (2009). *Introduction to Spatial Econometrics*. New York: CRC Press.

Maddala, G.S., Trost, R.P., Li, H., , & Joutz, F. (1997). “Estimation of short-run and long-run elasticities of energy demand from panel data using shrinkage estimators.” *Journal of Business and Economic Statistics*, 15, 90–100.

Ryan, D.L., & Ploure, A. (2009). "Empirical modeling of energy demand." Pages 112–143 of: Evans, J., & Hunt, L.C. (eds), *International handbook on the economics of energy*. Cheltenham, U.K.: Edward Elgar.

Swanson, B. (2010, September). "Answers Archive: Heating and Cooling Degree Days." Accessed January 2010.

US Bureau of Economic Analysis. (2012). "Gross Domestic Product by State." http://www.bea.gov, Accessed January 2012.

US Census Bureau. (2010). "Population Estimates." http://www.census.gov/popest/states/state.html, Accessed November 2011.

US Energy Information Administration. (2008). "State Energy Data Systems." http://www.eia.gov, Accessed December 2011.

US Energy Information Administration. (2012). "State-Level Energy-Related Carbon Dioxide Emissions, 2000-2009." http://www.eia.gov, Accessed January 2012.

US Nat'l Climate Data Center. (2010). "Heating and Cooling Degree Day Data." http://www.ncdc.noaa.gov, Accessed November 2011.