(2014) 44, 223–240

The Review of Regional Studies

The Official Journal of the Southern Regional Science Association

Forecasting U.S. State-Level Carbon Dioxide Emissions

J. Wesley Burnett[a] and Xueting Zhao[b]

[a]*Economics Department, College of Charleston, USA*
[b]*Regional Research Institute, West Virginia University, USA*

Abstract: This study explores the use of spatial models in forecasting U.S. state-level carbon dioxide emissions. We compare forecasts against empirical reality using panel data models with and without spatial effects. Understanding how to predict emissions is important for designing climate change mitigation policies. To determine if spatial econometric models can help us predict emissions, it is important to test these models to see if they are a valid strategy to describe the underlying data, in the context of forecasting. We find that a non-spatial OLS estimator performs best in all out-of-sample forecasts; however, the OLS model is not statistically distinguishable from a spatial panel data model with random effects.

Keywords: spatial panel data econometrics, forecasting, carbon dioxide emissions
JEL Codes: C33, C53, Q50

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. The geographic distribution of CO_2 emissions does not affect the global climatic impact, but 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 multilateral agreement to mitigate emissions such as the Kyoto Protocol, it must begin to look inward to determine how to reduce its major sources of emissions.

Most sources of CO_2 emissions come from energy-related activities, principally through coal-fired electricity generation and transportation (U.S. Department of Energy, 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. Therefore, understanding the subnational-level sources of emissions and spatial interactions of these sources across regions will be important for formulating policies to mitigate GHG emissions. That is, global climate change is an international problem in scope, yet domestic or regional policies can be implemented to mitigate CO_2 emissions. Perhaps one of the reasons that the U.S. has been slow to adopt a national mitigation scheme is due to uncertainty in state-level abatement costs.

Burnett is Assistant Professor, Economics Department, College of Charleston, Charleston, SC 28424. Zhao is Postdoctoral Research Fellow, Regional Research Institute, West Virginia University, Morgantown, WV 26506. *Corresponding Author*: J. W. Burnett E-mail: burnettjw@cofc.edu

ISSN 1553-0892, 0048-749X (online)
www.srsa.org/rrs

Current policy regimes often ignore location and dispersion characteristics of the sources of emissions (Fowlie and Muller, 2013).

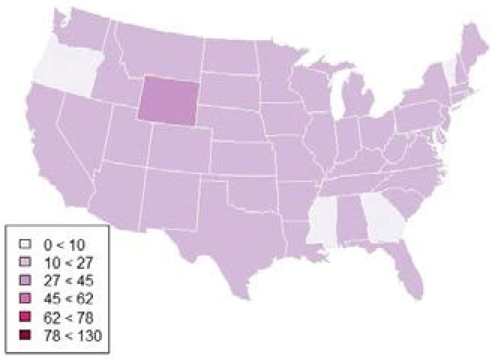
Spatial panel data models are a promising means to examine the spatial and temporal distribution of CO_2 emissions. Spatial econometrics is an applied field of econometrics that deals with sample data that is collected with reference to location measured as points in space. What distinguishes spatial econometrics from traditional econometrics is that the locational data may be characterized by spatial dependence or spatial heterogeneity (LeSage and Pace, 2009). As in time series, if this autocorrelation is present and unaccounted for then it could lead to biased or 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 dynamic, spatial panel data models that control for both spatial and temporal dependence within the underlying data (Yu, de Jong, and Lee, 2012).

However, we note that some spatial econometric specifications have come under criticism recently for problems associated with identification and for a lack of appeal to theoretical foundations. According to these criticisms, the problem of identification is similar to Manski's (1993) "reflection problem," where group average characteristics (neighboring state carbon dioxide emissions and structural characteristics) affect individual outcomes (local carbon dioxide emissions) but the parameters in the model are not identifiable (Partridge et al., 2012). That is, it is hard to separate out the effects of what causes emissions locally versus what causes emissions in neighboring states.

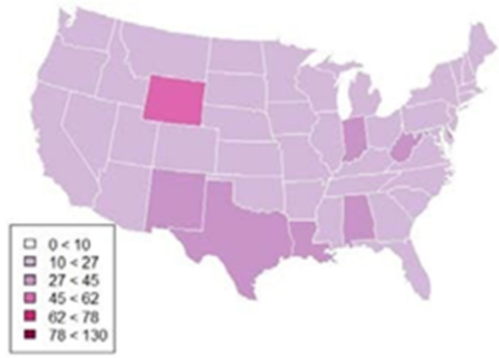
Nevertheless, we agree that there may be potential problems with the exogeneity of the spatial weighting matrix (among other things), so rather than appealing to causality we instead appeal to an alternative validation strategy that is less dependent on prior theory. That is, we take these models as a black box and test them against empirical reality (Freedman, 1991). Against this background, we compare forecasts of state-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 prediction against non-spatial panel data models in a forecasting error context. We compare the performance of several predictors for state-level CO_2 emissions for one-step-ahead iterated forecasts and for *n*-year-ahead forecasts. (Our data consist of yearly observations, so we will use the terms "one-step-ahead" and "one-year-ahead" forecasts interchangeably, unless explicitly stated otherwise). Based on forecast performance, we find that a simple, non-spatial ordinary least squares (OLS) estimator performs best for all forecast error metrics; however, the OLS estimator is not statistically different from the predictions of a spatial panel data model with random effects. These findings may suggest that for forecasting U.S. state-level carbon dioxide emissions, models that take into account spatial autocorrelation (and heterogeneity) provide the best "within sample" fit to the data, but not the best fit for "out of sample" predictive ability. To check the robustness of our findings, it may also be of interest to explore in future research the performance of different predictors along the lines of Kelejian and Prucha (2007).

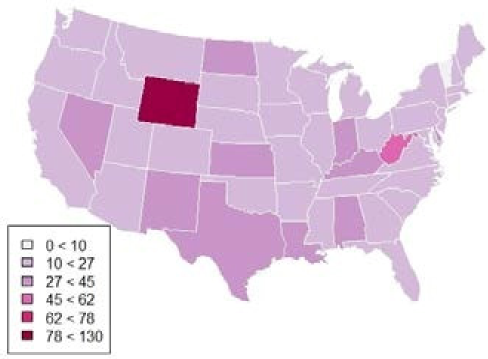
Figure 1: Per capita CO_2 Emissions, 1960-1980



(a) 1960



(b) 1970



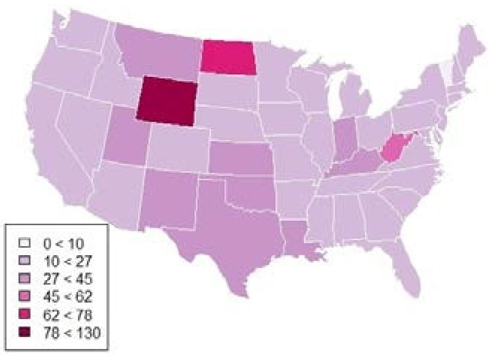
(c) 1980

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. From a policy perspective, an analysis of per capita emissions offers a more equitable measure for negotiating multilateral agreements. Based on these insights, we express a reduced form model similar to that of Burnett, Bergstrom, and Dorfman (2013). The structural and nonstructural factors we examine are GDP per capita, energy prices, and climatological factors.

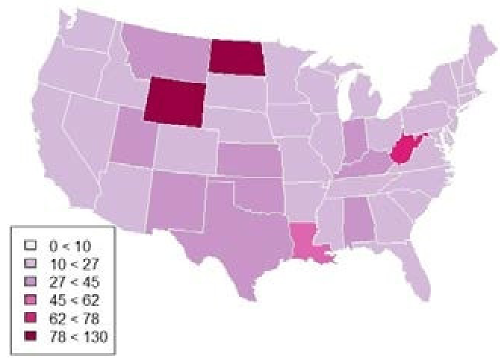
Chloropleth maps of state-level (aggregate) emissions for 1960-2009 (by decade) are offered in Figures 1 and 2. Each map's scale is based on the bins of per capita emissions in the year 2009. These maps show a general increase in per capita, state-level emissions, but a gradual easing of intensities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

Until fairly recently, not many papers have examined the forecasting ability of spatial econometric models—exceptions include Baltagi and Li (2006), Baltagi, Broniak, and Marland (2014), Elhorst (2014), and Kelejian and Prucha (2007), among others. However, there does seem to be a growing interest in the alternative validation strategy of prediction.

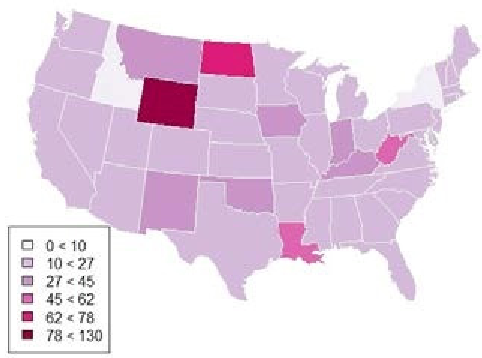
Figure 2: Per capita CO_2 Emissions, 1990-2009



(a) 1990



(b) 2000



(c) 2009

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).

2. METHODOLOGICAL APPROACH

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

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

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; and $\otimes$ (here and throughout) denotes the Kronecker product. All the terms are represented in natural logs so that the estimated coefficients can be interpreted as elasticities. The coefficient μ denotes the unobserved 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. $\boldsymbol{\iota}_N$ denotes a column vector of ones with the subscript denoting the dimension. $\mathbf{I}_N$ is an identity matrix of square subscripted dimension. We treat the unobserved individual effect as if randomly drawn from the population (Wooldridge, 2002). If the unobserved individual effects term is correlated with the explanatory variables, **X**, and we estimate (1) without controlling for it, then the estimates will be biased.

The unobserved individual effects, μ and η , can be eliminated by transforming the data as follows (where for ease of exposition we write in scalar notation)

(2) $$\bar{z}_{i.} = \frac{1}{T}\sum_{t=1}^{T} z_{it}, \quad \bar{z}_{.t} = \frac{1}{N}\sum_{i=1}^{N} z_{it}, \quad \bar{z}_{..} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} z_{it},$$

and

(3) $$\tilde{z}_{it} = z_{it} - \bar{z}_{i.} - \bar{z}_{.t} + \bar{z}_{..},$$

where z_{it} denotes any of the variables (dependent or independent) within the study. Applying this transformation to (1) yields

(4) $$\tilde{y}_{it} = \tilde{x}_{it}\beta + \tilde{u}_{it}.$$

The OLS estimator of (4) is typically referred to as the within estimator of the two-way model (1), or as the fixed effects estimator, or as the least squares dummy variable (LSDV) estimator (Baltagi, 2005). The fixed effects estimator is used 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. Throughout the rest of thearticle we will use the term “fixed effect” to denote that the data has been transformed to eliminate the heterogeneous fixed effect term alone (i.e., without eliminating the time effect). We will use the term “both effects” to denote that the data has been transformed to eliminate both the heterogeneous fixed effect and time effect.

We introduce spatial effects into the model by using a standard (pre-specified and non-negative) spatial weighting matrix, $\mathbf{W}_N$, as an *(N x 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

(5) $$\boldsymbol{W}_{NT} = \boldsymbol{I}_T \otimes \boldsymbol{W}_N.$$

For our estimations, we specify a simple first-order contiguity matrix where a neighboring state is assigned a value of one and zero otherwise if two states are not neighbors. The weighting matrix is then row-standardized so that the marginal effect estimates can be interpreted as an average effecte from neighbors. Burnett, Bergstrom, and Dorfman (2013) extended the analysis to consider a nearest-neighbor specification but did not find considerable differences in the estimated spatial autocorrelation coefficients. Further, LeSage and Pace (2010) argue that the marginal effect estimates and inference are *not* sensitive to the particular choice of the spatial weighting matrix.

2.1 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 autoregressive (SAR), spatial error (SEM), and spatial Durbin (SDM) models. Kelejian and Prucha (2007) report on the performance of different predictors for a cross sectional SARAR(1,1) model, and find that a reduced form predictor performs poorly; nevertheless, we use such predictors for our evaluation purposes. An expression for the first three models (the pooled OLS, the spatial autoregressive model, and the spatial error model) is provided in the Appendix. All of the spatial models are estimated via maximum likelihood following algorithms in LeSage and Pace (2009). The models were estimated in Matlab using code provided by Elhorst (2014).

Following Yu, de Jong, and Lee (2012), we extend the spatial models in Burnett, Bergstrom, and Dorfman (2013) to consider a dynamic spatial panel data model. Using the notation from Yu, de Jong, and Lee (2012) we consider the following model

$$(6)\ \boldsymbol{y} = \lambda(\boldsymbol{I_T} \otimes \boldsymbol{W}_N)\boldsymbol{y} + \gamma \boldsymbol{y_{t\text{-}1}} + \rho(\boldsymbol{I_T} \otimes \boldsymbol{W}_N)\boldsymbol{y_{t\text{-}1}} + \boldsymbol{X\beta} + (\boldsymbol{\iota_T} \otimes \boldsymbol{I}_N)\mu + (\boldsymbol{I_T} \otimes \boldsymbol{\iota}_N)\eta + \boldsymbol{\varepsilon},$$

where the subscript, *t-1*, indicates a temporally lagged variable. The model from Yu, de Jong, and Lee (2012) also allows for spatial lags of the explanatory variable, but we do not consider such an application in the present study. They provide a quasi-maximum likelihood method of estimating Equation (6) with both the state-level heterogeneous fixed effect, μ, and the time fixed effect, η. The specification in Equation (6) is slightly different from the previous spatial panel data models as it contains a temporally lagged dependent variable premultiplied by the scalar coefficient γ. The model also contains a temporally lagged, spatially weighted dependent variable premultiplied by the scalar coefficient ρ. Unlike the SDM model, Yu, de Jong, and Lee's specification in (6) does not include spatially weighted explanatory variables.

2.2 Forecasting

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. An explanation of the prediction formulas is provided in the Appendix.

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

$$(7)\qquad MAE = \sum_{t=1}^{T}\sum_{i=1}^{N}\frac{1}{N\cdot T}\left|F(t)-A(t)\right|,$$

$$(8)\qquad MAPE = \sum_{t=1}^{T}\sum_{i=1}^{N}\frac{1}{N\cdot T}\left|\frac{F(t)-A(t)}{A(t)}\right|,$$

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

The symbol T denotes the total number of periods and N denotes the total number of spatial units 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. According to Kennedy (2008), MAE is appropriate when the cost of forecast errors is proportional to the absolute size of the forecast error. MAPE is the average of the absolute values of the percentage forecast errors, and it has the advantage of being dimensionless. MAPE is more appropriate when the cost to forecast error is more closely related to the percentage error than to the numerical size of the error (Kennedy, 2008). A problem with MAPE is that it often performs under-forecasting. The errors in the RMSE metric are squared before averaging, so the RMSE gives a relatively higher weight to large errors – therefore, RMSE represents a quadratic loss function. RMSE is one of the most popular metrics in use. For the sake of robustness, we consider all of these forecasting metrics in the current study.

There are two principal types of forecasts. The first type of forecast compares predicted values to actual observations one year at a time. These one-step-ahead, iterated 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 forecast compares the predicted values to actual observations over the entire out-of-sample period. The n-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 (n 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 n-year-ahead forecasts, on the other hand, provide a metric for evaluating the medium-run predictive ability of the model. To compare each of the models, we will evaluate both types of forecasts.

2.3 Diebold-Mariano Statistic

Finally, we use a panel Diebold and Mariano (1995) (henceforth, DM) statistic to determine statistically if the model forecasts can be distinguished from the other forecasts. The panel version of the DM statistic was developed by Pesaran, Schuermann, and Smith (2009).. The illustration here closely follows the authors' original notation. Consider the following forecasting errors for method (or model) A relative to method B:

$$z_{it} = \left[e_{it}^{A}(1) \right]^2 - \left[e_{it}^{B}(1) \right]^2,$$
$$A \equiv \text{ Proposed forecast}$$
$$B \equiv \text{ Benchmark forecast,}$$

for i = 1, 2, …, N; t = 1, 2, …, T; and, $e_{it}(1)$ is the one-year-ahead forecast error. The term N denotes the number of states within our panel, and T denotes the forecast sample period. The panel DM statistic is developed as follows: for a given variable (state-level carbon dioxide emissions per capita), consider

$$z_{it} = \alpha_i + \varepsilon_{it},$$
$$H_0 : \alpha_i = 0,$$
$$H_1 : \alpha_i < 0, \text{ for some } i.$$

Under the null, we assume that the error term is independently and identically distributed, $\varepsilon_{it} \sim iid(0,\sigma_i^2)$, such that

$$\overline{DM} = \frac{\overline{z}}{\sqrt{V(\overline{z})}} \sim N(0,1),$$

where

$$\overline{z} = N^{-1}\sum_{i=1}^{N}\overline{z}_i, \quad \overline{z}_i = T^{-1}\sum_{t=1}^{T} z_{it},$$

and

$$V(\overline{z}) = (N \cdot T)^{-1} \cdot \left(N^{-1}\sum_{i=1}^{N}\hat{\sigma}_i^2 \right),$$

$$\hat{\sigma}_i^2 = \frac{\sum_{t=1}^{T}(z_{it} - \overline{z}_i)^2}{T-1}.$$

We restrict our panel DM tests to one-step-ahead forecasts to reduce the chance of having serial correlation within the underlying data, which could possibly invalidate our ability to infer from the DM tests. We made no adjustments for serial correlation because we are only forecasting one year in advance; therefore, it is somewhat reasonable to assume that the differentials are serially uncorrelated. For forecasts greater than one year, the panel DM statistic can be modified to deal with the serial correlation by using a Newey-West type estimator for $V(\overline{z}_i)$. We do not pursue this extension here. According to Pesaran, Schuermann, and Smith (2009), the degree of cross sectional dependence of the forecast errors has to be sufficiently weak for $N^{-1}\sum_{i=1}^{N}\overline{z}_i$ to tend to a normal distribution as $N \to \infty$. Unfortunately, due to the limited number of states within our analysis, we are not able to pursue such an extension in the current paper. The panel DM is a one-sided test, so the relevant one percent and five percent critical values are -2.326 and -1.634, respectively. A positive value of the panel DM statistic will present evidence against the proposed forecasting model defined as A above. As a sensitivity analysis, we use the panel DM tests to determine if the forecasting models are statistically distinguishable from the one another. The results of the panel DM statistic are listed in Section 4.2.

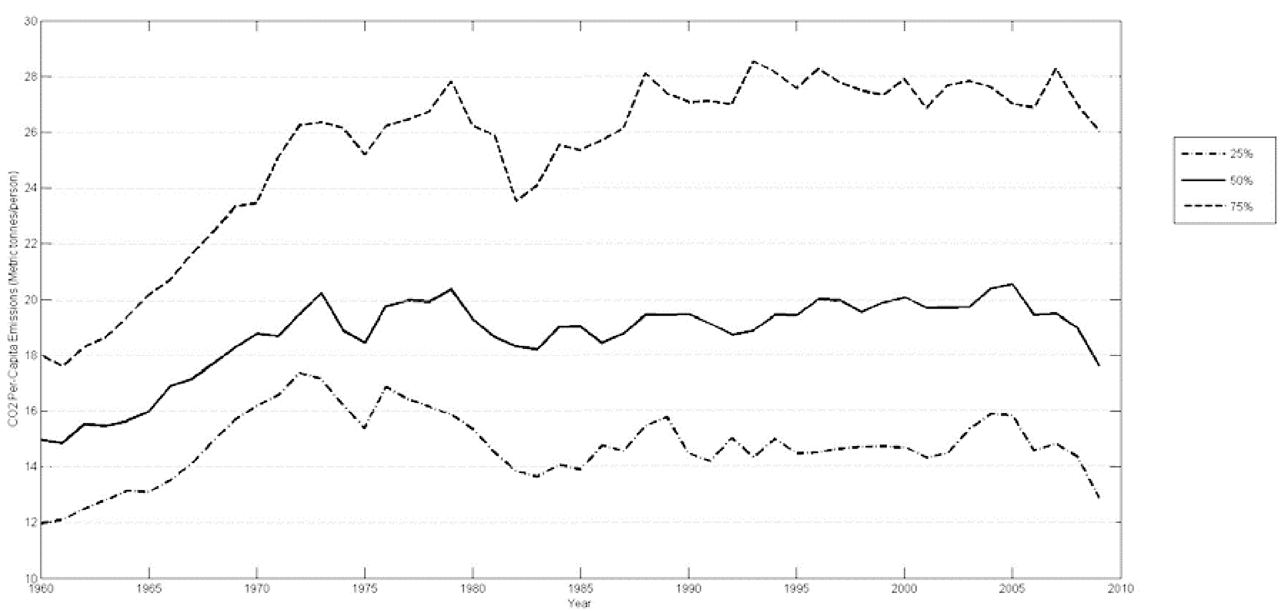
3. DATA

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 (CO_2), 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, CO_2 emissions are estimated based upon a state's observed energy use. Therefore, CO_2 emissions are not to be confused with actual CO_2 pollution emissions that are emitted from the end of a smokestack or tailpipe.[1]

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 Broniak, and Marland, 2004; US Department of Energy, 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 CO_2 emissions and not actual atmospheric emissions. Despite this deficiency, this measure of emissions is one of the more commonly used measures 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 Energy Information Administration (EIA) within the U.S. Department of Energy (2012). 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 tons for the 48 contiguous states, excluding the District of Columbia.

A plot of the quartiles of aggregate U.S. per capita CO_2 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 CO_2 Emissions, 1960-2009



[1] The CO_2 data should also not be confused with atmospheric CO_2 pollution, which following emission enters the upper atmosphere and is more global in scope.

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 CO_2 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, CO_2 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

$$(10) \qquad CO_{2,t} \equiv \alpha_t^{oil} \cdot E_t^{oil} + \alpha_t^{coal} \cdot E_t^{coal} + \alpha_t^{gas} \cdot E_t^{gas} + \alpha_t^{flare} \cdot E_t^{flare} + S_t,$$

where α_t^{oil}, α_t^{coal}, α_t^{gas}, $\alpha_t^{flare} > 0$ are the related thermal conversion factors. There are time subscripts on each of these coefficients because the thermal conversion rates change through time, such as technological improvements that reduce the amount of carbon emissions.

The GDP data were obtained from the Bureau of Economic Analysis (BEA) within the U.S. Department of Commerce (2010). The BEA offers annual state-level GDP estimates from 1963 to the near present. The estimates are based on the per capita nominal GDP by state. The estimates were converted to real dollars by using the BEA's GDP implicit price deflator.

To model climatic influences on energy demand, we use cooling-degree days (CDD) and heating-degree days (HDD), which were obtained from the U.S. National Climate Data Center within the National Oceanic and Atmospheric Administration (NOAA, 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 (NOAA, 2010). 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.

Energy prices were obtained from the EIA (US Energy Information Administration, 2012). 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 US Census Bureau (2010). 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: CO_2 and GDP represent per capita values. CO_2 is measured in metric tons. GDP and the price data are measured in real USD. Prices are measured as USD per unit of BTU.

4. EMPIRICAL ESTIMATION AND RESULTS

Burnett, Bergstrom, and Dorfman (2013) found that the SAR model provided the best fit to the entire sample of the data. Since spatial models have recently been criticized for being fundamentally unidentifiable, we extend their original work by comparing the forecasting ability of the different panel data models. That is, we use the same models to compare one-step-ahead and *n*-year-ahead forecasts to determine which model provides the best fit to the data out-of-sample.

4.1 Entire Sample Testing and Diagnostics

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

$$\ln(y_{it}) = \beta_0 + \beta_1 \ln(p_{it}^{c}) + \beta_2 \ln(p_{it}^{ng}) + \beta_3 \ln(p_{it}^{o}) + \beta_4 \ln(p_{it}^{e}) + \beta_5 \ln(GDP_{it}) + \beta_6 \ln(GDP_{it})^2 + \beta_7 \ln(CDD_{it}) + \beta_8 \ln(HDD_{it}) + \mu_i + \eta_t + u_{it}. \tag{11}$$

Equation (11) is a reduced-form model for energy demand (Ryan and 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, this model is very similar in nature to Aroonruengsawat, Auffhamer, and Sanstad's model (2012). The term y_{it} denotes real per capita CO_2 emissions in state i and 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_t^{oil}, p_t^{ng}, p_t^{c}, and p_t^{e} denote real state-level prices of crude oil, natural gas, coal, and electricity, respectively.

A full explanation of the within-sample regression results and diagnostic tests for equation (11), including the assumptions within the model, is available in Burnet, Bergstrom, and Dorfman (2013). To briefly recap the results, the diagnostic tests and regression analyses implied that the SAR model provided the best fit to the entire sample of the data.

4.2 Out-of-Sample Forecasting Ability

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 corroborates the findings in the previous sub-section, then perhaps the findings here will lend some credibility to the spatial panel data modeling approach.

Table 2: Forecast Error Performance of Iterated, One-Step-Ahead Forecasts

Models	2005	2006	2007	2008	2009	Average
Pooled OLS						
MAE	**0.2776**	**0.2826**	**0.2693**	**0.2914**	**0.2931**	**0.2828**
MAPE	**0.0842**	**0.0861**	**0.0814**	**0.0900**	**0.0932**	**0.0870**
RMSE	**0.4437**	**0.4424**	**0.4313**	**0.4719**	**0.4771**	**0.4533**
Non-Spatial (One-Way Effect)						
MAE	8.6761	7.5880	7.0660	7.9743	6.4720	7.5553
MAPE	2.9066	2.5784	2.3884	2.7334	2.2770	2.5767
RMSE	8.6853	7.5979	7.0758	7.9857	6.4839	7.5657
SAR (Two-Way Effects)						
MAE	35.1842	34.7006	34.6125	34.2855	33.6858	34.4937
MAPE	11.7440	11.7429	11.6520	11.6944	11.7900	11.7247
RMSE	35.1870	34.7036	34.6155	34.2885	33.6888	34.4967
SDM (Two-Way Effects)						
MAE	28.8656	27.7632	28.4562	28.6360	25.7347	27.8912
MAPE	9.6393	9.4005	9.5842	9.7718	9.0131	9.4818
RMSE	28.8693	27.7672	28.4600	28.6398	25.7389	27.8951
SAR RE (Two-Way Effects)						
MAE	0.3179	0.3233	0.3182	0.3219	0.3363	0.3235
MAPE	0.1005	0.1034	0.1004	0.1036	0.1131	0.1042
RMSE	0.4523	0.4590	0.4569	0.4563	0.4509	0.4551
SEM (Two-Way Effects)						
MAE	35.6073	34.7963	34.6187	34.3644	34.0266	34.6827
MAPE	11.8846	11.7748	11.6536	11.7207	11.9086	11.7884
RMSE	35.6101	34.7993	34.6216	34.3674	34.0296	34.6856
SDPD (One-Way Effect)						
MAE	5.9260	4.9630	5.5135	4.8652	4.6697	5.1875
MAPE	1.9767	1.6791	1.8544	1.6585	1.6327	1.7603
RMSE	5.9263	4.9636	5.5138	4.8656	4.6702	5.1879
SDPD (Two-Way Effects)						
MAE	5.6254	5.0609	5.9335	5.4468	5.1215	5.4376
MAPE	1.8757	1.7114	1.9946	1.8558	1.7896	1.8454
RMSE	5.6256	5.0613	5.9337	5.4470	5.1218	5.4379

Note: The term "One-way Effect" indicates that a heterogeneous, state-level fixed effect has been controlled for in the model. The term "Two-way Effect" indicates that both a heterogeneous fixed effect and a time fixed effect have been controlled for in the model. Numbers highlighted above indicate the smallest forecast errors among the group of estimators.

An explanation for the iterated, one-step-ahead forecasts and *n*-year-ahead forecasts was presented above in the Methodological Approach sections. The results for the iterated, one-step-ahead forecasts are provided in Table 2, whereas the *n*-year-ahead forecasts are provided in Table 3. We follow the basic prescription of Goldberger (1962), who claimed that a forecast should incorporate all available information. Therefore, where relevant we include both heterogeneous, state-level fixed effects and time fixed effects.

In the iterated, one-step-ahead forecasts in Table 2, the pooled OLS model hands-down performs best of all the panel data models. This can be observed by recalling that the smaller the forecast 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 most of the other estimators by orders of magnitude. The only exception in performance is offered by the spatial panel data model with random effects (SAR RE). Pooled OLS still outperforms SAR RE but only by a tiny margin. This suggests that the pooled OLS model performs better short-run forecasts of state-level CO_2 emissions.

The *n*-year-ahead forecasts in Table 3 offer a similar story to the one-step-ahead forecasts. Again, the pooled OLS estimator still provides the best forecasts across all of the different estimators. The SAR RE model is the only estimator that comes close in forecasting ability.

It is worth noting that the average forecast errors are smaller for the *n*-year-ahead forecasts over the one-year-ahead forecasts for the pooled OLS and SAR RE models. These results seem counter-intuitive. Casual observation of all the forecasts (not provided) suggests that the one-step-ahead forecasts are over-predicting the state-level emissions. This is arguably due to the fact that the rate of growth of emissions is declining in our out-of-sample observation period. Therefore, the one-step-ahead forecasts will over-predict emissions in the subsequent year. The *n*-year-ahead forecasts, on the other hand, seem to do a better job of fitting this trending behavior over the out-of-sample observation period. Thus, the average *n*-year-ahead forecast errors are slightly smaller than the one-year-ahead forecast errors.

The results of the panel DM tests largely corroborate the findings in Tables 2 and 3. That is, the DM test results imply that the pooled OLS model provides forecasts that are statistically distinguished from the benchmark forecasts at conventional significance levels. The only exception is for the spatial autoregressive model with random effects, which is statistically indistinguishable from the pooled OLS forecasts.

5. 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.

As evidenced by the empirical evaluations of Burnett, Bergstrom, and Dorfman (2013), the spatial autoregressive, panel data model provided the best fit to the within-sample data over the entire period of observation. We carried the empirical exploration one step further to compare the forecasting ability of the different estimators. Unlike the within-sample results, the pooled OLS estimator provided the best fit to the data in an out-of-sample forecasting context. This suggests, at least in this particular case, that spatial models perform best within-sample, but on average the pooled OLS estimators perform best out of sample. However, the pooled OLS estimator is not statistically distinguishable from the spatial autoregressive model with random effects, so such models are just as effective as predicting state-level carbon dioxide emissions.

Table 3: Forecast Error Performance of *n*-Year-Ahead Forecasts

Models	2005	2006	2007	2008	2009	Average
Pooled OLS						
MAE	**0.2776**	**0.2801**	**0.2766**	**0.2800**	**0.2814**	**0.2792**
MAPE	**0.0842**	**0.0852**	**0.0841**	**0.0856**	**0.0867**	**0.0851**
RMSE	**0.4437**	**0.4435**	**0.4388**	**0.4464**	**0.4508**	**0.4446**
Non-Spatial (One-way Effect)						
MAE	8.6761	8.5851	8.5420	8.4470	8.5123	8.5525
MAPE	2.9066	2.8956	2.8823	2.8611	2.9052	2.8902
RMSE	8.6853	8.5946	8.5513	8.4585	8.5241	8.5628
SAR (Two-way Effects)						
MAE	35.1842	35.1668	35.1556	35.1539	35.1776	35.1676
MAPE	11.7440	11.8223	11.8248	11.8657	11.9625	11.8439
RMSE	35.1870	35.1702	35.1590	35.1570	35.1810	35.1708
SDM (Two-way Effects)						
MAE	33.2276	33.2133	33.2064	33.2075	33.2323	33.2174
MAPE	11.0923	11.1668	11.1705	11.2100	11.3023	11.1884
RMSE	33.2306	33.2168	33.2099	33.2108	33.2359	33.2208
SAR RE (Two-way Effects)						
MAE	0.3179	0.3754	0.3720	0.3607	0.3702	0.3592
MAPE	0.1005	0.1217	0.1200	0.1176	0.1213	0.1162
RMSE	0.4523	0.4996	0.5025	0.4795	0.4993	0.4867
SEM (Two-way Effects)						
MAE	35.6073	35.5847	35.5706	35.5651	35.5901	35.5836
MAPE	11.8846	11.9622	11.9638	12.0038	12.1022	11.9833
RMSE	35.6101	35.5880	35.5739	35.5681	35.5934	35.5867
SDPD (One-way Effect)						
MAE	5.9260	5.9297	5.9226	5.9271	5.9368	5.9285
MAPE	1.9767	1.9921	1.9905	1.9991	2.0172	1.9952
RMSE	5.9263	5.9302	5.9230	5.9275	5.9373	5.9289
SDPD (Two-way Effects)						
MAE	5.6254	5.6297	5.6210	5.6247	5.6380	5.6277
MAPE	1.8757	1.8908	1.8885	1.8964	1.9150	1.8933
RMSE	5.6256	5.6301	5.6213	5.6250	5.6384	5.6281

Note: The term “One-way Effect” indicates that a heterogeneous, state-level fixed effect has been controlled for in the model. The term “Two-way Effect” indicates that both a heterogeneous fixed effect and a time fixed effect have been controlled for in the model. Numbers highlighted above indicate the smallest forecast errors among the group of estimators.

Table 4: Diebold Mariano Statistics for the Pooled OLS Model Relative to a Select Number of Benchmarks

One-Step-Ahead Forecasts	
Benchmark Models	**DM Stat**
Non-Spatial FE (One-way Effect)	-69.2339
SAR FE (Two-way Effects)	-481.2382
SAR RE (Two-way Effects)	-0.1801
SDM (Two-way Effects)	-173.7929
SEM (Two-way Effects)	-449.8875
SDPD (One-way Effect)	-75.4297
SDPD (Two-way Effects)	-114.2322

Note: The DM test statistic presented in the table is based on a one-tailed test, where the 1 percent and 5 percent critical values are -2.326 and -1.645, respectively.

The inference from these findings may be limited only to the framework within this particular study. That is, it is possible that the superior predictive ability of the pooled OLS model (or the spatial autoregressive model with random effects) is limited to U.S. state-level energy emissions during this particular timeframe. It is also possible that the predictive abilities of a majority of the spatial models underperformed OLS due a violation of the laws of geography. The first law of geography states: "[everything] is related to everything else, but near things are more related than distant things" (Tobler, 1970). The second law of geography states: "[everything] is related to everything else, but things observed at a coarse spatial resolution are more related than things observed at a finer resolution" (Arbia, Benedetti, and Espa, 1996). The results from Burnett, Bergstrom, and Dorfman (2013) would imply the first law as the regression results and diagnostic tests revealed spatial autocorrelation within the underlying data. The question though is if our geographical units of analysis (i.e., state boundaries) are of too coarse (low) resolution? The results of Burnett, Bergstrom, and Dorfman (2013) do not violate the second law per se, but perhaps the geographical units of analysis are not of fine (high) enough resolution to pick up on the proper spatial autocorrelation in the underlying data? In other words, perhaps finer resolute geographical data (such as county-level observations) would lead to better forecasts over the OLS models. Unfortunately, county-level carbon dioxide emissions are not available at this time, so we do not pursue this further within the current study.

Whether spatial panel data models 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 with random effects perform as well as pooled OLS in a forecasting error context. As mentioned in the introduction, future research may consider the robustness of our findings by examining the performance of different predictors along the lines of Kelejian and Prucha (2007). A next step would be to develop efficient prediction formulae, but unfortunately such an exercise is beyond the scope of the current analysis.

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. Spatial econometric models could help better determine the spatial and temporal distribution of CO_2 emissions in the U.S. 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.

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 spatial econometrics literature. Nevertheless, the spatial econometric models will need to be further tested against empirical reality in the future to help provide evidence of their validity.

REFERENCES

Arbia, Giuseppe, Roberto Benedetti, and Giuseppe Espa. (1996) "Effects of the MAUP on Image Classification," *Geographical Systems*, 3, 123–141.

Aroonruengsawat, Anin, Maximilian Auffhamer, and Alan H. Sanstad. (2012) "The Impact of State Level Building Codes on Residential Electricity Consumption," *Energy Journal*, 33, 31–52.

Auffhammer, Maximilian and Richard T. Carson. (2008) "Forecasting the Path of China's CO_2 Emissions Using Province-level Information," *Journal of Environmental Economics and Management*, 55, 229–247.

Auffhammer, Maximilian and Ralf Steinhauser. (2007) "The Future Trajectory of U.S. CO_2 Emissions: The Role of State vs. Aggregate Information," *Journal of Regional Science*, 47, 47–61.

Baltagi, Badi H. (2005) *Econometric Analysis of Panel Data*. John Wiley & Sons: Hoboken, NJ.

Baltagi, Badi H., Bernard Fingleton, and Alani Pirotte. (2014) "Estimating and Forecasting with a Dynamic Spatial Panel Data Model," *Oxford Bulletin of Economics and Statistics*, 76, 112–138.

Baltagi, Badi H. and Dong Li. (2006) "Prediction in the Panel Data Model with Spatial Correlation: The Case of Liquor," *Spatial Economic Analysis*, 1, 175–185.

Blasing, T.J., Christine T. Broniak, and Gregg Marland. (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," online memo, Carbon Dioxide Information Analysis Center Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, TN, U.S.A. doi 10.3334/CDIAC/00003. Available online in July 2015 at http://cdiac.ornl.gov/trends/emis_mon/stateemis/emis_state.html.

Burnett, James Wesley, John C. Bergstrom, and Jeffrey E. Dorfman. (2013) "A Spatial Panel Data Approach to Estimating U.S. State-Level Energy Emissions," *Energy Economics,* 40, 396–404.

Diaz, Henry F. and Robert G. Quayle. (1980) "Heating Degree Day Data Applied to Residential Heating Energy Consumption," *Journal of Applied Meterology*, 3, 241–246.

Diebold, Francis X. and Roberto S. Mariano. (1995) "Comparing Predictive Accuracy," *Journal of Business and Economic Statistics*, 13, 253–263.

Elhorst, J. Paul. (2014) "Matlab Software for Spatial Panels," *International Regional Science Review*, 37, 389–405

Itkonen, Juha V.A. (2012) "Problems Estimating the Carbon Kuznets Curve," *Energy*, 39, 274–280.

Kelejian, Harry H. and Ingmar R. Prucha. (2007) "The Relative Efficiencies of Various Predictors in Spatial Econometric Models Containing Spatial Lags," *Regional Science and Urban Economics*, 37, 363–374.

LeSage, James P. and R. Kelley Pace. (2009) *Introduction to Spatial Econometrics*. CRC Press: New York.

_____. (2010) "The Biggest Myth in Spatial Econometrics." Social Science Research Network. Available at: http://papers.ssrn.com.

National Oceanic and Atmospheric Administration, U.S. National Climate Data Center. (2010) "Heating and Cooling Degree Day Data," Last accessed online in July 2015 at http://www.cpc.ncep.noaa.gov/products/analysis_monitoring/cdus/degree_days/.

Partridge, Mark D., Marlon Boarnet, Steven Brakman, and Gianmarco Ottaviano. (2012) "Introduction: Whither Spatial Econometrics." *Journal of Regional Science,* 52, 167–171.

Pesaran, M. Hashem, Til Schuermann, and L. Vanessa Smith. (2009) "Forecasting Economic and Financial Variables with Global VARs," *International Journal of Forecasting*, 25, 642–675.

Ryan, David L. and André Plourde. (2009) "Empirical Modeling of Energy Demand," in Joanne Evans and Lester C. Hunt (eds.), *International Handbook on the Economics of Energy*. Edward Elgar: Cheltenham, U.K., pp. 112–143.

Tobler, R. Waldo. (1970) "A Computer Movie Simulating Urban Growth in the Detroit Region." *Economic Geography*, 46, 234–240.

U.S. Department of Commerce, Census Bureau. (2010) "Population Estimates," Available online in July 2015 at http://www.census.gov/popest/states/state.html.

U.S. Department of Commerce, Bureau of Economic Analysis. (2012) "Gross Domestic Product by State," available online in July 2015 at http://www.bea.gov.

U.S. Department of Energy, Energy Information Administration. (2012) "State-Level Energy-Related Carbon Dioxide Emissions, 2000–2009," available online in July 2015 at http://www.eia.gov.

Wooldridge, Jeffrey M. (2002) *Econometric Analysis of Cross Section and Panel Data.* MIT Press: Cambridge, MA.

Yu, Jihai, Robert de Jong, and Lung-fei Lee. (2012) "Estimation for Spatial Dynamic Panel Data with Fixed Effects: The Case of Spatial Cointegration," *Journal of Econometrics*, 167, 16–37.

APPENDIX

As previously mentioned, the forecasts in this study were performed for the various spatial and non-spatial models by using a maximum-likelihood algorithm, with Matlab code provided by Elhorst (2014). The various forecasting models are as follow:

$$
\begin{aligned}
&\text{POLS}: \boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon},\\
&\text{SAR}: \boldsymbol{y} = \lambda(\boldsymbol{I_T} \otimes \boldsymbol{W_N})\boldsymbol{y} + \boldsymbol{X\beta} + (\boldsymbol{\iota_T} \otimes \boldsymbol{I_N})\mu + (\boldsymbol{I_T} \otimes \boldsymbol{\iota_N})\eta + \boldsymbol{\varepsilon},\\
&\text{SEM}: \boldsymbol{y} = \boldsymbol{X\beta} + (\boldsymbol{\iota_T} \otimes \boldsymbol{I_N})\mu + (\boldsymbol{I_T} \otimes \boldsymbol{\iota_N})\eta + \boldsymbol{\varepsilon},\\
&\qquad \boldsymbol{\varepsilon} = \rho(\boldsymbol{I_T} \otimes \boldsymbol{W_N})\boldsymbol{\varepsilon} + \boldsymbol{u},\\
&\text{SDM}: \boldsymbol{y} = \lambda(\boldsymbol{I_T} \otimes \boldsymbol{W_N})\boldsymbol{y} + \boldsymbol{X\beta} + \gamma(\boldsymbol{I_T} \otimes \boldsymbol{W_N})\boldsymbol{X\beta}\\
&\qquad + (\boldsymbol{\iota_T} \otimes \boldsymbol{I_N})\mu + (\boldsymbol{I_T} \otimes \boldsymbol{\iota_N})\eta + \boldsymbol{\varepsilon},\\
&\text{SDPD}: \boldsymbol{y} = \lambda(\boldsymbol{I_T} \otimes \boldsymbol{W_N})\boldsymbol{y} + \gamma \boldsymbol{y_{t-1}} + \rho(\boldsymbol{I_T} \otimes \boldsymbol{W_N})\boldsymbol{y_{t-1}} + \boldsymbol{X\beta}\\
&\qquad + (\boldsymbol{\iota_T} \otimes \boldsymbol{I_N})\mu + (\boldsymbol{I_T} \otimes \boldsymbol{\iota_N})\eta + \boldsymbol{\varepsilon}.
\end{aligned}
\tag{A.1}
$$

To evaluate the forecasting performance, the parameters were estimated for each of the models above based upon the within sample data. Next, we used the estimated parameters from the previous step to forecast state-level emissions for the one-year-ahead and *n*-year-ahead predictions. Finally, we used the forecasting metrics (MAPE, MAE, and RMSE) to evaluate the error performance of the various estimators by comparing the forecasts against empirical reality. No additional assumptions about predictors, such as those outlined in Kelejian and Prucha (2007), were made with the models above. As the forecasting performance evaluation of these various models is still in its infancy, we leave additional prediction specifications and other extensions for future research.