

Regional patterns of U.S. household carbon emissions

William Pizer · James N. Sanchirico · Michael Batz

Received: 15 March 2009 / Accepted: 29 May 2009 / Published online: 9 September 2009
© Springer Science + Business Media B.V. 2009

Abstract Market-based policies to address fossil fuel-related externalities including climate change typically operate by raising the price of those fuels. Increases in energy prices have important consequences for a typical U.S. household that spent almost \$4,000 per year on electricity, fuel oil, natural gas, and gasoline in 2005. A key question for policymakers is how these consequences vary over different regions and subpopulations across the country—especially as adjustment and compensation programs are designed to protect more vulnerable regions. To answer this question, we use non-publicly available data from the U.S. Consumer Expenditure Survey over the period 1984–2000 to estimate long-run geographic variation in household use of electricity, fuel oil, natural gas, and gasoline, as well as the associated incidence of a \$10 per ton tax on carbon dioxide (ignoring behavioral response). We find substantial variation: incidence from the tax range from \$97 dollars per year per household in

Electronic supplementary material The online version of this article (doi:10.1007/s10584-009-9637-8) contains supplementary material, which is available to authorized users.

“One issue for U.S. lawmakers is that the impact of greenhouse gas restrictions would vary by region...” in “Climate Change Debate Hinges On Economics” *Washington Post*, July 15, 2007.

W. Pizer · J. N. Sanchirico
Resources for the Future, Washington, DC 20036, USA

J. N. Sanchirico (✉)
Department of Environmental Science and Policy, University of California,
Davis, CA 95616, USA
e-mail: jsanchirico@ucdavis.edu

M. Batz
Food Safety Program, Emerging Pathogens Institute, University of Florida,
Gainesville, FL, USA
e-mail: mbatz@ufl.edu

New York County, New York to \$235 per year per household in Tensas Parish, Louisiana. This variation can be explained by differences in energy use, carbon intensity of electricity generation, and electricity regulation.

1 Introduction

U.S. households account for approximately 20% of the energy consumed in the United States, spending an average of almost \$4,000 per household per year on electricity, fuel oil, natural gas, and gasoline.¹ Meanwhile, policymakers confronting issues of energy security, rising energy prices, and environmental protection are focusing on market-based policies, such as energy taxes or cap and trade programs that raise energy prices. For example, an analysis of cap and trade bills before the 110th Congress found prices of between \$20 and \$50 per ton of carbon dioxide (CO₂) and increases in fossil energy prices of 15–60% (Paltsev et al. 2007).²

A priori there is no reason to believe that the incidence of market-based policies in the energy sector will be uniform across the United States, as there are significant variations in climate, electricity generation, and historical fuel use. Yet, no fine-scale publicly available spatial data on household energy use exist, because published microdata has the geographic identifiers removed to protect respondents' confidentiality. In this paper, thanks to an agreement with the Bureau of Labor Statistics giving us access to such geographic identifiers, we quantify the geographic variation in energy use and the potential regional impacts of policies that could raise energy prices—particularly those policies related to carbon emissions.

Understanding the spatial incidence and distribution of energy policies is important for at least two reasons. First, if variation among regions is significantly smaller than variation within regions, then regional concerns need not necessarily figure into policy design. But if it is large, as we hypothesize, quantifying this variation can help policymakers design a more equitable program, presumably with more aid and assistance focused on those regions facing greater impacts. Regions that drive more, that are more dependent on coal-based electricity, or that require air conditioning all will face relatively higher burdens under a market-based climate policy. This is particularly relevant for the political economy of carbon regulation: government representation is organized by geography, and members of Congress will want to know whether their constituents will bear an “unfair” burden (Borenstein 2007; Mufson 2007).

Second, with an increasing emphasis on local initiatives, understanding the regional incidence of energy price increases can help with the analysis of state and regional initiatives such as the Regional Greenhouse Gas Initiative in the Northeast or California's new tailpipe standards (RGGI 2005; CARB 2002).

¹Between 2000 and 2005, expenditures on energy rose over 30%, from \$2,985 in 2000 to \$3,887 in 2005. This data is available at <<http://www.bls.gov/cex/home.htm>>. Last accessed April 12, 2007.

²Note that market-based policies that place a price on CO₂ emissions can both raise prices paid by fossil energy consumers and *depress* prices received by fossil energy producers. The aforementioned study shows that typically the bulk of the price effect falls on consumers.

Despite the policy relevance of regional variation in household impacts of energy price increases, there has been little effort to quantify regional patterns of household energy use. State-level data on residential energy use are available from the Energy Information Administration (EIA, www.eia.doe.gov) for more recent periods, but residential end-use is not disaggregated by fuel type, which is an important dimension when investigating the incidence of carbon taxes. Furthermore, states are not necessarily the most interesting geographic unit. States like California, New York, and Florida have potentially important variation within their borders and dense urban areas often times cross state lines (e.g., Metropolitan New York, Washington, Kansas City, and Philadelphia).

Under confidentiality agreements with the Bureau of Labor Statistics, we use non-publicly available data from the U.S. Consumer Expenditure Survey (CES) over the period 1984–2000 to estimate geographic variation in household use of electricity, fuel oil, natural gas, and gasoline, as well as the associated emissions of carbon. Although the non-parametric kernel “smoother” we employ can create estimates at an arbitrary spatial scale across the United States and improve our understanding of the interactions between economic activity and geography, we focus on county-level estimates of annual household use of the four main fuels. But one can easily use the same technique for other levels of detail, such as congressional districts and postal zip codes. In addition to mapping fuel use, we estimate the associated incidence of a \$10 per ton carbon dioxide tax as a means to illustrate the potential value in having finer-scale carbon dioxide emissions data to design carbon policies.

We note that our primary effort is not to estimate the behavioral response to market-based mitigation policies, which have been well-studied elsewhere in the literature—including the responsiveness of the residential sector to higher energy prices (Dahl 1993).³ Instead, we wish to demonstrate the regional variation in burden. Behavioral responses, estimated in the aggregate, can then be layered on top of the regional variation, a calculation we do in the text.

2 Methods

Even with the CES data pooled over 15 years, the geographic coverage is not complete enough to compute county or even sub-state averages at a sufficiently fine level of detail. Out of more than 3,100 counties in the lower 48 states, the CES has sampled households in only 666 over this period; these 666 sample counties are disproportionately in urban areas, leaving large areas of the rural west unsampled. If data were available at the level of detail for every region (county) that interests us, we could simply use those averages. Alternatively, we could construct simple regressions of energy use on household demographic variables and dummy-variables for each region, again allowing us to pick out county averages—this time, controlling for demographics. Here, the coefficients on the regional dummy-variables would

³While a behavioral response is likely, the degree to which previous estimates that were calculated during a period when consumer choices were limited due to regulations on energy efficiency is unclear as many consumers might have been “forced” to choose greater levels of efficiency than they otherwise preferred.

provide the information we sought. But because the data are thin or non-existent over wide areas of less populated, rural areas, these techniques are not appropriate.

Depending on the questions asked, the literature on estimating household (or unit-level) energy demand utilizes a wide variety of techniques. For example, research focused on durable equipment purchases and long-run versus short-run demand responses typically use two-stage reduced form models that capture the purchase decision and decision on how to use the equipment (Baker and Blundell 1991; Bernard et al. 1996; Dubin and McFadden 1984). On the other hand, forecasting energy use often entails detailed demographic models (Raupach et al. 2007; O'Neill and Chen 2002). In both cases, regional variation (if present) is addressed using dummy variables for broad geographic identifiers available in the data. Where regional variation has been the object of a statistical effort, demographic and degree-day data has often been used to substitute for detailed geographic information using econometric modeling (typically geographic identifiers are removed from available data for confidentiality reasons).⁴

Spatial econometric analysis is another approach suited for cases where interest centers on a regression model and there is a need to address spatial autocorrelation and heteroskedasticity in the residuals (Anselin 1988; Anselin et al. 2004). The spatial econometric approach is comparable to our formulation when we employ the demographic information.⁵ However, given our simple focus on the data's spatial pattern itself, we utilize a non-parametric, kernel regression (Silverman 1986). Here, the term "regression" is almost misleading; this technique provides a transparent way to estimate local energy as an average of spatially near observations and a statistically consistent framework to construct standard errors and handle covariates.

In other words, kernel regression applied to spatial data creates an estimate of energy use at a particular geographic point based on an average of energy use at nearby, sampled points. It also allows for the fact that data for some of the sampled counties, with many more observations than others, contains less sampling error compared to other sampled counties with fewer observations.

We begin with a sample of observations of fuel quantities Y_1, \dots, Y_I in a set of sampled counties $1, \dots, I$ located at coordinates $\mathbf{x}_1, \dots, \mathbf{x}_I$. We assume

$$Y_i = f(\mathbf{x}_i) + e_i \quad (1)$$

where $f(\mathbf{x})$ describes the relevant spatial pattern of fuel use, and e_i are errors with $E[e_i] = 0$. Given our knowledge of the number of observations collected for each

⁴Various researchers have used localized information—on household characteristics, fuel use, or climate data (heating/cooling degree days) at the county or finer level—coupled with regression models of aggregate data on these variables to spatially disaggregate data without detailed geographic identifiers. See, for example, May et al. (1996), Osborn et al. (1999).

⁵In our demographic model, we construct an initial estimate of the regional energy use pattern and then we estimate a demographic model of energy use in a second stage with the spatial effects netted out. This is analogous to a spatial econometric approach that first estimates the spatial correlation pattern, then estimates the regression model on a transformed data set that is independently and identically distributed. This approach is very similar to a geographically weighted regression (for more information on these type of analysis, see Fotheringham et al. 2002).

county, n_i , we assume $E[e_i^2] = \sigma^2/n_i$, where σ^2 is the variance associated with an individual household observation.

Our estimate of energy use at an arbitrary point \mathbf{x} , $\hat{f}(\mathbf{x})$, given our set of I observations is

$$\hat{f}(\mathbf{x}) = \sum_{i=1}^I \left(\frac{n_i w_i}{\sum_j n_j w_j} \right) \cdot Y_i = \sum_{i=1}^I \left(\frac{n_i w(\|\mathbf{x} - \mathbf{x}_i\|/b(\mathbf{x}))}{\sum_j n_j w(\|\mathbf{x} - \mathbf{x}_j\|/b(\mathbf{x}))} \right) \cdot Y_i \tag{2}$$

where $\|\mathbf{x} - \mathbf{x}_i\|$ is the Euclidian distance between points \mathbf{x} and \mathbf{x}_i , $w(\cdot)$ is the “kernel” or weighting function, and $b(\mathbf{x})$ is the “bandwidth” that can vary over location \mathbf{x} . We include the number of observations n_i in the weighting scheme, because there are n_i observations of Y_i located at \mathbf{x}_i . By doing so, we are weighting county observations by their precision (one over the variance), which is proportional to n_i . Note that kernel regression is simply a weighted average, where the art is defining the kernel function and bandwidth (Watson 1964; Nadraya 1964).⁶

As a weighted average, it is straightforward to compute the variance associated with our estimator. First we estimate the sampling error for individual observations, $\hat{\sigma}^2(x)$ given by

$$\hat{\sigma}^2(x) = \sum_{i=1}^I \left(\frac{n_i w(\|\mathbf{x} - \mathbf{x}_i\|/b(\mathbf{x}))}{\sum_j n_j w(\|\mathbf{x} - \mathbf{x}_j\|/b(\mathbf{x}))} \right)^2 \sigma_i^2 \tag{3}$$

where σ_i^2 are observations of sampling error in observed counties. Then, we divide this sampling error for individual observations by the effective number of observations used to compute $\hat{f}(x)$ implied by an estimate of the observation density $\hat{g}(x)$, chosen bandwidth $b(x)$, and total number of observations in the sample n , namely $\hat{g}(x)nb^2(x)$.⁷

Turning to the kernel function, we can limit our search by noting that the function should be non-negative and monotonically decreasing, so that weights are never negative and decrease as distance increases.⁸ There are still many possibilities, including constant, linear, quadratic, and Gaussian functions. We choose the Gaussian function where $w(d/b) = \exp(-(d/b)^2)$. The effect of this function is that, relative to the weight at a distance of zero, the weight at $d/b = 0.5$ is 0.78, the weight at $d/b = 1$ is 0.37, and the weight at $d/b = 2$ is 0.02. The principal advantage of this function is that it is not piecewise defined and therefore quite smooth—a useful property when we next determine the bandwidth.

To choose the bandwidth, we use cross validation (Stone 1974) that minimizes the difference between the predicted and observed values at the sample points \mathbf{x}_i . That is, pick b to minimize $\sum_{i=1}^I (\hat{Y}_i - Y_i)^2$, where \hat{Y}_i is the predicted value using

⁶We can alternatively estimate the regional variation controlling for demographics by replacing Y_i in Eq. (2) with $Y_i - \hat{\beta}z_i$ where z_i are covariates and $\hat{\beta}$ is an partial regression coefficient on the covariates. This is further discussed in the [Electronic supplementary material](#).

⁷Imagine a two-dimensional box of size $b(x)$ by $b(x)$ with observations spread out with $\hat{g}(x)n$ per square unit.

⁸Negative weights have been proposed as a way to reduce bias but they also introduce many problems; see Section 3.6 of Silverman (1986).

all observations except the i th one (see [Electronic supplementary material](#) for more information).

We make one further modification by allowing b to vary by location based on the density of observations. That is, where there are fewer observations we use a larger bandwidth. In particular, we specify $b(\mathbf{x}_i) = (\hat{g}(\mathbf{x}_i)/\bar{g})^{-1/5} b$, where $\hat{g}(\mathbf{x}_i)$ is again an estimate of the density of observations at point \mathbf{x}_i , \bar{g} is the geometric mean of the density estimate, and b , again, is a global bandwidth parameter now chosen by cross validation with the specified spatial variation.⁹ We considered further estimates of local bandwidths, attempting to numerically minimize estimated mean-squared error but with little success. The supplementary material includes additional details on the adaptive-bandwidth approach.

We use the adaptive cross-validation approach to compute bandwidths for each fuel independently. This same fuel-specific adaptive bandwidth is also used to estimate the variance in household fuel demand, and, in turn, the variance of our county-level estimates.

For each of the three primary fuels, we compute CO₂ incidence by multiplying estimated fuel demand by the carbon content of the fuel and the reference price of CO₂ used in our analysis (\$10 per ton). Electricity is considerably more complicated because not only does the carbon content of electricity vary by location, the pricing of that carbon content varies depending on whether a particular region has competitive or regulated electricity markets. To provide a realistic estimate of how much consumers will be affected by a market-based CO₂ policy, we use regional estimates of electricity price impacts associated with a given CO₂ price and apply that to our reference CO₂ price of \$10 per ton (Burtraw and Palmer 2006). The CO₂ incidence associated with electricity is then the estimated county-level electricity demand multiplied by this change in regional electricity prices.

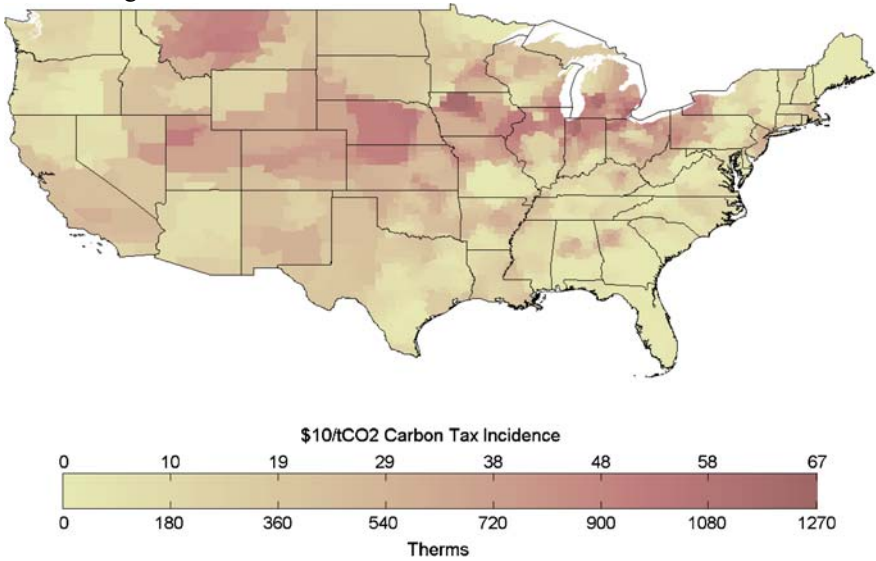
We describe the methods for including covariates in Eq. (1), the data used to run ordinary least squares regressions of household demographics and characteristics on the use of each of the four fuels, and the regression results in the [Electronic supplementary material](#). Overall, the regression results are compelling for a number of reasons. Although one always has to be careful not to place too much emphasis on R^2 values, the model's explanatory power is within reasonable limits. Second, the parameter estimates are largely consistent across the fuels, although the estimates for fuel oil do not perform as well.

3 Results

Figure 1 shows our estimates for the patterns of electricity, natural gas, fuel oil, and gasoline use at the county level. The scale on each figure indicates both the energy use and associated incidence from a \$10 per ton carbon dioxide emissions tax, except in the case of electricity, where carbon intensity varies by location. Figure 2 shows total CO₂ incidence and CO₂ incidence of electricity only with and without demographic controls. CO₂ incidence refers to the amount fuel prices will rise based on a given

⁹Although the intuition that we want a larger bandwidth where the density is smaller is straightforward, the specific exponent $-1/5$ is not; see Silverman (1986), p. 104.

A: Natural gas



B. Fuel oil

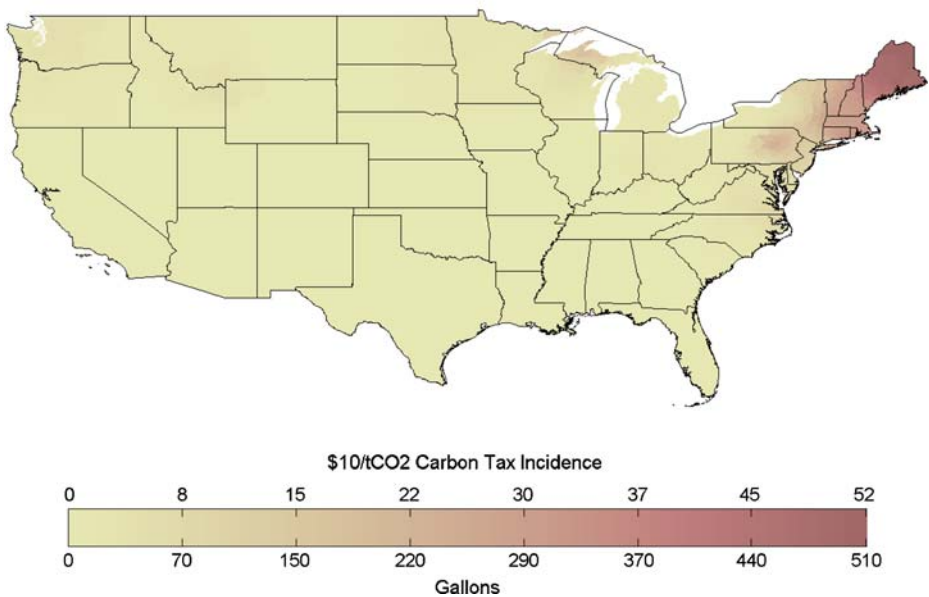
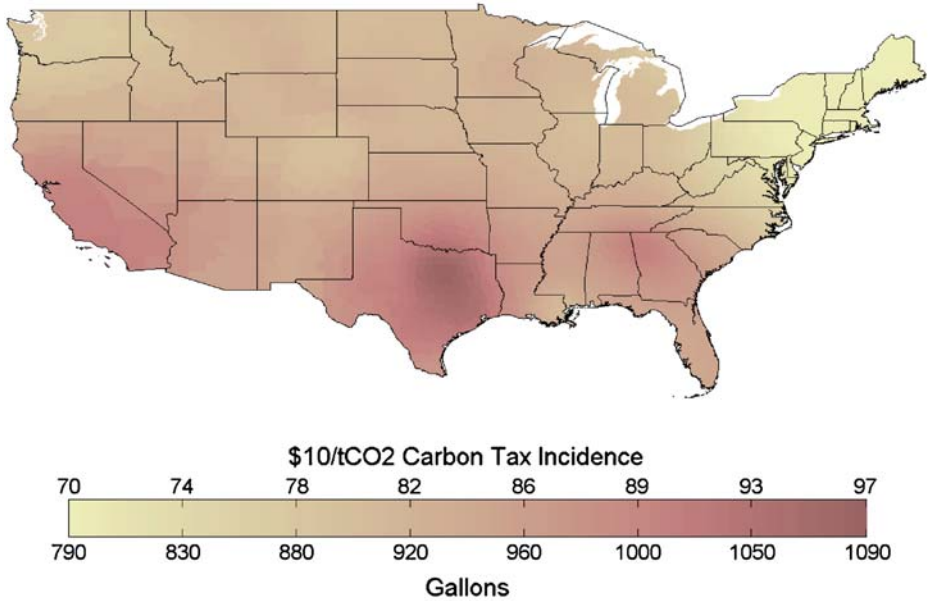


Fig. 1 Regional fuel use and carbon dioxide tax incidence. Note: panel (a) is natural gas (carbon emissions are calculated by converting to MMBTU and multiplying by 11.7), panel (b) is fuel oil (carbon emissions are 19.6 per gallon), panel (c) is gasoline (carbon emissions are 22.4 per gallon), and panel (d) is the quantity of fuel use for electricity. The carbon dioxide incidence for electricity is illustrated in Fig. 2 panel (b), as it depends on the generation of electricity

C. Gasoline



D. Electricity

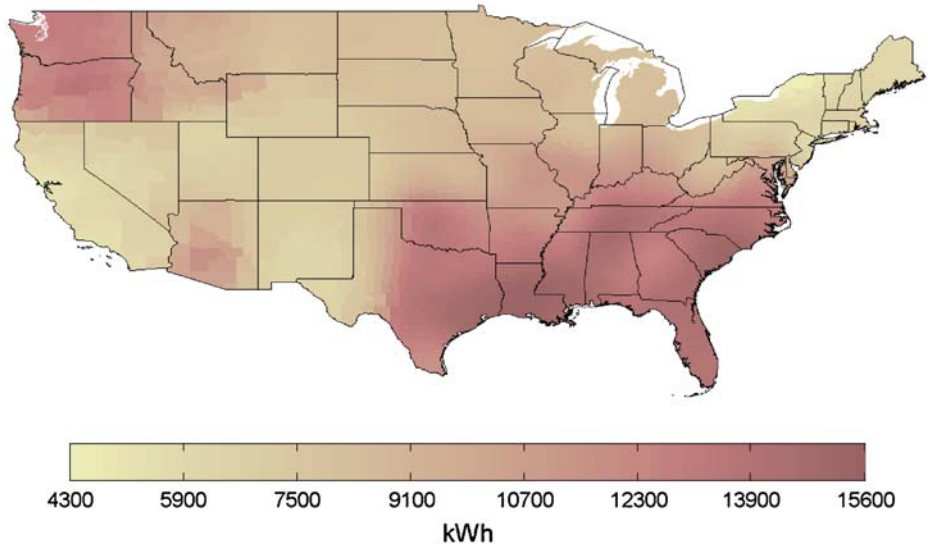
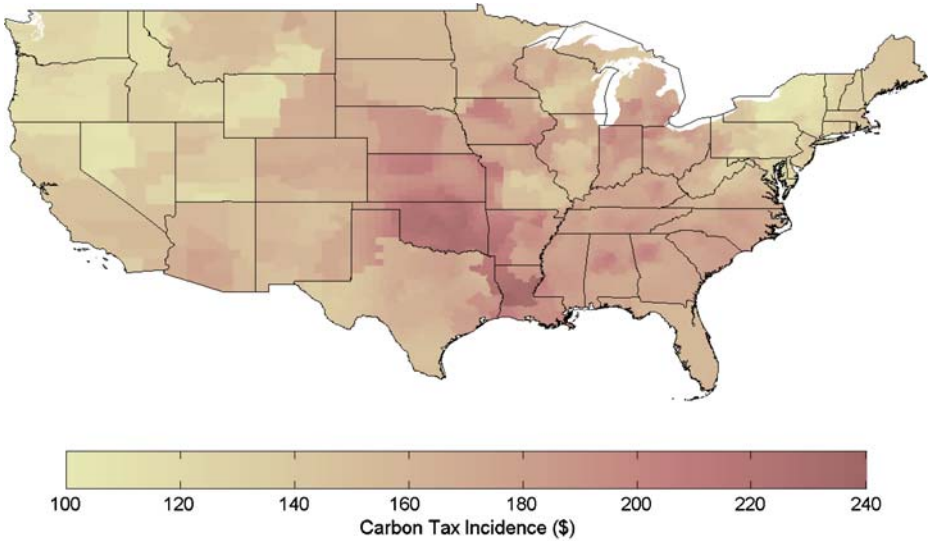


Fig. 1 (continued)

CO₂ price and is directly proportional to the carbon content of the fuels. The CO₂ incidence for electricity is more complex, depending on the market structure in a given region (see [Electronic supplementary material](#)).

The estimated pattern of natural gas use in Fig. 1, panel a, is clearly concentrated in the Midwest, and rather sparse in the South. This is consistent with EIA estimates

A: Total Carbon incidence



B: Electricity only

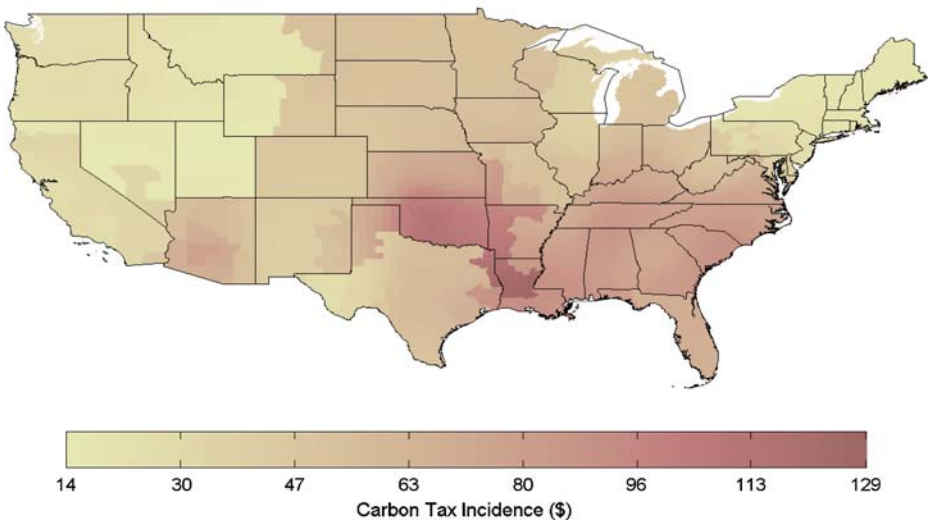
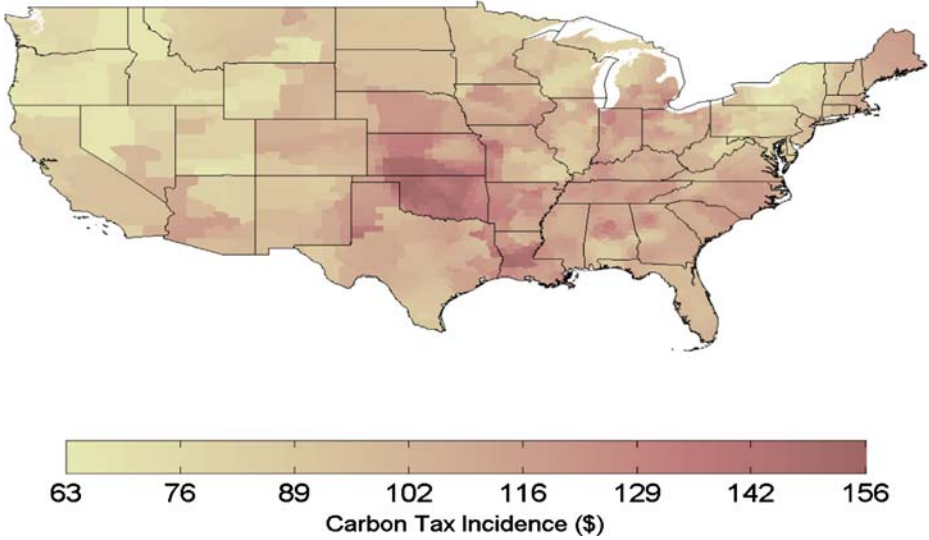


Fig. 2 Total carbon tax incidence. Note: panel (a) is the carbon dioxide tax incidence across all fuels, panel (b) is carbon dioxide tax incidence for electricity only, panel (c) is the total, and panel (d) is carbon dioxide tax incidence for electricity after controlling for household demographics

in Table 1,¹⁰ showing use in the Midwest to be two to three times use in the South and West. Compared to the other fuels in Fig. 1, we also find more local variation

¹⁰The census regions are Northeast (ME, NH, VT, MA, RI, NY, PA, NJ, CT), Midwest (OH, MI, IN, IL, WI, MN, IA, MO, ND, SD, NE, KS), South (DE, MD, DC, WV, VA, NC, SC, GA, FL, KY, TN, AL, MS, LA, AR, OK, TX), and West (WA, OR, CA, NV, MT, ID, WY, UT, CO, AZ, NM, AK, HI). We do not include Hawaii and Alaska in our analysis.

C. Total carbon incidence for median household



D. Electricity only for median household

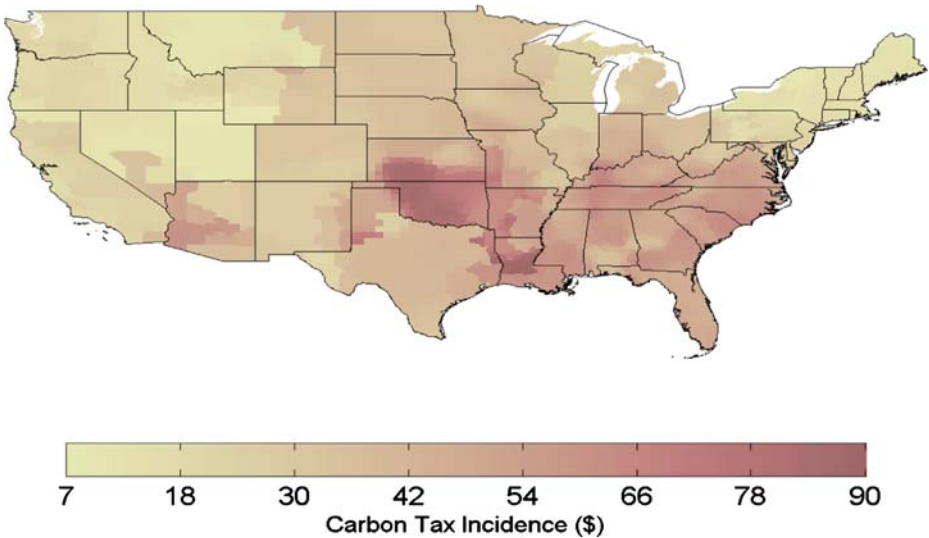


Fig. 2 (continued)

in natural gas usage, with adjacent counties sometimes using significantly different levels of natural gas.

Our estimates of fuel oil use in Fig. 1, panel b, show what we would expect: heavy use in New England and the upper peninsula of Michigan, trailing down into portions of New York, New Jersey, and Pennsylvania, but otherwise virtually

Table 1 Annual household consumption estimates relative to EIA estimates

	Median county	Mean county	Std. errors	EIA 2001	EIA 1997	EIA 1993*
Electricity (kWh)						
Northeast	6,303	6,454	117	7,354	6,780	6,731
Midwest	8,501	9,012	202	9,210	8,681	8,688
South	13,694	13,192	310	14,163	13,276	12,980
West	6,481	7,710	241	7,916	8,294	8,052
National	8,796	9,762	233	10,332	9,822	9,615
Natural gas (therms)						
Northeast	430	384	59	470	489	543
Midwest	769	716	120	712	875	827
South	192	216	67	289	307	347
West	435	405	88	369	416	444
National	383	406	82	440	502	527
Fuel oil (gallons)						
Northeast	137	165	15	199	276	267
Midwest	11	11	5	17	30	37
South	1	10	3	11	15	23
West	2	8	4	3	7	11
National	4	40	6	46	69	74
Gasoline (gallons)						
Northeast	798	801	7	868	705	695
Midwest	904	906	11	1,074	918	857
South	953	955	13	1,104	974	908
West	1,001	967	14	982	880	882
National	915	916	12	1,025	886	846

Note: The source for regional electricity, natural gas, and fuel oil is the Energy Information Administration (EIA), Residential Energy Consumption Survey (<http://www.eia.doe.gov/emeu/recs/>), and the source for gasoline is EIA Household vehicle energy use survey (http://www.eia.doe.gov/emeu/rtecs/nhts_survey/2001). All years of the two surveys align, except that the gasoline survey is 1994 not 1993 (labeled with an *). The EIA estimates are average consumption for a household in the region and our estimates include both averages and medians (across county estimates). EIA estimates have been adjusted for a slight discrepancy between EIA and BLS household definitions (“housing unit” versus “consuming units”) and the RTECS basis of “household with a vehicle”

nothing.¹¹ Within these states, however, household use varies considerably within each state as does use across states. For example, in Maine, there is a 30% difference between the highest and lowest counties. Given New York’s varied landscape, it is no surprise that it has significant within-state variation both in terms of percentage differences and levels. Although sporadic use exists here and there, there is little evidence of use in the South and West.

Gasoline usage shown in Fig. 1, panel c, shows remarkably little variation. From a high of 1,090 gallons per year in Texas to a low of 790 gallons in New York and Maine, gasoline use fluctuates (geographically) much less than any other fuel. EIA estimates confirm the relative uniformity of gasoline use. Perhaps predictably, southern California, the Houston–Dallas corridor in Texas, and Atlanta show higher

¹¹Despite lower use per household in New York, New Jersey, and Pennsylvania, *total* fuel oil use is larger in these states than in New England because of their higher populations.

household usage, and the more densely populated northeastern region shows some of the lowest.¹²

Looking at panel d in Fig. 1, the principal observation is that electricity usage in the South and Southeast is much higher than most other regions of the country. This understandably arises from the need to operate air conditioning units most of the summer and drives usage to more than 11,000 kWh in the South. According to EIA, air conditioning and space heating comprise 26% of the end-use consumption of residential electricity consumption in 2001. Another observation is the degree of heterogeneity within each state and census region. For example, the western census region has highs of about 10,700 kWh in the Washington and Oregon and lows of about 7,000 kWh in California.

With natural gas, fuel oil, and gasoline, we can easily convert fuel use to emissions using CO₂ coefficients from the Department of Energy and then multiply by \$10 per ton to get predicted CO₂ incidence levels. However, when we shift to carbon emissions from electricity production in Fig. 2, panel B, we need to take into account the location and type of generation. The implication is that the pattern we observe in fuel use does not directly map onto CO₂ incidence.

Although coal-fired power plants lead to high levels of carbon emissions per kilowatt-hour, hydropower in the Northwest and nuclear power in parts of New England and the Southeast have no emissions. Natural gas generation contains roughly half the carbon emissions per Btu compared to coal and has a considerably lower heat rate/higher conversion efficiency. In addition to the generation mix, differences in market structure are equally important in deriving CO₂ incidence measures because market structure affects how electricity prices will change in response to underlying fuel price changes. In particular, the Northeast is deregulated, and prices are set by the marginal producer in the market (sometimes gas, sometimes coal). The South, on the other hand, is regulated and prices are set at average cost, across all fuels, by regulators. Other regions face some mix of regulation and deregulation. The result of these differences is a concentration of electricity-related household CO₂ incidence in the Texas–Louisiana region as well as around southern Indiana and Ohio. The southern census region has an average incidence level approximately 65% greater than the northeastern region, which has the lowest mean level.

Figure 2, panel a, aggregates up the CO₂ incidence of the \$10 per ton CO₂ tax across the four fuels, and we find average CO₂ incidence levels per household ranging from \$240 per year in portions of the southern plains and Texas to less than \$120 in the Northeast and scattered throughout the West. Table 2 presents the differences in incidence within each census region. These results sharply focus the concerns we noted at the outset: climate change policies that attempt to discourage carbon dioxide by raising the price of fossil fuels in relation to their CO₂ emissions will have different effects on households in different parts of the country. Impacts in some regions of the

¹²As a check on our results, we can aggregate up our estimates to total gasoline consumption as follows. First, we take our national average of about 900 gallons per household over 1984–2000 and multiply it by 110 million households, which yields about 100 billion gallons. Next, we account for the fact that non-household motor-gasoline use accounts for about 15% of total gasoline use, yielding 115 billion gallons of total demand. Finally, when we adjust for the 10% growth from our sample period to 2001 and the 10% since, we have roughly total annual motor gasoline consumption as found in *Annual Energy Review 2006*, Washington: Energy Information Administration.

Table 2 Annual household incidence of a \$10/t CO₂ tax by census region

	Min County	Mean County	Median County	Max County
All energy				
Northeast	97.58	132.87	131.09	172.43
Midwest	126.52	172.57	171.91	218.35
South	117.64	169.82	171.23	235.44
West	105.84	142.72	144.49	173.29
National	97.58	157.42	155.24	235.44
Electricity				
Northeast	15.54	24.62	21.56	54.03
Midwest	36.64	52.69	52.26	96.93
South	27.44	72.21	72.42	129.12
West	13.50	34.00	28.67	77.67
National	13.50	50.21	51.25	129.12

country are double the impact in others, and impacts in some counties are double the impacts in other counties within the same region.

The flip side of CO₂ incidence is that the counties with higher incidence have greater levels of household carbon emissions or are “dirtier” than other counties. The counties with total incidence greater than the 95 percent quartile are in Louisiana, Oklahoma, Arkansas, Texas, Kansas, Missouri, and Alabama. The cleanest counties—those with incidence levels less than the 5 percent quartile—are mostly in New York, Washington, Oregon, Idaho, Vermont, Pennsylvania, and Nevada. The median county is Santa Cruz County in Arizona with total incidence of \$116.

So far our discussion has ignored the fact that the CES collects extensive data on the characteristics of households participating in the survey. Many of these—notably income and household size—would be expected to have a large influence on energy use. To the extent that there is substantial and persistent regional variation in these characteristics, the influence of these characteristics will be associated purely with regional variation based on the preceding approach.

Whether we want to estimate regional variation in CO₂ incidence controlling for this regional variation in household characteristics depends on the particular question. If we ask about the regional incidence of higher energy prices, there is no need to control for these variables; we simply need to look at the regional patterns of energy use. However, we could ask how the energy use of a *particular* kind of household would vary depending on where it is located. In this case we would want to hold constant characteristics and isolate the effect of geographic variation. For example, policymakers might be interested in how low-income households might be affected in different parts of the country.

Figure 2, panels c and d, illustrate the incidence levels of the median household (e.g., median income, household size, etc.) if you placed them in any particular county.¹³ Overall, the strong geographical patterns found in panels a and b remain, but the tax burdens are lower, everything else equal, reflecting the skewed distribution of income and its impact on energy demand (median income is roughly half

¹³Note that the variation in Fig. 2 panels c and d would be the same for any type of household based on our model of energy use; only the level varies with demographic characteristics.

mean income). The amount of relative variation (ratio of highest to lowest values) is similar in panels c and d, reflecting the fact that the bulk of overall variation is driven by geography, not demographic differences.¹⁴

Although the patterns of energy use and CO₂ incidence appear to coincide with our intuition, it is useful to verify our estimates against other sources. Unfortunately, we are not aware of another set of estimates at the county level. Table 1 compares our estimates aggregated (using population weights) to census regions versus household average levels produced by the EIA in 1993, 1997, and 2001. In general, we find that the pattern of our mean aggregate estimates generally matches the EIA within the standard error of our estimates.¹⁵ The notable discrepancies are lower estimates in our data for natural gas and fuel oil everywhere but the West.

There are several potential explanations for these discrepancies in Table 1. First, our sample covers 1984–2000 versus available EIA data on regional household energy use, which is only available for 1993, 1997, and 2001. Looking at national data from EIA indicates this difference explains about one-quarter of the total discrepancy for natural gas (and does nothing for fuel oil). A second explanation is that the CES oversamples urban areas¹⁶ and urban households are less likely to use natural gas and especially fuel oil. This effect is corrected in the demographic model where a control variable for urban/rural is used to predict fuel use.

As noted earlier, these estimates ignore any behavioral response. That is, the estimates assume consumers pay higher energy costs without investing in higher efficiency equipment or choosing to forego energy services. In reality, we would expect both to occur, with larger responses as more time elapses and capital stock turns over. Any of these kinds of mitigation efforts necessarily improve the consumers' outcome relative to no behavioral response—in this way our estimates provide an *upper bound* on actual impacts. Further, assuming behavioral responses are roughly proportional, say, offsetting 15% of the initial impact absent any adjustment, the *relative* impacts across counties will remain the same. Those counties with relatively high impacts will continue to experience relatively high impacts.

Consulting the literature (Dahl 1993) we find short run elasticities of around 0.15 and longer-run elasticities approaching unity or higher. These values suggest that, in the short run, the pattern of impacts is unlikely to be much different than indicated in Figs. 1 and 2 with only 15% of the impact on energy prices offset by lower energy consumption. In the longer run, reduced energy use might fully offset higher

¹⁴This may not be true, however, at the very detailed level—in a particular county and year and for a particular fuel. For instance, when we look at variation driven by solely by demographics (based on the 2000 Census and our demographic control estimates), we find factors of 2, 3, 1.6, and 3 for natural gas, fuel oil, gasoline, and electricity, respectively, between the lowest and highest estimates. We find, therefore, that the demographic variation is no larger than the geographic variation we found for all fuels except perhaps gasoline (comparing Fig. 2a and c). Having said that, the demographic factors suggest it is possible that at a fine level, demographic data could reveal additional variation. A recent study that lends support to our methods and results was a cross-country analysis that found that access to natural resources and weather were significant explanatory variables in explaining differences in emissions even between countries with very similar income levels (Neumayer 2002).

¹⁵A (small) portion of the difference in the West is because EIA includes Hawaii and Alaska in their estimates and we do not.

¹⁶The purpose of the Consumer Expenditure Survey is to provide expenditure weights for the Consumer Price Index, which is then constructed for various metropolitan areas.

energy prices. In this case, mitigation costs would be reflected in lower welfare (from reduced energy services) and higher capital costs (for greater energy efficiency). It seems reasonable that these costs would continue to be roughly equal to some fraction of the initial impact on fuel expenditures.¹⁷

4 Discussion

Increases in energy prices—whether arising from carbon policies or from other world/national events—will have different impacts on households in different regions of the country. Electricity use is higher in the South, fuel oil is used almost entirely in the Northeast, and natural gas is much more prevalent in the Midwest. Gasoline usage fluctuates very little.

Policymakers need to be keenly aware of these differences as they contemplate climate change policies that would put a price on CO₂ emissions. Such policies place a heavier burden on regions where geography and history have led to higher emissions per household. The South requires air conditioning. The Midwest is reliant on coal for electricity generation. Californians drive more.

Using data from the CES over the period 1984–2000, this paper has tried to quantify exactly how much these regional differences matter. While general patterns are well-known, our analysis implies that one size does not fit all. In fact, we find that there exists considerable variation in the type of fuel used and the average level of use across regions, states, and even counties. We also find substantial variation in the incidence of a CO₂ tax. This variation can be explained both by differences in energy use as well as differences in the carbon intensity of electricity generation.

An interesting observation is that the regions of the country where incidence is lower—the Northeast and West Coast—happen to be those where recent initiatives have been implemented to address climate change through state programs. Whether the lower incidence is partly responsible for those efforts or whether it reflects a much longer-term trend toward cleaner/more efficient energy use is unclear.

What do our results quantitatively imply about the regional incidence of climate change policies? The overall result for carbon emissions is that households in the Northwest United States would likely bear half the burden of households in the Midwest and Texas where per household incidence is almost double. If equity is a concern, policymakers will need to contemplate ways that this burden can be equalized—a process that most likely would occur in the design phase of the policy. We should point out, however, that while the focus of this work has been regional incidence, it is also important to consider the distribution across other key household variables, such as income (Dinan 2000) and more generally, how one might measure interregional equity (Rose and Zhang 2004).

¹⁷With roughly linear marginal mitigation costs and a unitary long-run elasticity, it is easy to show that the long-run net cost impact would be $x(1 + \frac{1}{2}x)/(1 + x)$, where x is the original, fractional increase in energy costs. Here, the long-run impact is some fraction of the original impact x , reflecting the cost of fuel savings that range linearly between the original price p and new, higher price $(1 + x)p$.

Acknowledgements We want to especially thank Steve Henderson, Wolf Weber, Chris Chapman, and Mark Vendemia at the Office of Prices and Living Conditions at the U.S. Bureau of Labor Statistics for their support, Drew Bagliano and David Evans for their valuable research support, Dallas Burtraw and Karen Palmer for information on U.S. electricity prices, and three anonymous referees and Steve Schneider for valuable comments. Funding for this research was provided by the Energy Foundation and Resources for the Future. William A. Pizer is currently the deputy assistant secretary for environment and energy at the Treasury Department; his work on this paper was completed while he was a full-time senior fellow at RFF.

References

- Anselin L (1988) *Spatial econometrics: methods and models*. Kluwer, Dordrecht
- Anselin L, Florax RJGM, Rey S (eds) (2004) *Advances in spatial econometrics: methodology, tools, and applications*. In: *Advances in spatial science series*, p 513, ISBN: 978-3-540-43729-1. Springer
- Baker P, Blundell R (1991) The microeconomic approach to modeling energy demand: some results for UK households. *Oxf Rev Econ Policy* 7(2):54–76
- Bernard J-T, Bolduc D, Belanger D (1996) Quebec residential electricity demand: a microeconomic approach. *Can J Econ* 29(1):92–113
- Borenstein, S (2007) Planet's woes? Blame coal: Texas leads in carbon emissions overall; Wyoming is tops per capita. Associated Press
- Burtraw D, Palmer K (2006) "Shifting the burden and spreading the wealth: who bears the cost of climate policy in the electricity sector?" RFF discussion paper 06–52 (July). Available at www.rff.org
- California Air Resources Board (CARB) (2002) *Reducing climate change emissions from motor vehicles, vol 2004*. Sacramento, California Air Resources Board. Available at <http://www.arb.ca.gov/cc/factsheets/ccfactsheet.pdf>
- Dahl C (1993) A survey of energy demand elasticities in support of the development of NEMS. http://ftp.eia.doe.gov/pub/oiaf/elasticitysurvey/elasticitysurvey_dahl.pdf
- Dinan T (2000) Who gains and who pays under carbon-allowance trading? The distributional effects of alternative policy designs. Congressional Budget Office, Washington
- Dubin JA, McFadden DL (1984) An econometric analysis of residential electric appliance holdings and consumption. *Econometrica* 52(2):345–362
- Fotheringham AS, Brunson C, Charlton ME (2002) *Geographically weighted regression: the analysis of spatially varying relationships*. Wiley, Chichester
- May C, Brown R, Banwell P, Foery K, Offutt S, Wilson A (1996) Increasing the use of energy-efficient HVAC equipment in homes through the use of voluntary and profitable pollution prevention programs. In: *Proceedings of the 1996 ACEEE summer study on energy efficiency in buildings*. American Council for an Energy Efficient Economy, Washington
- Mufson S (2007) Climate change debate hinges on economics; lawmakers doubt voters would fund big carbon cuts, p A01. *Washington Post*
- Nadraya EA (1964) On estimating regression. *Theory Probab Appl* 15:134–137
- Neumayer E (2002) Can natural factors explain any cross-country differences in carbon dioxide emissions? *Energy Policy* 30:7–12
- O'Neill, BC, Chen BS (2002) Demographic determinants of household energy use in the United States. In: *Population and development review* 28, Supplement: population and environment: methods of analysis, pp 53–88
- Osborn J, Marnay C, McMahon J (1999) Modeling the effects of appliance standards through geographic analysis. 1999 ESRI user conference proceedings. <http://gis.esri.com/library/userconf/proc99/navigate/proceed.htm>
- Paltsev S, Reilly JM, Jacoby HD, Gurgel AC, Metcalf GE, Sokolov AP, Holak JF (2007) *Assessment of U.S. cap-and-trade proposals, report #146*. MIT Joint Program on the Science and Policy of Global Change, Cambridge, MA
- Raupach M, Marland G, Ciais P, Le Quéré C, Canadell JG, Klepper G, Field CB (2007) Global and regional drivers of accelerating CO₂ emissions. *PNAS* 104(24):10288–10293
- Regional Greenhouse Gas Initiative (RGGI) Staff working group (2005). In: *RGGI memorandum of understanding*. RGGI Staff Working Group, Albany
- Rose A, Zhang ZX (2004) Interregional burden-sharing of greenhouse gas mitigation in the United States. *Mitig Adapt Strategies Glob Change* 9(3):477–500

- Silverman BW (1986) Density estimation for statistics and data analysis. In: Cox D, Reid N, Rubin DB, Silverman BW (eds.) Monographs on statistics and applied probability, vol 26. Chapman and Hall, London
- Stone M (1974) Cross-validated choice and assessment of statistical predictions (with discussion). *J R Stat Soc B* 36:111–147
- Watson GS (1964) Smooth regression analysis. *Sankhya Ser A* 26:359–372