Carbon Mitigation Costs for the Commercial Sector: Discrete-Continuous Choice Analysis of Multifuel Energy Demand

Richard G. Newell and William A. Pizer

June 2005 • Discussion Paper 05-13

RESOURCES
FOR THE FUTURE

1616 P Street, NW
Washington, DC 20036
Telephone 202-328-5000
Fax 202 939-3460

Carbon Mitigation Costs for the Commercial Sector: Discrete-Continuous Choice Analysis of Multifuel Energy Demand

Richard G. Newell and William A. Pizer

Abstract

We estimate a carbon mitigation cost curve for the U.S. commercial sector based on econometric estimation of the responsiveness of fuel demand and equipment choices to energy price changes. The model econometrically estimates fuel demand conditional on fuel choice, which is characterized by a multinomial logit model. Separate estimation of end uses (e.g., heating, cooking) using the 1995 Commercial Buildings Energy Consumption Survey allows for exceptionally detailed estimation of price responsiveness disaggregated by end use and fuel type. We then construct aggregate long-run elasticities, by fuel type, through a series of simulations; own-price elasticities range from –0.9 for district heat services to –2.9 for fuel oil. The simulations form the basis of a marginal cost curve for carbon mitigation, which suggests that a price of $20 per ton of carbon would result in an 8% reduction in commercial carbon emissions, and a price of $100 per ton would result in a 28% reduction.

Key Words: commercial energy demand, carbon policy, climate change, discrete choice

JEL Classification Numbers: Q28, Q48, Q41, C35, C15

Table of Contents

Carbon Mitigation Costs for the Commercial Sector: Discrete-Continuous Choice Analysis of Multifuel Energy Demand

Richard G. Newell and William A. Pizer*

1. Introduction

The commercial sector is one of four important energy end-use sectors—along with residential, industrial, and transportation uses—and was responsible for 245 million metric tons of carbon in 1998, or 16% of U.S. energy-related carbon emissions (U.S. EIA 2002). Since 1990, commercial emissions have grown at an average of about 2% per year, faster than any of the other end-use sectors. These emissions result from the lighting, heating, cooling, and other energy operating requirements of commercial buildings such as stores, offices, restaurants, hotels, religious organizations, schools, and other public buildings. Any assessment of the likely magnitude and cost of policies for reducing carbon emissions from the commercial sector therefore quickly becomes a problem of estimating changes in the quantity and types of energy used by commercial buildings.

Given the commercial sector's importance, as well as the large number of econometric studies estimating residential and transportation energy demand, it is surprising how few econometric studies of commercial energy demand exist. We respond by using the most comprehensive available data on commercial building energy consumption to estimate a

* Newell is a Senior Fellow and Pizer is a Fellow at Resources for the Future, Washington, DC. We thank Martin Heintzelman and Kenny Gillingham for research assistance, and we acknowledge financial support from U.S. Department of Energy Grant DE-FG02-98ER62702.

discrete-continuous choice model of fuel demand, and then use these estimates to simulate carbon mitigation cost curves for the commercial sector.

1.1 Commercial and residential energy demand literature

The microeconomic literature on energy demand in the commercial sector is not very deep. Bohi (1981) and Dahl (1993) provide thorough reviews of the energy demand literature both in general and specifically regarding the commercial sector.[1] The residential energy demand literature is considerably deeper than that for the commercial sector, with modeling taking many different forms ranging from simple ordinary least squares (OLS) to two- and three-stage models using a variety of estimation techniques. For comparability, our discussion concentrates on literature which uses disaggregate cross-sectional data, rather than studies using aggregate or large time-series datasets.

There are three general methodological frameworks employed in modeling residential energy demand. The first is standard OLS or generalized least squares (GLS) estimation of energy use on such factors as energy prices, income, weather, house structure variables, and appliance stock variables (Branch 1993). This simplified approach has its advantages, but it fails to address issues of fuel choice. The second general class of models, the simultaneous equation model, estimates an energy demand equation and an appliance stock equation (and maybe a price equation) simultaneously (Garbacz 1984a, 1984b) to address endogeneity in the appliance stock. These equations are estimated using two-stage least squares.

The final class of models uses a two- or multiple-stage model involving qualitative choice analysis (Baker and Blundell 1991; Bernard et al. 1996; Dubin and McFadden 1984; Parti

[1] The most recent work in this area (Sutherland 1990), uses the 1986 Non-Residential Buildings Energy Consumption Survey (NBECS) to estimate separate ordinary least squares models of electricity, natural gas, and fuel oil use, by four U.S. regions, as a function of fuel prices and various building characteristics. He finds own-price elasticities greater than unity for all but one fuel in one region.

and Parti 1980; Train 1986). This approach is based on a framework in which energy provides utility not directly but indirectly through the use of appliances, implying endogeneity of appliance stocks. Because of this dependency on appliance use, elasticities should be estimated not solely on the basis of one energy equation, but also on the choice of fuels for heating and cooling and on the stock of other appliances. These models differ from the second approach in that the equations are estimated sequentially, not simultaneously. For example, Dubin and McFadden (1984) use a logit estimation of heating fuel choice and variations of OLS and instrumental variables to estimate the energy demand equation. Our approach is closest to the last category of discrete-continuous choice models.

2. Discrete-Continuous Choice Model of Energy Demand

2.1 Model of fuel choice and energy demand

We posit that building owners and developers face a two-stage decision process when determining their energy demand for particular end uses (e.g., heating, water heating, cooking, miscellaneous, and other electricity). In the first stage, they choose *which* fuel or combination of fuels to use for each end use, or perhaps they decide to use no fuel at all. Choice options for cooking, for example, include electricity, natural gas, and joint use of electricity and natural gas. We call this first stage the *fuel choice* decision—it is based on comparing the cost of each alternative, which will depend on fuel prices, equipment costs, and individual building characteristics. Conditional on the first stage, in the second stage, *energy demand* decision, managers determine *how much* of each fuel to use. Similar observed and unobserved variables influence these related discrete and continuous decisions, and care must be taken to correctly account for that relationship.

Our approach, based on Dubin and McFadden (1984), specifies that for a given end use, the cost $C_{i,j}$ of each option i for building j is

$$C_{i,j} = f_i\left(p_{i,j}, r_{i,j}, \mathbf{Z}_j, s_j\right) + \varepsilon_{i,j}, \tag{1}$$

where $p_{i,j}$ is the price of energy, $\mathbf{Z}_j$ is a vector of building-specific characteristics (e.g., building type, size, age, climate), s_j is a given desired level of building energy services, $r_{i,j}$ is the price of equipment, f_i is a function relating these variables to total cost, and $\varepsilon_{i,j}$ is an independently and identically distributed random disturbance. In cases where the building uses only one fuel for each end use, $p_{i,j}$ is simply equal to the price of that fuel. In cases where multiple fuels are used for an end use, $p_{i,j}$ is a vector of prices of the utilized fuels.

The first-stage choice among alternative fuels or fuel combinations explains which fuel option is used based on weighing the relative costs of different options. Assuming the disturbances $\varepsilon_{i,j}$ have extreme value distributions with

$$\Pr\left(\varepsilon_{i,j} < \varepsilon\right) = \exp\left(-\exp(-\varepsilon)\right),$$

the probability that option i has the lowest cost and is therefore chosen is given by

$$\Pr\left(\delta_{i,j} = 1\right) = \frac{\exp\left(f_i\left(p_{i,j}, r_{i,j}, \mathbf{Z}_j, s_j\right)\right)}{\sum_{i'} \exp\left(f_{i'}\left(p_{i,j}, r_{i,j}, \mathbf{Z}_j, s_j\right)\right)}, \tag{2}$$

where $\delta_{i,j}$ is a dummy variable equal to one for the option i selected by building j and zero otherwise, and i' indexes all fuel choice options. A multinomial logit likelihood function can be formed based on Equation (2) and the parameters estimated using maximum likelihood.

The second stage explains the level of fuel use conditional on the fuel option chosen. Using Equation (1) coupled with Shephard's lemma, we can derive the demand for energy, x, by building j with fuel option i as

$$x_{i,j} = \frac{\partial C_{i,j}}{\partial p_{i,j}} = g_i\left(p_{i,j}, r_{i,j}, \mathbf{Z}_j, s_j, \eta_{i,j}\right), \tag{3}$$

where $\eta_{i,j}$ is an error term. An important observation in the literature that we raised above is that the εs and ηs are likely to be correlated. Unobserved characteristics that affect the demand for a particular fuel are likely to affect the costs—and thus the likelihood—of choosing that fuel option. Knowing, for example, that the natural gas option was chosen despite strong predictions to the contrary suggests that perhaps the demand for natural gas will be different than the predicted value. In effect, there is an omitted variable related to fuel choice in the energy demand equation. As a consequence, the conditional expectation of $\eta_{i,j}$ is not zero (i.e., $E[\eta_{i,j} \mid \delta_{i,j} = 1] \neq 0$) but rather a function of the choice probability. This violates the standard OLS assumption that guarantees consistency and is analogous to the sample selection problem discussed by Heckman (1979). Unless a "selectivity correction" is made, the estimated coefficients of the demand model will be biased.

Assuming that $\eta_{i,j}$ is correlated with $\varepsilon_{i,j}$, but not the $\varepsilon_{i'\neq i,j}$ (i.e., the choice errors for the options not chosen), Dubin and McFadden (1984) show that $E[\eta_{i,j} \mid \delta_{i,j} = 1]$ is proportional to $\ln(\Pr(\delta_{i,j} = 1))$. Thus, we can correct for the selection problem by including the log of a consistent estimate of the choice probability (i.e., the predicted probability from the estimated choice equation) as an additional explanatory variable in the energy demand equation, so that the remaining portion of the disturbance in Equation (3) has an expectation of zero.

2.2 Econometric specification

For each of the fuel choice options *i*, we assume the functional form for the fuel choice equation is

$$C_{i,j} = \beta_i + \boldsymbol{\beta}_{i,p} \cdot \ln \mathbf{P}_j + \tfrac{1}{2} \ln \mathbf{P}_j \cdot \boldsymbol{\beta}_{i,pp} \ln \mathbf{P}_j + \beta_{i,y} \ln y_j + \boldsymbol{\beta}_{i,py} \cdot \ln \mathbf{P}_j \ln y_j + \boldsymbol{\beta}_{i,z} \cdot \mathbf{Z}_j + \varepsilon_{ij}, \quad (4)$$

where the β are parameters and $\boldsymbol{\beta}$ are vectors of parameters, $\mathbf{P}_j$ is a vector of fuel prices, and we have broken out y_j, the size of building j, from the other building characteristics, $\mathbf{Z}_j$, to allow for scale biases in fuel choice. Employing this functional form, we estimate the first-stage fuel choice model given by Equation (2) as a maximum likelihood multinomial logit model (see Maddala 1983 for estimation details).

From the estimated choice model, we construct the selectivity correction for each option in each building, $\hat{\lambda}_{i,j}$, which we subsequently include in the energy demand equation(s) for each option. Our estimation procedure for the energy demand model differs only slightly depending on whether the fuel choice involves a single fuel or multiple fuels. For single-fuel options, we specify the energy demand equation for each fuel choice option as

$$\ln x_{i,j} = \gamma_i + \gamma_{i,p} \cdot \ln p_{i,j} + \gamma_{i,y} \ln y_j + \boldsymbol{\gamma}_{i,z} \cdot \mathbf{Z}_j + \gamma_{i,\lambda} \hat{\lambda}_{i,j} + \mu_{i,j}, \quad (5)$$

where the γ are parameters and $\boldsymbol{\gamma}$ is a vector of parameters, $\hat{\lambda}_{i,j}$ is the selectivity correction term from the fuel choice equation, and $\mu_{i,j}$ is an iid random disturbance specific to each option and building.[2] For fuel choice options involving combinations of more than one fuel, we simultaneously estimate two equations of the form (5) and include the prices of both fuels in each demand equation.[3] We therefore estimate a total of five logit choice equations (with a total of 20 fuel use choices and five nonuse options) and 28 demand equations (one demand equation

[2] Note that we do not impose the functional form or cross-equation parameter restrictions that are implied by the structural relationship between Equations (1) and (3)—namely, that the demand function is a derivative of the cost function. In moving toward empirical estimation, one must be careful not to place too much structural restriction on the econometric relationship of the equations for fuel choice, energy demand, and fuel share, in part because the decisions may occur at different times under different conditions. Any such structural relationship is also complicated by the discrete-choice nature of the first-stage choice model, which estimates choice probabilities rather than directly estimating the parameters of the cost function.

[3] Note that unlike share equations or unlogged demand equations, there is no automatic restriction on the cross-price elasticities.

for electricity-only end uses, 12 for heating, six for water heating, five for miscellaneous end uses, and four for cooking). These equations contain a total of several hundred parameter estimates.

Simple aggregate model. For comparison purposes, we also estimate a simplified model that estimates aggregate energy use by fuel and does not treat the fuel choice decision or estimate separately by end use. The four equations (electricity, natural gas, fuel oil, district services) are of the form (5), except that $\hat{\lambda}_{i,j}$ is not included.

3. Data and Estimation

3.1 Commercial Buildings Energy Consumption Survey

The data for this study come from the 1995 Commercial Buildings Energy Consumption Survey (CBECS) conducted by the U.S. Energy Information Administration (1998a). CBECS is a quadrennial (previously triennial) survey of U.S. commercial buildings, which are defined to include buildings used by the service sector—wholesale and retail stores, hotels, and hospitals—as well as other institutions, such as churches and schools. The 1995 survey results include observations for 5,766 buildings sampled to represent about 4.6 million commercial buildings and 59 billion square feet of floor space. The dependent variables are based on KBtus of each of four fuels (electricity, natural gas, fuel oil, and district heat) consumed in each building, by each of five end uses (heating, water heating, cooking, miscellaneous, and other end uses that use only electricity, such as lighting, cooling, office equipment, refrigeration, and ventilation). Table 1 shows fuel expenditure shares by end use and fuel type. Electricity alone accounts for about 81% of expenditures, with natural gas accounting for most of the remainder.

Table 1 provides descriptive statistics on the explanatory variables. The explanatory variables include fuel prices; size (square feet); percentage heated and cooled; temperature (cooling and heating degree days); hours open weekly; whether building is owner occupied; age (in 20-year intervals); location (North, South, Midwest, West); whether building is in part of a multibuilding facility; whether building has a central physical plant; and the type of building (restaurant, warehouse, hospital, public, lodging, retail, or other). All explanatory variables were included in all equations in the form given by Equations (4) and (5), with all nonpercentage continuous variables first being logged. Percentage lighted was also included in equations for the electricity-only end use. To capture demand for heating (cooling) services, we interact percentage heated (cooled) by heating (cooling) degree days rather than adding these variables individually. The logit fuel choice equations also include dummy variables for whether the building is in a metropolitan statistical area (MSA; i.e., is urban) or is government owned. Other potentially relevant variables available in the CBECS data were excluded either because there were many missing observations or because the variable was estimated to have negligible effect.

3.2 Estimation and simulation

Data imputation. Perhaps the most vexing problem that we face is a rather extensive amount of missing data—in particular, price information for the fuel choice equations. Prices are derived from the ratio of fuel expenditures to fuel quantity. When a fuel is not used by a particular facility, for any end use, the price is not observed. We are also missing data on the percentage of building cooled, percentage of building heated, percentage of building lighted, whether the building is owner-occupied, number of hours the building is open weekly, and cooling degree days. The number of observations across different variables is shown in Table 1, illustrating the missing data problem.

To address missing data, we use a multiple imputation approach (Rubin 1987). We estimate a relationship between each missing variable and other observed variables using a simple linear model based on complete observations. We then impute values for missing variables based on the predictions of the linear model *plus* a random disturbance based on the estimated error in the linear model.[4] We repeat this imputation process five times, with different random disturbances for each imputation. Our analysis is performed separately on all five imputed data sets, with overall parameter estimates based on the average estimate across the five imputations and overall parameter variance estimates based on the sum of between and within variation measured in the five parameter estimates.[5]

Multinomial logit choice options. Once we have a complete (imputed) data set, we can implement the two-step discrete choice, linear demand model outlined in the text to estimate model parameters and simulate demand response. We do this for each of five end uses: heating, water heating, cooking, miscellaneous, and other electrical (miscellaneous includes nonexclusively electrical end uses). Several practical issues arise that require some attention as we estimate the model. The first is the number of choices in the discrete-choice step. There are four main fuels that account for the vast majority of all reported commercial energy use: electricity, natural gas, fuel oil, and district heat. Because these fuels substitute for one another, their use needs to be considered jointly, giving rise to a total of 16 combinations.

Estimating a choice model with 16 options, especially when some options are particularly uncommon, is impractical. In our work, we identify eight combinations that occur frequently

---

[4] When we use the imputed price of district heat services to estimate the fuel choice model (2), we inflate the imputed price by a factor of 10 for a randomly chosen 90% of the imputed observations. This helps us account for the fact that the use of district heat in most locations that do not already use district would likely involve very high fixed installation and access costs.

[5] That is, if our estimates of a parameter θ equal θ_1, θ_2, θ_3, θ_4, and θ_5, with variance estimates of σ_1^2, σ_2^2, σ_3^2, σ_4^2, and σ_5^2, across the five imputed data sets, the overall parameter estimate would be $\bar{\theta} = (\theta_1 + \theta_2 + \theta_3 + \theta_4 + \theta_5)/5$ and the overall parameter variance estimate would be $(\sigma_1^2 + \sigma_2^2 + \sigma_3^2 + \sigma_4^2 + \sigma_5^2)/25 + (\theta_1^2 + \theta_2^2 + \theta_3^2 + \theta_4^2 + \theta_5^2)/25 - \bar{\theta}^2/5$.

enough to warrant inclusion, and for most end uses we consider a smaller subset. Table 3 shows the pattern of fuel use combinations by end use that we consider. Observations that do not fit into our groupings are not used to estimate the model but are used to predict fuel use and demand response to price changes.[6] After estimating the fuel choice model, we use it to predict the probability of each possible fuel combination for each observation, including those observations that were ignored in the estimation based on their use of an unusual fuel combinations.[7] We do this for both the benchmark prices as well as simulated price changes associated with varying levels of a carbon price.

Sampling weights in aggregate prediction. After estimating the fuel demand equations for each fuel choice combination based on the observations that actually chose that fuel choice, we predict the fuel demand for *all* observations for all fuel choice combinations for both benchmark prices and simulated price changes. This raises an issue as we shift from predictions of *log* energy demand (the form in our model in (5)) to *level* energy demand. The issue is that mean-zero predictions of log energy demand do not translate into mean-zero predictions of level energy demand because exponentiation is a nonlinear function and nonlinear functions generally do not preserve expectations. Mathematically, we have

$$E\left[y\right] = E\left[\hat{y} + e\right] = \hat{y}$$

where y is the true log energy demand, $\hat{y}$ is the predicted value in logs, and e is the prediction error. But,

$$E\left[\exp y\right] = E\left[\exp\left(\hat{y} + e\right)\right] = \exp\hat{y} \cdot E\left[\exp e\right] \neq \exp\hat{y}$$

[6] This amounts to only 1.75% of observations.

[7] For example, an observation that fuel oil was used for cooking would be ignored in the estimation, but we would still go back and predict the probability that this observation alternatively used electricity, natural gas, or both (the three fuel combinations included in the estimation).

because $E[\exp e] \neq 1$. Assuming a normal distribution for *e*, one can compute $E[\exp e] = \exp \hat{\sigma}^2/2$ where $\hat{\sigma}^2 = \frac{1}{n}\sum e_i^2$ is the estimated variance of *e* based on observed residuals e_i (see Goldberger 1968).

We actually take a slightly different approach incorporating sampling weights so that our aggregate prediction at observed prices matches actual aggregate energy use when the individual predictions are aggregated using those weights. The sampling weights, provided with the data set, indicate how many buildings each observation represents in the population. That is, we want to construct an estimate of $E[\exp e]$ such that

$$\sum_i w_i \cdot E[\exp e] \exp \hat{y}_i = \sum_i w_i \exp y_i ,$$

where again y_i are the true log energy demands (for the observed data), $\hat{y}_i$ are the predicted log energy demands, and w_i are the sampling weights. The following consistent estimator of $E[\exp e]$ satisfies the above condition:

$$\widehat{E[\exp e]} = \frac{\sum_i w_i \exp \hat{y}_i \cdot \exp e_i}{\sum_i w_i \exp \hat{y}_i} ,$$

which in effect weights using both the sample weights and the predicted level of energy use.[8] Once we calculate our estimate of $E[\exp e]$, we scale up all of our predictions ($\exp \hat{y}$) of levelized fuel demand by this value for each end use, fuel combination, and fuel. We compute estimates of expenditure by multiplying by price. The advantage of this approach is that we maintain the benchmark prediction of aggregate energy demand using the estimated sample when the sample is aggregated.

[8] Note that any weighted average of observed values of e_i would provide a consistent estimate of $E[\exp(e)]$.

Once we have computed the levelized (unlogged) fuel use and expenditure predictions for each observation, end use, fuel combination, and fuel, we aggregate by computing

$$\sum_{i,j,k} w_i \hat{p}_{i,j,k} \exp \hat{y}_{i,j,k,l} E\left[\exp e_{i,j,k,l} \middle| j,k,l\right] ,$$

where w_i is the sampling weight for observation i, $\hat{p}_{i,j,k}$ is the probability of the kth fuel combination, for the jth end use, and observation i, $\hat{y}_{i,j,k,l}$ is predicted log fuel demand for observation i, end use j, fuel combination k, and fuel l, and $e_{i,j,k,l}$ is similarly the error in the prediction of log fuel demand for observation i, end use j, fuel combination k, and fuel l. This delivers estimates of aggregate fuel demand for fuel l across all commercial buildings for both the benchmark and the simulated price changes. Note that alternatively, we can also compute

$$\sum_{i,j,k} w_i 1\left(k = Z_{i,j}\right) \exp \hat{y}_{i,j,k,l} E\left[\exp e_{i,j,k,l} \middle| j,k,l\right],$$

where k indicates a particular fuel choice combination, and $Z_{i,j}$ is the observed fuel combination choice for observation i and end use j. This represents the predicted energy demand holding the choice of fuel combination fixed. Based on our weighting to compute $E\left[\exp e\right]$, we would exactly match the benchmark aggregate fuel use estimates for the sample *except* we do not predict energy demand for certain fuel choice combinations that are observed in the sample.[9]

Carbon price simulations. We simulate the effect of carbon taxes (or equivalent permit system prices) on commercial energy demand and carbon emissions using parameter estimates from the fuel choice and demand models described above. We assess carbon taxes ranging from \$10 per metric ton of carbon (\$/tC) to \$150/tC by increasing all fuel prices simultaneously based on the carbon content of each fuel,[10] and assuming these increases are passed on to building

[9] See footnote 6.

[10] The emissions factors used for converting Btus of energy consumption to carbon emissions were as follows, in million metric tons of carbon per quadrillion Btus: 49.2 for electricity (U.S. DOE 2001, 128, 151); 14.47 for natural gas (U.S. EIA 2002, B-1); 19.95 for fuel oil (U.S. EIA 2002, B-1); and 38.36 for district heat services. The district

occupants (e.g., the price of electricity and district heat services rises based on average carbon content). We do not consider the possible response of electricity or district heat generators to adjust their carbon content (via fuel switching), so our results are best thought of as partial equilibrium in nature.[11] For comparison with other studies and for other modeling efforts, we also compute demand elasticities by fuel type based on demand responses to simulated price increases, where fuel prices are altered one at a time. We report average elasticities for price increases resulting from a $50/tC carbon tax.

4. Estimation and Simulation Results

4.1 Estimation results

As described earlier, our disaggregated approach involves estimation of five logit fuel choice equations (with a total of 20 fuel-using and five nonuse choices) and 28 fuel demand equations, containing several hundred parameter estimates. We focus our attention here on parameter estimates for energy prices because these have the most direct effect on the resultant cost estimates for carbon mitigation via carbon tax–induced energy price increases. The full set of estimation results is shown in Tables A.1–A.10 in the Appendix.

Energy cost minimization implies that an increase in the price of a particular fuel relative to other fuels should decrease the probability of that fuel's being chosen for a particular end use. The logit fuel choice parameter estimates support this relationship, with all but one of the 28 linear fuel price coefficients being either significantly negative or statistically insignificant (17 were significantly negative, one was significantly positive, and 10 were insignificant). Recall

heat emissions factor was based on the portion of primary fuels used in district heat combustion (50% natural gas, 18% electricity, 27% coal, and 5% fuel oil; U.S. EIA 1993, 29), using an emissions factor for coal of 26 (U.S. EIA 2002, B-1) and a heat input-output ratio of 1.59 (U.S. EIA 1993).

[11] Our estimates also hold constant the level of building services, consistent with EIA general equilibrium analyses (U.S. EIA 1998b), which suggest that such output effects in the commercial building sector would be negligible.

that the predicted probability of each fuel option ($\hat{\lambda}_{i,j}$) also enters into each of the demand equations to control for correlation between the fuel choice and fuel utilization decisions. The estimated effects of these selectivity corrections were statistically significant at the 5% level in four of the 28 demand equations. All of the significant estimates were negative, consistent with the idea that lower demand (and cost) for a particular choice is associated with a higher probability of making that choice.

Turning to the fuel demand equations, downward-sloping demand functions imply that own-price demand elasticities should also be negative. As shown in Table 3, own-price elasticity estimates for all 28 demand equations were negative, with all but two of these estimates being significant at the 5% level. The own-price elasticities ranged from –0.31 for district services for heating to –4.21 for fuel oil heating when used in combination with electricity. Over three-quarters of the own-price elasticity estimates were found to be greater than unit elasticity, implying that energy price increases should result in not only commercial fuel demand decreases but also net decreases in commercial energy expenditures. Cross-price elasticities in demand equations for fuel options with two fuels were typically imprecisely estimated; only three of 16 cross-price elasticities were statistically significant.[12]

Finally, for the main end uses where fuels were sometimes used in combination (i.e., heating and water heating), the estimates tended to be more elastic when the fuels were used in combination with another fuel than when only one fuel was used. For example, the elasticity for fuel oil for heating was –2.07 when fuel oil was used by itself, but when it was used in

[12] The significant elasticities—the electricity price in the natural gas demand equation for the heating, water heating, and "miscellaneous" end uses when both electricity and natural gas are used—were negative, suggesting complementarity.

combination with electricity for heating, the elasticity rose to –2.83. This is consistent with the immediate presence of more fuel options' leading to greater flexibility in demand adjustment.[13]

Table 4 shows results for aggregate own-price demand elasticities for each type of fuel, based on simulation of fuel price increases and demand responses of the end use–specific models. The table shows elasticity estimates allowing the fuel choice to be both fixed and variable, based on whether adjustment using the logit fuel choice model was included. As expected, the own-price elasticities treating fuel choice as variable were all higher than those treating fuel choice as fixed, with the combined elasticity ranging from –0.88 for district services to –2.95 for fuel oil. Aggregate cross-price elasticities for each fuel were all found to be nonnegative but fairly low, as shown in Table 5. The elasticity of natural gas with respect to electricity price changes was 0.21, while the elasticity of fuel oil with respect to natural gas price changes was 0.20. All other aggregate cross-price elasticities were between zero and 0.14.

Those elasticities are in the range of other econometrically estimated long-run own-price and cross-price elasticity estimates for the commercial sector, which tend to be around –1 for electricity and natural gas, and more elastic for fuel oil (Dahl 1993; Wade 1999; Bohi 1981); we know of no other elasticity estimates for district services. These estimates are distinctly more elastic than the implicit long-run elasticity of the National Energy Modeling System (NEMS) commercial demand module, however, which has implied own-price elasticities of –0.45 for electricity, –0.40 for natural gas, and –0.39 for fuel oil (Wade 2003).[14]

Interestingly, two of the three nonnegligible NEMS cross-price elasticities are higher than our aggregate cross-price elasticities (shown in Table 5). NEMS commercial demand has

[13] The one exception is the elasticity of demand for electricity, which is lower when electricity is used in conjunction with district heat services, a relatively inelastic energy source.

[14] NEMS uses a short-run price elasticity of demand of –0.25 for all commercial end uses except refrigeration, which uses –0.10, and office equipment and miscellaneous end uses, which employ a –0.05 elasticity. The long-run

implicit cross-price elasticities of 0.86 for natural gas in response to electricity price increases (we estimated the cross-price elasticity at 0.09), and 0.75 for fuel oil in response to natural gas price increases (we estimated 0.20). Thus, NEMS assumptions imply that commercial demand for a particular fuel will fall less and demand for most substitute fuels will rise more in response to a fuel price increase than suggested by our estimates. Both these differences in price responsiveness between our estimates and NEMS are reflected in our estimates (below) of the cost of carbon mitigation from the commercial sector, which are much higher in analyses using NEMS.

Comparison with simple demand model. For comparison purposes, we also estimate a simple model of fuel demand, which aggregates fuel usage by fuel type (i.e., it does not separately treat end uses) and does not address the discrete fuel choice decision. We estimate this simple model using both the full imputed data set (averaging estimates over imputations as described above), and the data set without imputed values (where demand for a particular fuel is estimated using only those building observations that actually consumed that fuel).

The results for the simple fuel demand equations are shown in the rightmost columns of Table 3. Elasticities exhibit the same pattern across fuels as the aggregate elasticities based on the disaggregated discrete-continuous choice modeling framework (e.g., district heat < electricity < natural gas < fuel oil), but there are some significant differences in magnitude. The estimated elasticities for natural gas and fuel oil are substantially higher in the simple aggegate model, while the district services elasticity is much lower. The imputation approach does not substantially affect the estimates in the simple aggegate model.

price elasticity for commercial energy demand is also a function of altered equipment choices in response to changing fuel prices.

4.2 Carbon price simulation results

Figure 1 and Table 6 present our estimates of the effect of increases in the price of carbon up to \$150/tC. The figure shows that the econometric estimates imply a slightly convex marginal carbon cost curve, with a \$20/tC and \$100/tC carbon tax being associated with 8% and 28% carbon reductions, respectively. The cost of equivalent percentage reductions based on the NEMS commercial demand module is five to six times higher, as suggested earlier by the much lower responsiveness to fuel price changes implicit in that model. Note that these estimates all hold the carbon intensity of electricity constant, even though it would be expected to fall if carbon policy covered the electricity sector. Thus, a full general equilibrium assessment of the costs of reducing carbon attributable to the commercial sector would show greater emissions reductions and lower costs for a given carbon tax than suggested here.[15]

Table 6 provides some additional detail on the carbon policy simulation results. We estimate that the long-run annual total cost[16] of carbon reductions from commercial buildings via increased carbon prices ranges from \$132 million for an 8% reduction in emissions (\$20/tC carbon tax) to about \$2.2 billion for a 28% reduction (\$100/tC carbon tax). These carbon reductions result from decreases in fuel consumption, which range from 6% to 23% for electricity, 16% to 44% for natural gas, and 22% to 67% for fuel oil for a carbon price of \$20–\$100/tC. Individual end uses exhibit a somewhat smaller range of carbon reductions (6%–39%) over the same carbon price range.

---

[15] We compute the percentage carbon reduction from the NEMS commercial sector based on fuel demand responses from the U.S. Energy Information Administration (1998b) multiplied by the carbon emissions factors given in footnote 10. The results for NEMS in 2010 and 2020 reflect snapshots as of those years, with the 2020 snapshot reflecting lower costs due to greater time for adjustment to price changes.

[16] The total cost is given by the area under the marginal cost curve, which we estimate by integrating a quadratic formulation of the marginal cost curve, which fits the data extremely well.

Electricity use, which accounts for 70% of baseline carbon emissions from commercial buildings, also constitutes the majority (50%–57%) of reductions due to carbon price increases. Natural gas accounts for 15% of baseline carbon emissions and a proportionately higher 30%–24% of reductions over the carbon price range \$20–\$100/tC. Fuel oil accounts for only 3% of baseline emissions but a proportionately much higher 7%–6% of reductions. This pattern of reductions across fuels is consistent with their relative elasticities. Similarly, the last few rows of Table 6 show that individual end uses contribute to carbon reductions in rough proportion to their contribution to baseline emissions, although heating emissions fall proportionately more than the electricity-only end use. This is again consistent with the relative elasticity of the underlying fuels predominant in each end use.

5. Conclusion

We demonstrate that it is possible to formulate a highly disaggregated, end use– and fuel-specific model of energy demand that addresses both the fuel choice and fuel use decisions, to estimate these models using publicly available data, and to construct a marginal cost curve for carbon reductions in the commercial sector based on these estimates. Such modeling efforts help bridge a significant methodological and empirical gap between often highly aggregated energy-economy models and underlying microeconomic behavior. We apply this modeling approach to the seldom-studied commercial sector, finding quite elastic demand response to fuel price increases.

Although our estimates are in the range of other econometric studies, they are much more elastic than the implied price-responsiveness of the NEMS commercial models, NEMS being the official modeling tool used by the Department of Energy’s Energy Information Administration for forecasting future energy prices and quantities and evaluating the cost of policies to reduce energy use and carbon dioxide emissions. Our simulations of the effectiveness of carbon price increases (due to a carbon tax or carbon permit system) suggest that \$20–\$100/tC, carbon price would induce carbon reductions of 8%–28%, respectively, from the commercial sector (holding constant the carbon intensity of electricity). Because of its lack of price responsiveness, NEMS suggests carbon prices that are about five to six times higher for equivalent commercial-sector reductions (again, holding constant the carbon intensity of electricity).

Table 1: Descriptive Statistics

Variable	*Number of observations*	*Mean*	*Standard deviation*
Electricity price (¢/KBtu)	5,609	2.60	1.30
Natural gas price (¢/KBtu)	3,708	0.59	0.32
Fuel oil price (¢/KBtu)	878	0.57	0.14
District services price (¢/KBtu)	583	0.72	0.30
Square footage	5,766	127,606	265,892
Percentage lighted	5,575	0.89	0.23
Percentage heated	5,369	0.89	0.25
Percentage cooled	4,947	0.74	0.34
Cooling degree days	5,755	1,312	879
Heating degree days	5,766	4,357	2,249
Hours open weekly	5,646	79	48
Owner occupied (0/1)	5,653	0.79	0.41
Government owned (0/1)	5,766	0.21	0.40
Multibuilding facility (0/1)	5,766	0.43	0.50
Central physical plant (0/1)	5,766	0.13	0.34
Age≤20 (0/1)	5,766	0.41	0.49
20<Age≤40 (0/1)	5,766	0.34	0.47
40<Age≤60 (0/1)	5,766	0.12	0.33
60<Age≤80 (0/1)	5,766	0.07	0.25
80<Age (0/1)	5,766	0.05	0.22
Northern U.S. (0/1)	5,766	0.17	0.37
Southern U.S. (0/1)	5,766	0.36	0.48
Midwestern U.S. (0/1)	5,766	0.24	0.43
Western U.S. (0/1)	5,766	0.24	0.42
Located in MSA (0/1)	5,766	0.80	0.40
Building type			
Restaurant (0/1)	5,766	0.02	0.15
Warehouse (0/1)	5,766	0.14	0.35
Hospital (0/1)	5,766	0.04	0.19
Public (0/1)	5,766	0.23	0.42
Lodging (0/1)	5,766	0.07	0.25
Retail (0/1)	5,766	0.20	0.40
Other (0/1)	5,766	0.31	0.46

Table 2. Expenditure Shares by End Use and Fuel

End use	Fuel				Total
	Electricity	Natural gas	Fuel oil	District services	
Electricity only	1.00				0.69
Lighting	1.00				0.37
Cooling	0.99				0.11
Office equipment	1.00				0.10
Refrigeration	1.00				0.06
Ventilation	1.00				0.05
Heating	0.24	0.50	0.08	0.19	0.15
Miscellaneous	0.88	0.10	0.02		0.07
Water heating	0.22	0.50	0.05	0.23	0.07
Cooking	0.29	0.71			0.02
Total	0.81	0.13	0.02	0.04	1.00

Table 3. Own-Price Elasticities for Demand Equations by Fuel Option and End Use (Given choice of fuel option)

Fuel options	End uses				
	Electricity only	Heating	Water heating	Cooking	Miscellaneous
Electricity	–1.14**	–1.37**	–1.19**	–1.07**	–1.03**
Natural gas		–1.79**	–1.24**	–1.51**	
Fuel oil		–2.07**	–1.46*		
District services		–0.31*	–0.41**		
Electricity and natural gas					
Electricity		–1.69**	–1.36**	–0.97**	–0.82**
Natural gas		–2.12**	–1.60**	–1.29**	–1.89**
Electricity and fuel oil					
Electricity		–1.44**			–1.06**
Fuel oil		–4.21**			–2.54**
Electricity and district services					
Electricity		–0.91**			
District services		–0.90**			
Natural gas and fuel oil					
Natural gas		–2.07**			
Fuel oil		–2.83**			

Note: "Electricity only" includes lighting, cooling, office equipment, refrigeration, and ventilation. Asterisks denote statistical significance: ** = 95%, * = 90%.

Table 4. Own-Price Elasticities for Aggregate Fuel Demand

Fuel	Detailed model, then aggregated		Simple aggregate model	
	Fuel choice fixed	Fuel choice variable	Without imputations	With imputations
Electricity	–1.12	–1.14	–1.07	–1.29
			(0.04)	(0.04)
Natural gas	–1.39	–1.60	–2.25	–2.28
			(0.06)	(0.06)
Fuel oil	–2.00	–2.95	–4.45	–4.33
			(0.24)	(0.22)
District services	–0.44	–0.88	–0.38	–0.42
			(0.08)	(0.08)

Note: Standard errors are shown in parentheses for the simple aggregate model.

Table 5. Own- and Cross-Price Elasticities for Aggregate Fuel Demand (variable fuel choice using detailed model)

	with respect to change in price of…		
	Electricity and district services	Natural gas	Fuel oil
Elasticity of…			
Electricity	–1.14	0.01	0.00
Natural gas	0.09	–1.60	0.07
Fuel oil	0.21	0.20	–2.95
District services	–0.88	0.14	0.09

Table 6. Cost and Distribution of Carbon and Fuel Demand Reductions

	Carbon price ($/tC)				
	$20	$40	$60	$80	$100
Total cost ($million)	132	461	939	1,531	2,211
% reduction *in* total carbon emissions	8	14	19	24	28
% reduction *in* carbon emissions and fuel demand, by fuel					
Electricity	6	11	15	19	23
Natural gas	16	25	33	39	44
Fuel oil	22	38	50	59	67
District services	8	15	21	26	31
% reduction *in* carbon emissions, by end use					
Electricity only	6	11	15	19	23
Heating	13	21	28	34	39
Water heating	10	18	25	31	36
Cooking	9	17	23	29	34
Miscellaneous	8	13	18	22	26
% *of* carbon reductions from each fuel					
Electricity (70% of baseline emissions)	50	53	55	56	57
Natural gas (15% of baseline emissions)	30	27	26	24	24
Fuel oil (3% of baseline emissions)	7	7	7	6	6
District services (12% of baseline emissions)	13	13	13	13	13
% *of* carbon reductions from each end use					
Electricity only (60% of baseline emissions)	43	45	47	48	49
Heating (21% of baseline emissions)	35	33	31	30	30
Water heating (10% of baseline emissions)	13	13	13	13	13
Cooking (2% of baseline emissions)	2	3	3	3	3
Miscellaneous (7% of baseline emissions)	6	6	6	6	6

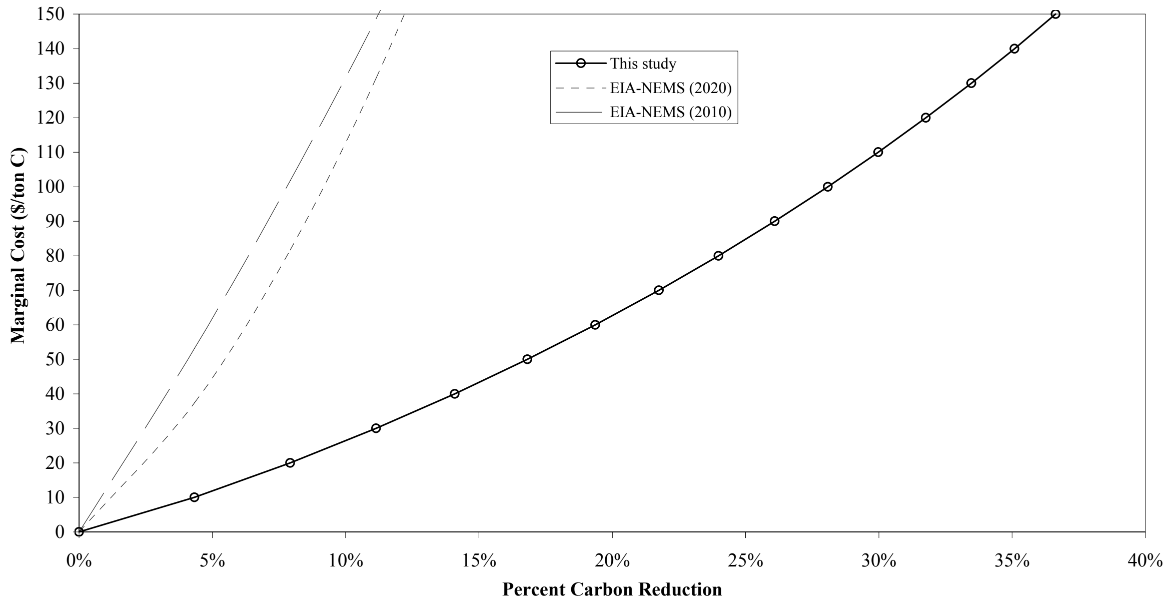

Figure 1. Carbon Cost Curve for the Commercial Sector

Note: Figure shows percentage carbon reduction from the NEMS commercial sector based on fuel demand responses from EIA (1998b) multiplied by the carbon emissions factors given in the data section.

References

Baker, Paul, and Richard Blundell. 1991. The Microeconometric Approach to Modeling Energy Demand: Some Results for UK Households. *Oxford Review of Economic Policy* 7(2).

Bernard, Jean-Thomas, Denis Bolduc, and Donald Belanger. 1996. Quebec Residential Electricity Demand: A Microeconometric Approach. *Canadian Journal of Economics* 29(1).

Bohi, Douglas R. 1981. *Analyzing Demand Behavior: A Study of Energy Elasticities*. Baltimore: Resources for the Future and Johns Hopkins University Press.

Branch, E. Raphael. 1993. Short Run Income Elasticity of Demand for Residential Electricity Using Consumer Expenditure Survey Data. *The Energy Journal* 14(4).

Dahl, C. 1993. A Survey of Energy Demand Elasticities in Support of the Development of the NEMS., Washington DC: U.S. Department of Energy.

Dubin, Jeffrey A., and Daniel L. McFadden. 1984. An Econometric Analysis of Residential Electric Appliance Holdings and Consumption. *Econometrica* 52(2).

Garbacz, Christopher. 1984a. Residential Electricity Demand: A Suggested Appliance Stock Equation. *The Energy Journal* 5(2).

———. 1984b. A National Micro-Data Based Model of Residential Electricity Demand: New Evidence on Seasonal Variation. *Southern Economic Journal* 51(1).

Goldberger, A.S. 1968. The Interpretation and Estimation of Cobb-Douglas Functions *Econometrica* 35(3-4): 464–72.

Heckman, J. 1979. Sample Selection Bias as a Specification Error. *Econometrica* 47(1): 153–61.

Maddala, G.S. 1983. *Limited Dependent and Qualitative Variables in Econometrics*. Cambridge: Cambridge University Press.

Parti, M., and C. Parti. 1980. The Total and Appliance-Specific Conditional Demand for Electricity in the Household Sector. *Bell Journal of Economics* 11(1): 309–21.

Rubin, D.B. 1987. *Multiple Imputation for Nonresponse in Surveys*. New York: John Wiley and Sons.

Sutherland, Ronald J. 1990. The Demand for Energy in Commercial Buildings. *Energy Systems and Policy* 14: 237–56.

Train, Kenneth E. 1986. *Qualitative Choice Analysis*. Cambridge, MA: MIT Press.

U.S. Department of Energy (DOE). 2001. Annual Energy Outlook 2001. DOE/EIA-0383.

Washington, DC: Department of Energy.

U.S. Energy Information Administration (EIA). 1993. Assessment of Energy Use in Multibuilding Facilities. DOE/EIA-0555(93)/1. Washington, DC: Energy Information Administration.

———. 1998a. A Look at Commercial Buildings in 1995: Characteristics, Energy Consumption, and Energy Expenditures. DOE/EIA-0625(95). Washington, DC: Energy Information Administration.

———. 1998b. Impacts of the Kyoto Protocol on U.S. Energy Markets and Economic Activity., Washington, DC: Energy Information Administration.

———. 2002. Emissions of Greenhouse Gases in the United States 2001. DOE/EIA-0573(2001). Washington, DC: Energy Information Administration.

Wade, Steven. 1999. Price Responsiveness in the NEMS Buildings Sector Models. In *Issues in Midterm Analysis and Forecasting 1999,* pp. 55-63. DOE/EIA-0607(99). Washington, DC: Energy Information Administration.

———. 2003. Personal communication. Washington, DC: Energy Information Administration.