The View from the Top: Top-Down Estimation of Program Savings Using Utility-Level Data in Massachusetts

Chris Russell, Ph.D., NMR Group, Inc., Bryan, TX
Ferit Ucar, Ph.D., Environmental Defense Fund, Princeton, NJ
Lawrence Masland, Massachusetts DOER, Boston, MA
Noel Stevens, DNV GL, Burlington, MA
Monica Cohen, Columbia Gas of Massachusetts, Westborough, MA

ABSTRACT

There has been a growing interest in top-down approaches to measuring the impacts of energy efficiency programs because these approaches provide a relatively inexpensive option for estimating the overall net savings from a portfolio of programs in a given geographic area. The Massachusetts Program Administrators (PAs) funded a study to explore the potential for using top-down techniques as an additional tool to complement bottom-up approaches to measuring impacts from energy efficiency programs because top-down techniques have the potential to capture interactive effects between programs and market effects.

In this paper, we developed two top-down models—one residential and one commercial—that used long-term aggregate data from the PAs and municipal utilities in Massachusetts to estimate the net electric program savings. The data for the municipal utilities were included in the model in order to provide a measure of the baseline level of program activity because, historically, the municipal utilities in Massachusetts have had no or significantly lower levels of energy program activity compared to the PAs. While the top-down net saving estimates, in general, were comparable to the traditional bottom-up estimates, the results were somewhat sensitive to the model specification. The findings from the preferred model indicated that a one-dollar increase in energy efficiency expenditures this year would decrease electricity consumption by 4.3±2.7 kWh per year in the residential sector and 3.3±2.5 kWh per year in the C&I sector. The top-down annual net saving estimates were 187%±117% of the bottom-up annual net saving estimates in the residential sector and 101%±78% of that in the C&I sector between 2003 and 2012.

Introduction

The Massachusetts PAs funded a study to explore the potential for top-down techniques to complement bottom-up approaches to estimating impacts from energy efficiency programs as well as exploring whether top-down modeling should play a role in net energy impact evaluation. Top-down modeling is an econometric approach to estimating program impacts that employs aggregate cross-sectional and time series data to model aggregate energy consumption as a function of exogenous variables including program activity, price, and other economic factors. The goal of this modeling is to isolate the effect of program activity from policy and economic factors, as well as other naturally occurring changes.

One of the primary motivations for the PA-municipal top-down approach was an important initial analysis conducted by Lawrence Masland of the Massachusetts Department of Energy Resources. Mr. Masland examined trends in per-customer residential energy consumption for PAs and municipal utility customers for 1990 through 2011 using the data from the Energy Information Administration's (EIA) Annual Electric Power Industry Reports (EIA-861 data files). His analysis showed that the average annual residential electricity consumption per customer for both the PAs and municipal utilities increased from 1990 to 2011, but the rate of increase was significantly higher for the municipal utilities.

Another primary motivation for this model was to establish the counterfactual (no-program) scenario. Understanding the true extent of a program's impacts requires information regarding the level of consumption absent any programmatic activity. This PA-municipal approach extended the timeframe long enough to include a period with no programmatic activity, at least for municipal utilities.

In this paper, we developed a version of a macro-consumption model using aggregate electricity consumption data for PAs (at the PA level) and municipal utilities in Massachusetts. We modeled these data as a function of exogenous variables including program activity, price, weather, and other demographic and economic factors affecting consumption. We ran separate models for the residential and commercial and industrial (C&I) sectors. By controlling for other factors that could cause the diverging trends in electricity consumption between the PAs and municipal utilities, this top-down model sought to isolate the effect of energy-efficiency programs on consumption. The substantial differences in energy-efficiency program expenditures across the PAs and the municipal utilities in a given year and within PA and municipal utilities over time provided the identifying variation for the model.

An important part of the study was to identify the most appropriate data sources and gather data on energy consumption, energy program activity, and other factors. The main explanatory variable of interest was electric program activity. We attempted to collect detailed data on electric program activity from the PAs and municipal utilities for 1990 through 2012. The only piece of electric program data that was consistently available across all PAs, municipal utilities, and years was the total residential and C&I electric program expenditures. As a result, we used the annual total program expenditures as a proxy for the annual energy program activity.

Residential Model

This section provides details on residential model specification and presents the model results. First, we discuss the data collected for this project. Next, we present the model specification. We then present and discuss the residential model results. Finally, we present a comparison of top-down and bottom-up estimates of residential net savings.

Data Collection

We collected time-series data on residential electricity consumption and factors that could affect consumption for all Massachusetts PAs/utilities and towns from 1990 to 2012, including the following data elements:

- *Electricity Consumption and Price Data* – We collected data on the total residential electricity sales, revenue, and customers in Massachusetts from the EIA's 861 files for 1990 to 2012 for each PA and municipal utility. We computed the annual energy consumption per customer and average price per kWh using these data.
- *Energy Efficiency Programmatic Activity* – We assessed the quality of demand-side management program data reported to EIA on the EIA-861 form. The assessment revealed that data were missing and/or inconsistent for some PAs and municipal utilities for some years. Because it was crucial to gather accurate information for the main explanatory variable of interest for the model in order to produce reliable estimates, we made a substantial effort to collect the energy efficiency program expenditures data by sector and year from the PAs, the municipal utilities, and their association.[1]
- *Weather Data* – We gathered daily temperature data for all weather stations in Massachusetts from the National Oceanic and Atmospheric Administration (NOAA) from 1990 through 2012. We first computed the annual heating degree days (HDDs) and cooling degree days (CDDs)

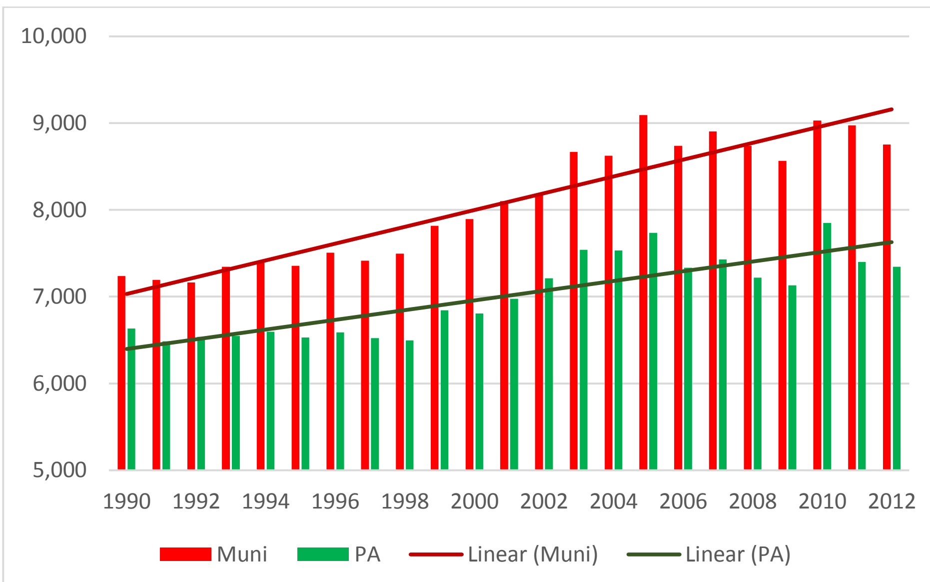
[1] Despite these efforts, we could not get program data for 12 municipal utilities in the state. These utilities were excluded from the analysis. We would especially like to thank Kim Boas of the Massachusetts Municipal Wholesale Electric Company (MMWEC) for providing data for its members and the Massachusetts PAs for providing the data for their organizations.

for each station. Next, we matched each town to the nearest weather station. Finally, we computed a weighted average of annual HDDs and CDDs for each PA/utility service area using the number of housing units in each town as the weight.

- *Economic and Demographic Data* – We gathered town-level economic and demographic data from the following sources:
  - US Census American Community Survey (ACS) – Contains annual residential socioeconomic data at the census block level of granularity—the smallest geographic unit used by the US Census Bureau—since 2005.
  - US Decennial Census – Contains residential socioeconomic data at the census block level of granularity. Conducted in 1990, 2000, and 2010.[2]
  - US Census Building Permits Survey – Contains annual construction statistics by permit-issuing place (usually the township) on new privately owned residential housing units authorized by building permits.
  - Bureau of Labor Statistics – Contains annual labor force, employment, and unemployment counts at the town level of granularity.

Model Specification

As shown in Figure 1, average annual residential electricity consumption per customer for both PAs and municipal utilities increased from 1990 to 2012, but the rate of increase was significantly higher for the municipal utilities than for the PAs.

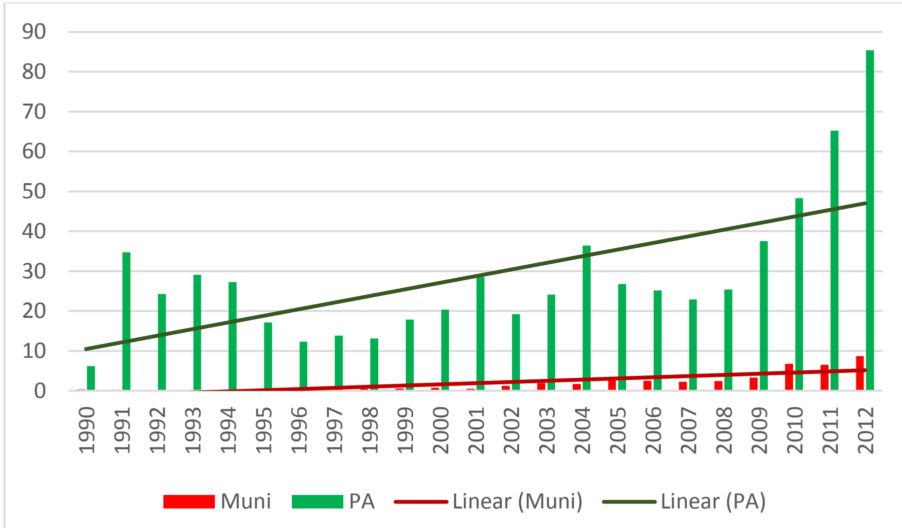


Source: Annual Electric Power Industry Report (EIA-861 data file)

Figure 1. Residential Electricity Consumption per Customer (in kWh), Massachusetts

While most, if not all, municipal utilities had residential energy efficiency programs during the same period, the municipal utilities were slow to embrace the funding of energy efficiency programs, and funding levels were significantly below those of the PAs, as shown in Figure 2.

[2] Decennial Census data (1990 and 2000) were used for the period before the annual ACS data were available. In order to make the decennial data fit into a data set with yearly time points, we estimated the difference between the two points (1990 and 2000, for example) and evenly distributed the difference annually between the two data collection points, thereby forcing the decennial data to vary from year to year.



Source: Massachusetts Program Administrators, municipal utilities, and the Massachusetts Municipal Wholesale Electric Company (MMWEC)

Figure 2. Residential Electric Program Expenditures per Customer (in $), Massachusetts

While the lower levels of increase in consumption in PA territories could be due to greater programmatic activity, accurately estimating programmatic impacts would require controlling for structural and exogenous trend factors, which would allow for isolation of the effect of program activity from natural changes and policy variables.

Figure 3 shows the electric service territories by town for the PAs and municipal utilities in Massachusetts in 2015. The towns served by the municipal utilities that were included in the residential model were fairly well distributed across the state.

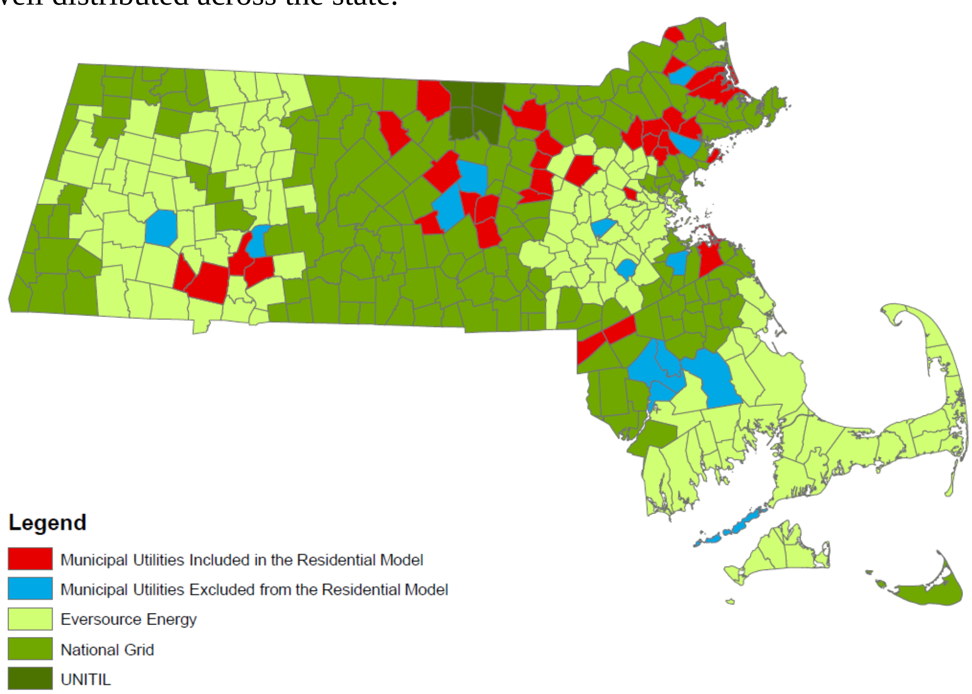


Figure 3. Electric Service Territories by Town, Massachusetts, 2015

For the residential sector, we specified a fixed-effects panel regression model. This type of regression model allows each PA/utility to act as its own control.[3] Program activity, the variable of interest for estimating program impacts, was incorporated through program expenditure data. The model was specified with lagged program activity variables to account for the time between program implementation and program-induced reductions in electricity consumption, and because of the fact that energy efficiency investments continue to yield savings for the life of measures installed.

The top-down model sought to estimate the impact of energy efficiency program expenditures on electricity consumption by separating that effect from other causes of changes in usage. The fixed-effects model estimates electricity consumption as a function of PA/utility current-year and past-year electric energy efficiency program expenditures, electricity prices, weather, and economic and demographic factors. The regression form used is as follows:

$$\log(EC_{it}) = \beta_1 \, log(P_{it}) + \beta_2 \log(HDD_{it}) + \beta_3 \log(CDD_{it}) + \beta_4 \log(I_{it}) + \beta_5 \, EH_{it} + \beta_6 VAL_{it}$$
$$+ \beta_7 \, NC_{it} + \beta_8 \, SF_{it} + \beta_9 \, RENT_{it} + \beta_{10} \, EMP_{it} + \sum_{j=0}^{n} \alpha_j EE_{it-j} + \beta_{11} \tau_t + \delta_i + \varepsilon_{it}$$

Where:

$\log(EC_{it})$	=	Natural logarithm of annual consumption per residential customer in PA/utility service area *i* and year *t*;
$\log(P_{it})$	=	Natural logarithm of electricity price in 2012 dollars;[4]
$\log(HDD_{it})$	=	Natural logarithm of annual heating degree days (base 65);
$\log(CDD_{it})$	=	Natural logarithm of annual cooling degree days (base 70);
$\log(I_{it})$	=	Natural logarithm of median household income in 2012 dollars;
EH_{it}	=	The share of households using electricity as the primary heating fuel;
VAL_{it}	=	The median house values in 2012 dollars;
NC_{it}	=	The share of new construction in residential housing, computed as the total number of residential new construction permits divided by the total number of housing units;
SF_{it}	=	The share of single-family homes in residential housing;
$RENT_{it}$	=	The share of renters;
EMP_{it}	=	The employment rate, computed as the number of employees divided by the number of people in the labor force;
EE_{it-j}	=	Total residential electric energy efficiency program expenditures per residential customer-*j*. The coefficient α_j measures the percentage change in electricity consumption in year *t* from a one-dollar change in energy efficiency program expenditures in year *t-j*. The sum of α_0 through α_n measures the percentage change in electricity consumption in year *t* from a one-dollar change in energy efficiency program expenditures in year *t* and the previous *n* years.[5]
τ_t	=	Time-trend variable that is equal to 1 in 1990 and increasing by one unit annually. This accounts for the naturally occurring change in electricity consumption not captured by the variables included in the model;[6]

[3] The inclusion of fixed effects in the model ensures that the estimated regression coefficients are not biased due to non-time-varying (i.e., PA/utility-specific) characteristics. A random-effects specification is more efficient, but using random effects does not fully control for all utility-specific characteristics. Hausman tests were used to determine which model specification to use. The findings from those tests showed that fixed effects were more appropriate for this analysis.

[4] Nominal prices were adjusted to reflect 2012 dollars using the GDP implicit price deflator from the Federal Reserve Economic Data.

[5] We also tested specifications with distributed lag models with a special parameterization of lagged energy efficiency expenditures variables in order to account for the possible non-linear and delayed effects of energy efficiency program activity on consumption. The results were similar.

[6] As a robustness check, we also tested specifications with non-linear (a natural cubic spline, or some second- or third-degree

δ_i = PA/utility fixed effects that capture time-invariant PA/utility-specific fixed effects in electricity consumption. There may be a certain PA/utility-level variation in the data that is not necessarily related to energy efficiency programmatic activity;

ε_{it} = Regression error term.

Since there is a significant variation in the size of PAs and municipal utilities, the models were weighted by the amount of residential electricity sales to properly represent the different magnitudes of spending and potential savings across the PAs and municipal utilities in Massachusetts.

Finally, the Massachusetts PAs have had residential upstream lighting programs since 1998, and these programs have accounted for a significant share of program-claimed savings for the PAs. The incentive structure of these programs does not allow for assurances that each purchaser of a program bulb is a residential customer in the sponsoring PA's service territory. Therefore, some program bulbs may have been purchased by customers served by municipal utilities. This leakage means that some of the program expenditures in the neighboring area are affecting consumption in the municipal utility area. Based on a component of the Massachusetts Residential Customer Profile study, which allocated total upstream program rebate dollars to Census block groups in Massachusetts from 2010 through 2013, we reallocated a portion of PA electricity program expenditures to municipal utilities.

Model Results

Table 1 shows the coefficient estimates for the key explanatory variables from six different residential models.

Table 1. Residential Model Results

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Annual residential energy efficiency program expenditures per customer in year t	-0.00014 (0.0002)	0.00038+ (0.0002)	0.00031 (0.0003)	-0.00012 (0.0002)	0.00040 (0.0002)	0.00032 (0.0003)
Annual residential energy efficiency program expenditures per customer in year t-1		-0.00046** (0.0001)	-0.00033 (0.0003)		-0.00049** (0.0001)	-0.00037 (0.0003)
Annual residential energy efficiency program expenditures per customer in year t-2		-0.00028 (0.0004)	-0.00030 (0.0003)		-0.00029 (0.0004)	-0.00032 (0.0003)
Annual residential energy efficiency program expenditures per customer in year t-3		-0.00066* (0.0003)	-0.00073** (0.0003)		-0.00068* (0.0003)	-0.00078** (0.0003)
Annual residential energy efficiency program expenditures per customer in year t-4		-0.00150** (0.0004)	-0.00128** (0.0002)		-0.00153** (0.0004)	-0.00132** (0.0003)
Annual residential energy efficiency program expenditures per customer in year t-5			-0.00110 (0.0011)			-0.00111 (0.0011)
Annual residential energy efficiency program expenditures per customer in year t-6			-0.00019 (0.0009)			-0.00023 (0.0009)
Estimation method	FE	FE	FE	FE	FE	FE

polynomials) time trends. This had little effect on the results. Similarly, including the indicator variables for individual years instead of a time trend did not result in a significant change in the model results.

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Cumulative residential energy efficiency program expenditures per customer in years t-4 through t	N/A	-.00252** (0.0007)	-.00234** (0.0005)	N/A	-.00259** (0.0008)	-.00247** (0.0005)
Cumulative residential energy efficiency program expenditures per customer in years t-6 through t	N/A	N/A	-.00363** (0.0006)	N/A	N/A	-.00380** (0.0006)
Observations	438	422	414	438	422	414
Within R^2	0.64	0.69	0.71	0.64	0.69	0.71
Years included	2000-2012	2000-2012	2000-2012	2000-2012	2000-2012	2000-2012
Account for leakage of PA-supported CFLs to municipal utility customers	NO	NO	NO	YES	YES	YES
Number of utilities	35	35	35	35	35	35

Notes: In all models, the dependent variable is the natural logarithm of annual electricity consumption per customer. All independent variables are in natural log forms except the variables expressed as percentages and energy efficiency program expenditure variables. Observations are weighted by PA/utility annual total residential sales. + $p<0.10$, * $p<0.05$, ** $p<0.01$.

In Model 1, current electricity consumption is modeled as a function of current-year energy efficiency expenditures and other factors affecting electricity consumption. While the coefficient -0.00014 of current-year annual residential energy efficiency program expenditures per customer has the expected negative sign, it is not statistically significant at a 90% confidence level. This model does not capture the lagged impact of the energy efficiency programs on energy consumption. In addition, the impact of current program expenditures on current consumption could be twice as large if program expenditures were distributed uniformly in a given year because, in that case, each dollar of current-year expenditures would affect only one-half of current-year consumption.

In Model 2, current electricity consumption is modeled as a function of current-year energy efficiency expenditures and those of the previous four years, as well as other factors affecting electricity consumption. The lagged energy efficiency expenditures included in the model capture the impact on current consumption of the measures installed in the previous four years, as well as the market effects. While all of the lagged energy efficiency program expenditure coefficients have the expected negative sign, the current-year energy efficiency program expenditures have a positive sign. The first- and fourth-year lag coefficients are statistically significant at a 99% confidence level. The coefficients of energy efficiency program expenditures are also jointly significant at a 99% confidence level ($F(5,34)=8.1$, $p=0.000$). The sum of the current and four lagged energy efficiency expenditure coefficients is -0.00252 with a standard error of 0.0007, which is also statistically significant at a 99% confidence level. The average annual residential electricity consumption in Massachusetts for years 2000 through 2012 was 7,533 kWh per customer. The model results suggests that one dollar spent on energy efficiency per customer this year would decrease per-customer residential electricity consumption by a total of 18.98 kWh over the next four and one-half years, with a 95% confidence interval of [8.2 kWh, 29.8 kWh] or 4.2±2.4 kWh per year.

In Model 3, current electricity consumption is modeled as a function of current-year energy efficiency expenditures and those of the previous six years, as well as other factors affecting electricity consumption. The third- and fourth-year lag coefficients are statistically significant at a 99% confidence level. The coefficients of energy efficiency program expenditures are also jointly significant at a 99% confidence level ($F(7,34)=23.3$, $p=0.000$). The sum of the current and six lagged energy efficiency expenditure coefficients is -0.00363, with a standard error of 0.0006 (significant at a 99% confidence level). The model suggests that one dollar spent on energy efficiency per customer this year would decrease per-customer residential electricity consumption by a total of 27.34 kWh over the next six and one-half years, with a 95% confidence interval of [18.7 kWh, 35.9 kWh] or 4.2±1.3 kWh per year.

Models 4 through 6 repeat Models 1 through 3 except that Models 4 through 6 contain the adjustments to energy efficiency program expenditures by PAs and municipal utilities to account for the PA-supported program bulbs that were purchased by municipal utility customers. The results indicate that these adjustments improve the estimates of the impact of energy efficiency program expenditures on consumption, but only slightly.

The relatively long time-series data allowed us to test several finite distributed lag models to empirically determine the appropriate lag length. We selected the four-lag model as the most appropriate model through a statistical significance test.[7] Among the six models whose results are shown in Table 2, Model 5 with a four-year lag is our preferred model because 1) it accounts for the lagged impact of energy efficiency program expenditures on energy consumption, 2) it accounts for the leakage of PA lighting program rebate dollars to municipal utility service territories, 3) the coefficients of the first, the third, and the fourth lag are statistically significant, and 4) the coefficients of current and lagged energy efficiency expenditure variables are jointly statistically significant. This being said, the fact that the six-year lag model produces very similar results indicates that the fixed-effects model produces stable results across models with different lags.

In Model 5, the sum of the current and four lagged energy efficiency expenditure coefficients is -0.00259. This suggests that one dollar spent in energy efficiency expenditures per customer this year would decrease per-customer residential electricity consumption by a total of 19.5±12.2 kWh over the next four and one-half years, or 4.3±2.7 kWh per year.

A Comparison of Residential Top-down and Bottom-up Saving Estimates

Top-down saving estimates are not intended to replace traditional bottom-up estimates but can help validate them. Table 2 provides a comparison of top-down annual net saving estimates, using the Model 4-6 results, with the PA-reported bottom-up annual net savings in Massachusetts from 2003 through 2012. The table shows the annual net savings estimates and the corresponding lower and upper bounds of the 90% confidence intervals. The table also expresses top-down estimated savings as a percent of the annual bottom-up saving estimates to provide a top-down to bottom-up estimate ratio. The no-lag model does not capture the lagged impact of the energy efficiency programs on energy consumption. The four- and six-lag models, respectively, account for the impact of up to four and six previous years' programmatic activity on the current year's consumption. The four-lag model, which provided the best statistical fit to the data, shows a top-down to bottom-up ratio of 187%, but the 90% confidence interval ranges from 92% to 282%.

Table 2. Comparison of Residential Top-down and Bottom-up Estimates, 2003-2012

	Top-down Annual Net Saving Estimates (GWh)			Top-down Annual Net Saving Estimates (% of Net Bottom-up Estimates)		
Lag Structure	**Lower Bound**	**Point Estimate**	**Upper Bound**	**Lower Bound**	**Point Estimate**	**Upper Bound**
No Lag	-1,366	784	2,935	-68%	39%	146%
Four Lags	1,851	3,762	5,674	92%	187%	282%
Six Lags	2,829	3,821	4,814	141%	190%	240%

Notes: The source of residential electric program reported savings and expenditures is Massachusetts Division of Energy Resources' (DOER's) PARIS database. Upper and lower bounds are for a 90% confidence interval.

[7] In this method, the way to choose the length of a lag is to start with a long lag, test the statistical significance of the coefficient at the longest lag—the "trailing lag"—and shorten the lag by one period if one cannot reject the null hypothesis that the effect at the longest lag is zero. One continues shortening the lag until the trailing lag coefficient is statistically significant.

Commercial and Industrial Model

This section provides details on C&I model specification and presents the model results. First, we discuss the C&I data collected for this project. Next, we present the model specification. We then present and discuss the C&I model results. Finally, we present a comparison of top-down and bottom-up estimates of C&I net savings.

Data Collection

For the C&I model, the electricity consumption, price, electric program activity, and weather data were from the same sources as the residential model. The C&I model included data from all PAs and the same municipal utilities included in the residential model plus one additional municipal utility. We gathered economic and firmographic data from the following sources:

- US Census ZIP Business Patterns – Contains business (establishment) and employee counts by size and by North American Industry Classification System (NAICS) industry type, summaries by ZIP code (without industry breakdown) for employment, payroll, and counts by employment size.
- Bureau of Labor Statistics – Contains annual labor force, employment, and unemployment counts at the town level of granularity.
- McGraw Hill Dodge C&I New Construction Database – This database, purchased as part of the Massachusetts C&I Program Evaluation, contains information on project square footage, value, type, and location for all non-residential new construction projects in Massachusetts from 1996 through 2011.

Model Specification

As with the residential model, we specified the C&I models as a fixed-effects panel regression model. We first considered using C&I electricity consumption intensity—i.e., electricity use per square foot of floor space—as the dependent variable. However, reliable information on square footage was not available from the public data sources. As an alternative way of normalizing consumption across the C&I customers, we used C&I consumption per employee as our dependent variable in our models.[8] The top-down model estimates C&I electricity consumption as a function of PA/utility current- and past-year C&I program expenditures per employee, electricity prices, weather, and economic and firmographic factors. The regression form used is as follows:

$$\log(EC_{it}) = \beta_1 \, log(P_{it}) + \beta_2 \log(HDD_{it}) + \beta_3 \log(CDD_{it}) + \beta_4 \log(EINC_{it}) + \beta_5 NC_{it} + \beta_6 EMP_{it} + \sum_{k=1}^{20} \gamma_k \, NAICS_{k,it} + \sum_{j=0}^{n} \alpha_j EE_{it-j} + \beta_7 \tau_t + \delta_i + \varepsilon_{it}$$

Where:

$\log(EC_{it})$	=	Natural logarithm of annual consumption per employee in PA/utility service area *i* and year *t*;
$\log(P_{it})$	=	Natural logarithm of electricity price in 2012 dollars;
$\log(HDD_{it})$	=	Natural logarithm of annual heating degree days;
$\log(CDD_{it})$	=	Natural logarithm of annual cooling degree days;
$\log(EINC_{it})$	=	Natural logarithm of mean annual employment income per employee, in 2012 dollars, computed as total annual payroll divided by total number of employees;
NC_{it}	=	Square footage of C&I new construction per employee;
$\text{NAICS}_{k,it}$	=	The percent of establishments in a two-digit NAICS industry code *k*. The γ_k is a vector of coefficients that capture the differences in building energy use by business type;

[8] We also tested models in which C&I consumption was expressed as per-customer or per-establishment. The results from these models were comparable.

EMP_{it}	=	The employment rate;
EE_{it-j}	=	Total C&I energy efficiency program expenditures per employee in PA/utility service area *i* and year *t-j*;
τ_t	=	Time-trend variable that is equal to 1 in 1990 and increasing by one unit annually;[9]
δ_i	=	PA/utility fixed effects that capture time-invariant, PA/utility-specific fixed effects in electricity consumption;
ε_{it}	=	Regression error term *t*.

Model Results

Table 3 shows the coefficient estimates for the key explanatory variables from four different C&I models.

Table 3. C&I Model Results

Variable	Model 1	Model 2	Model 3	Model 4
Annual C&I energy efficiency program expenditures per employee in year t	-0.00029+ (0.0002)	-0.00018 (0.0001)	-0.00018 (0.0001)	-0.00017 (0.0001)
Annual C&I energy efficiency program expenditures per employee in year t-1		-0.00025* (0.0001)	-0.00024* (0.0001)	-0.00018+ (0.0002)
Annual C&I energy efficiency program expenditures per employee in year t-2		-0.00011 (0.0002)	-0.00009 (0.0002)	-0.00008 (0.0002)
Annual C&I energy efficiency program expenditures per employee in year t-3		-0.00036** (0.0001)	-0.00033* (0.0001)	-0.00026+ (0.0002)
Annual C&I energy efficiency program expenditures per employee in year t-4			0.00010 (0.0001)	0.00011 (0.0002)
Annual C&I energy efficiency program expenditures per employee in year t-5				0.00044** (0.0001)
Annual C&I energy efficiency program expenditures per employee in year t-6				0.00043* (0.0002)
Estimation Method	FE	FE	FE	FE
Cumulative C&I energy efficiency program expenditures per customer in years t-3 through t		-0.00091* (0.0004)		
Cumulative C&I energy efficiency program expenditures per customer in years t-4 through t			-0.00075 (0.0005)	
Cumulative C&I energy efficiency program expenditures per customer in years t-6 through t				0.00029 (0.0007)
Observations	379	379	379	379
Within R^2	0.39	0.40	0.40	0.43
Years Included	2002-2012	2002-2012	2002-2012	2002-2012
Number of Utilities	36	36	36	36

Notes: In all models, the dependent variable is the natural logarithm of annual electricity consumption per employee. All independent variables are in natural log forms except the variables expressed as percentages, C&I new construction, and energy efficiency expenditures variables. Observations are weighted by PA/utility annual total C&I sales. + $p<0.10$, * $p<0.05$, ** $p<0.01$

In Model 1, current electricity consumption is modeled as a function of current-year energy efficiency expenditures and other factors affecting electricity consumption. The coefficient -0.00029 of current-year annual C&I energy efficiency program expenditures per employee, which is statistically

[9] We also tested specifications with non-linear (a natural cubic spline, or some second or third degree polynomials) time trends. This had little impact on the results.

significant at a 90% confidence level, suggests that a one-dollar increase in C&I program expenditures per employee in a given year would decrease the per-employee electricity consumption by about 0.029% in that year. The average annual C&I electricity consumption per employee in Massachusetts for years 2002 through 2012 was 12,560 kWh. This suggests that one dollar spent in energy efficiency expenditures per employee in a given year would decrease per-employee C&I electricity consumption by a total of 3.6 kWh in that year with a 95% confidence interval of [0.5 kWh, 6.8 kWh].

Model 2 adds the previous three years' energy efficiency expenditures to the specification. While the current and lagged energy efficiency program expenditure coefficients all have the expected negative sign, only the first- and third-year lag coefficients are statistically significant at a 95% confidence level. The coefficients of energy efficiency program expenditures are also jointly significant at a 99% confidence level ($F(4,35)=6.63$, $p=0.0004$). The sum of the current and three lagged energy efficiency expenditure coefficients is -0.00091 with a standard error of 0.0004, which is also statistically significant at a 95% confidence level. The model suggests that one dollar spent in energy efficiency expenditures per employee this year would decrease per-employee C&I electricity consumption by a total of 11.4 kWh over the next three and one-half years, with a 95% confidence interval of [2.55 kWh, 20.3 kWh] or 3.2±2.54 kWh per year.

Model 3 adds the previous four years' energy efficiency expenditures to the specification. While the current and previous three years' energy efficiency program expenditure coefficients all have the expected negative sign, the fourth year's coefficient is positive. Similar to Model 3, the first and third lag coefficients are negative and statistically significant. In addition, the sum of the current and four lagged energy efficiency expenditure coefficients is not statistically significant.

Model 4 adds the previous six years' energy efficiency expenditures to the specification. While the current and previous three years' energy efficiency program expenditure coefficients all have the expected negative sign, the fourth, fifth, and sixth year's coefficients are all positive. Moreover, the fifth and sixth year coefficients are statistically significant. The sum of the current and four lagged energy efficiency expenditure coefficients is also positive but not statistically significant.

In general, the C&I model results were less consistent than the corresponding residential models, and the model fit is not as strong. We think that the C&I top-down models did not work as well as the residential ones for the following reasons:

- Electricity consumption in the C&I sector when combined into a single sector is much more volatile than in the residential sector due to high variability in industrial consumption.
- The economic and firmographic variables included in the models were not sufficient as exogenous controls.
- Normalizing consumption by the number of customers, establishments, or employees may not work as well as normalizing it by the square footage of floor space.

A Comparison of C&I Top-down and Bottom-up Saving Estimates

Table 4 provides a comparison of top-down annual net saving estimates, using the Model 4-6 results, with the PA-reported bottom-up annual net savings in Massachusetts from 2003 through 2012. The three-lag model, which provided the best statistical fit to the data, shows a top-down to bottom-up ratio of 101%, and the 90% confidence interval ranges from 28% to 174%. The fact that C&I models with different lag lengths produce different results is expected given that consumption in the C&I sector is more volatile than that in the residential sector and the customer base is more heterogeneous.

Table 4. Comparison of C&I Top-down and Bottom-up Estimates, 2003-2012

Lag Structure	Top-down Annual Net Saving Estimates (GWh)			Top-down Annual Net Saving Estimates (% of Net Bottom-up Estimates)		
	Lower Bound	Point Estimate	Upper Bound	Lower Bound	Point Estimate	Upper Bound
No Lag	-501	3,727	7,956	-15%	112%	240%
Three Lags	925	3,342	5,758	28%	101%	174%
Four Lags	-207	2,142	4,491	-6%	65%	136%
Six Lags	-2,850	-573	1,703	-86%	-17%	51%

Notes: The source of C&I electric program reported savings and expenditures is Massachusetts Division of Energy Resources' (DOER's) PARIS database. Upper and lower bounds are for a 90% confidence interval.

Study Results and Conclusions

While the findings from this pilot study are preliminary, initial model results look promising as an alternative method that may potentially support and help validate the bottom-up estimates of net savings. While the preliminary indicators suggest that the program effects identified by bottom-up approaches were real and may even understate the program-induced savings in Massachusetts, further research is needed to explore the stability and sensitivity of the model results.

The study also draws attention to an inherent limitation of top-down methods. While our models were able to detect energy savings, they were not estimated precisely. The confidence intervals around the top-down estimates of savings were quite large. If the further research establishes that the results are stable against alternative model specifications, then we recommend exploring ways to reduce the width of the confidence intervals around the estimates.

References

Auffhammer, M., University of California, B., Blumstein, C., & Fowelie, M. (2007). Demand-Side Management and Energy Efficiency Revisited.

CADMUS Group, I. (2012). *CPUC Macro Consumption Metric Pilot Study (Final Report).* Prepared for the California Public Utilities Commission.

Demand Research, L. (2012). *Macro Consumption Metrics Pilot Study Technical Memorandum -Final Report.* Prepared for the California Public Utilities Commission.

Horowitz, M. J. (2007). Changes in Electricy Demand in the United States from the 1970s to 2003. *The Energy Journal, Vol. 28, No. 3*.

Loughran, D. S. (2004). Demand-Side Management and Energy Efficiency in the United States. *The Energy Journal, Vol. 25 No. 1*.

Rivers, N., & Jaccard, M. (2011). Electric Utility Demand Side Management in Canada. *The Energy Journal, Vol. 32, No. 4*.