Rolling Up IOU Account- Level Data to Measure Savings from the Top-Down

Noel Stevens, DNV GL, Burlington, MA
Monica Cohen, Columbia Gas of Massachusetts, Westborough, MA
Andrew Stryker, DNV GL, Oakland, CA
Jon Farland, DNV GL, Oakland, CA
Pam Rathbun, Tetra Tech Madison, WI

ABSTRACT

Traditional bottom-up approaches to energy efficiency program evaluation have well-documented technical and cost limitations. Regulatory agencies and investor-owned utilities (IOUs) have begun to explore top-down modeling, as a supplemental or alternative approach to measuring net energy impacts. The goal of this modeling technique is to isolate the effect of program activity using a holistic approach, estimating program impacts across all energy-efficiency programs in a given geographical region or service territory, rather than estimating savings with separate studies for each program or measure/end-use within a program. This study develops a set of statewide top-down models to estimate the change in aggregate consumption per unit (e.g., household or employee) in a geographic region (i.e., town or county) as a function of program expenditures, or ex ante savings, and economic conditions. This modeling approach uses account level investor-owned utility (IOU) billing and program tracking data which is aggregated to the town and county levels. That data is used to measure changes in normalized annual consumption resulting from changes to programmatic activity and other exogenous factors. In contrast to many existing top-down studies, the use of account -level data provides detailed information regarding program activity level and provides the flexibility to aggregate by the dimensions that are relevant to the specific study, such as county, town or IOU service territory. Weather normalizing consumption at the account level and aggregating to the desired level of geography also allows us to examine separate models for by sector and isolate separate impacts by measure type.

Introduction

Top-down modeling is an econometric approach to measure program impacts using aggregate cross-sectional and time series data. The top-down models measure changes to aggregate energy consumption relative to changes in energy efficiency programmatic activity, prices, and other economic factors. The goal of this type of modeling is to isolate the effect of program activity from other natural changes and policy variables. Top-down techniques use a holistic approach by estimating program impacts across all energy-efficiency programs in a given geographical region or service territory, rather than running separate studies for each program (or measure/end-use within a program).

The Massachusetts Program Administrators (PAs) expressed interest in determining whether top-down modeling should play a role in net energy impact evaluation. The PAs are exploring the potential for top-down techniques to be an additional tool to complement bottom-up approaches to measuring impacts from energy efficiency programs, as top-down techniques have the potential to capture interactive effects between programs, and market effects.[i]

The goal of this study is to develop and apply multiple top-down methods for Massachusetts, and to understand the strengths and limitations of those methods relative to the traditional bottom-up approaches to measuring net energy impacts. To address this goal, the evaluation team focused on the following research objectives:

- Review existing top-down modeling techniques, and recommend specific methods to be used in Massachusetts
- Obtain the necessary data for employing one or more agreed-upon approaches
- Implement multiple agreed-upon approaches in parallel to evaluate and recommend top-down modeling techniques that may contribute to longer on-going evaluation efforts
- Establish the preferred model specifications and data requirements for the Year 2 and beyond top-down modeling efforts for regional and national models using different approaches. Due to time and data limitations, we restricted this year's analysis to electric consumption models only.

Analytical Framework for Reviewing Top-down Modeling Approaches

This section summarizes the expected outcomes and requirements for the use of a top-down method in evaluating the expected savings from a portfolio of programs. In developing this framework, we reviewed 15 top-down research studies that have been used to estimate impacts associated with energy efficiency programs. First, we discuss the possible benefits of including top-down modeling in the set of tools used to evaluate energy efficiency programs. Next, we discuss some important limitations to the use of top-down methods in program evaluation. Then, we discuss the necessary requirements for successful top-down models, which will form the basis for our comparison of techniques in the following sections. Finally, we provide a brief review of previous top-down literature.

Evaluators and commissioners are intrigued by the potential of top-down techniques because of their potential to provide low-cost supplemental or alternative estimates of net program savings. The appeal of this structure is that, in principle, it captures the full program effect, including free-ridership, spillover, market effects, and snapback. A properly structured top-down model can potentially make the following contributions to the set of tools used to evaluate energy efficiency programs:

- Provide relatively inexpensive estimates of program-induced savings estimates for all units in the study. Because top-down models can examine the change in total energy consumption over time, relative to changes in the level of program activity, it is theoretically possible to measure the overall impact of all programmatic activity across the portfolio of energy efficiency programs net of free-riders, spillover, and snapback. However, results do include incremental program induced savings as well as some level of naturally occurring savings due to the inability to fully remove free-riders from self-selected program participants.
- Provide expected program-induced savings for a unit with particular characteristics. Top-down models are essentially macro-economic models that relate energy consumption to a series of program and non-program (exogenous) variables. The non-program variables are used to describe the relative impact of a population's demographic or firm-o-graphic characteristics. Consequently, by using assumed values for these characteristics, evaluators, program designers, and policy makers can estimate the expected level of savings from populations with the prescribed characteristics, given an assumed level of program activity.
- Provide combined effects of all cumulative program activity for a particular unit, including spillover and snapback. By measuring the change in aggregate consumption relative to the level of program activity across the combined set of programmatic activities over time, top-down models naturally account for the interaction among program impacts. Consequently, they provide measures of the cumulative effect of the portfolio over time, including spillover (i.e., untracked program attributable savings).

- Provide confidence intervals and precision levels for net energy savings from the portfolio of programs. The statistical estimation of top-down models yields estimates of the standard error that can be used to construct confidence intervals around the savings estimates.

  Despite the potential advantages, there are several challenges to developing robust top-down models:

- Models provide average, not specific, program effects. The extent to which some programs are more efficient than others will not be reflected in the model results.

- Data availability limits possible model specifications. To estimate such a model, we need economic, price, and programmatic activity data for each area and time period. Compromise is typically necessary between the ideal specification and the types of data available at various levels of aggregation.

- Spillover between study units reduces apparent program effects. To the extent that program activity in one area and time period reduces consumption in another area or time period, the program effect estimated by the regression will be dampened. The difference in consumption per unit across areas and time periods with different levels of program activity will be reduced, so that the measured program effect will be less.

- Omitted or incorrectly specified variables can bias the results in unknown ways. A key premise of the top-down model is that all non-program factors affecting consumption are accounted for in the model. However, model specifications tend to be limited by available data. To the extent that there are other drivers of consumption not accounted for in the model structure, or not incorporated with the best metric or form of relationship, all the model coefficients including the program coefficient can be biased.

- Self-selection effects can bias the results. If program activity tends to be higher in areas and times where customers would have a natural tendency to adopt more (or less) energy efficiency on their own, and there is no good metric of this natural tendency that can be incorporated to control for it, the estimated program effect will be overstated (or understated). This is a special case of omitted variable bias.

These technical constraints result in the following limitations in top-down models:

- Inability to obtain savings estimates net of free riders. While some authors have argued that top-down models provide true measures of net savings, net of free-riders, this is not necessarily the case for a number of reasons. First, top-down models that look at varying levels of program activity within a geographical region (i.e., a state or PA territory) are unlikely to contain a true measure of the "no program" scenario. As has been discussed extensively regarding bottom-up approaches, nonparticipants are not good measures of participant behavior absent the program, as participants are inherently different from nonparticipants. Second, while it is possible to use comparison groups (i.e., other states), comparison groups are likely to have a lower percentage of energy efficient equipment than a true control group within the study population.

- Inability to provide separate free-ridership, spillover, and market effects estimates. Top-down models may provide measures of "near net" savings, but separate estimates for free-ridership, spillover, and market effects are imbedded in the final "near net" savings estimates.

- Inability to provide an isolated effect of a particular program and year. Information concerning the relative contribution of separate programs to the overall savings associated with a portfolio of programs is important to policy makers and program designers. This information provides for more efficient allocation of resources to meet savings goals.

- Inability to identify which groups of measures are performing better, or worse, given their characteristics. Such information is valuable to policy makers, program designers, and implementers.

While it is unlikely that any one tool will contain all of these attributes, we established the following set of criteria for top-down models to provide useful tools for evaluating energy efficiency programs. The choice of specific top-down method will result in different strengths and weaknesses. No one model will provide a "silver bullet" to address all the relevant concerns. Rather, running a variety of models over time is most likely to provide a more comprehensive view of net program savings that can be used in conjunction with bottom-up techniques to triangulate the net impact of programmatic activity on energy consumption. Ability to establish the counterfactual (no-program) scenario – Understanding the true extent of a program's impacts requires information regarding the level of consumption absent from any programmatic activity. Without using a random experimental design with a true "control group," the counterfactual scenario is typically simulated using a number of techniques, each of which has certain limitations.

- Diversity of program activity levels across units of observation (time-geography combinations) – In order to detect the impact of programmatic changes on consumption, the level of programmatic activity for each unit of observation must be reported consistently. Moreover, the program variables must provide sufficient variation from one location-time combination to relate changes in consumption to the program variables.
- Consistent relationship between program activity and savings – The influence of program variables and consumption must be consistent across units of observation.
- Minimal effect of one area on another (cross-area spillover) – Non-program tracked spillover from one area to another will assign exaggerated program effects in the region experiencing the spillover, and reduce the apparent impacts of the region that is the source of the spillover.
- Appropriate and consistent use of exogenous explanatory variables – The model must account for exogenous differences between units of observation. This requires inclusion of relevant explanatory variables that are reported at a comparable level of analysis as the dependent variable, scaling variable, program variables, and other independent variables used in the model.
- Ability to measure program activity at the most granular geographic level –While explanatory variables measure the relative influence of demographic factors on savings, top-down estimates based on data aggregated at too high of a geographic area (i.e., state) may lose the ability to provide meaningful estimates of variables important to all interested parties. To retain this desirable level of information, top-down models should seek to measure program activity at the most granular geographic level possible.
- Long enough time series to detect and isolate program impacts – It is important that the time series is sufficiently long to contain adequate variability in the program variable(s) that is distinguishable from exogenous factors.
- Account for the lag structure of program impacts – The time series must be sufficiently long to allow for the lag structure in program impacts, relative to the level of program activity. There is a lag between program marketing efforts and the realization and reporting of program savings. There must be sufficient data in the time series to capture programmatic activity from at least one or two periods prior to the first observation in the consumption history.

Massachusetts PA-data Commercial and Industrial (C&I) Top-down Pilot Study Methodology

Based on the analytical framework for conducting successful top-down models, we developed the top-down modeling pilot study described in the sections that follow. This section describes the modeling specifications and process we used to construct a pilot study of top-down modeling for Commercial and Industrial (C&I) electric customers using the Massachusetts program administrators' (PAs) account level billing and tracking data. Equation 1 presents the PA-data C&I electric top-down model. We use δ (delta) to highlight the use of the first difference (i.e., year-over-year change) in variables used in the model specification. This is a standard billing analysis approach, extended to the cross-sectional economic aggregate, that works well when there is a limited number of time series observations, as exists during the initial model estimation phase in this study.

Equation 1. The PA Commercial/Industrial Model

$\delta(\text{NAC})\ tsgf = \beta 0\ sgf + \beta 1*[\delta\text{Employment}]tsg + \beta 2*[\delta\text{EE \$ Program Activity}]\ tsgf + \beta 4*£\ sgf + \beta 4*¥\ tsf\ \varepsilon\ tsgf$

Where each variable in Equation 1 2 is defined as follows:

- β0 sgf is a fixed effects variable for sector (s), within geographic region (g), and by fuel type (f)
- (NAC) tsgf = Normalized (C&I) Annual Energy Consumption in year (t), sector (s), within geographic region (g), and by fuel type (f). For the county-level models, all variables are divided by gross domestic product to provide a measure of energy intensity per unit of output. For the town-level models, population is used in place of GDP due to data limitations.
- Employmenttsg = Economic activity measured as the total employment per GDP or population, for county and town-level models, respectively, within year (t), sector (s), within geographic region (g)
- Program activity = We considered two separate measures of programmatic activity separately:
- EE $ Program Expenditure Vbl(s) tsgf is one or more EE program variables measured in $s, reflecting program expenditures as reported in the PA program tracking data, in year (t), sector (s), within geographic region (g), and by fuel type (f), and,
- EE Program Energy Savings Vbl tsgf is a measure of estimated EE savings, as reported in the PA program tracking data, in year (t), sector (s), within geographic region (g), and by fuel type (f = electricity or natural gas).
- £ sgf = Parameter for geographic fixed effects for county or town g in sector s, and fuel type f.
- ¥ tsf = Parameter for annual fixed effects for year t in sector s, and fuel type f.

  We used the following steps to develop the county- and town-level models:

- Model of total NAC verses economic activity – Before introducing program activity and other variables, we first investigated whether changes to NAC could be explained by changes in employment, as well as geographic and annual fixed effects. We constructed these simplified consumption models at two separate levels of analysis:
  - County-level model – This model used NAC per unit of GDP as a dependent variable

  - Town-level model – Due to lack of available GDP data at more granular levels of analysis, the town-level model used NAC per capita as a dependent variable. Since only three years of data were available, we explored the non-differenced version of each model to provide for estimation across all three years of available data
- Introduce measures of program activity – After determining that we could successfully model NAC as a function of the economic variable, we introduced the following two measures of programmatic activity separately. We considered the measures separately to limit collinearity:
  - Aggregate energy efficiency expenditures per unit – We obtained account- and measure-level downstream program expenditures and measure- and location-specific upstream data from the PA tracking data
  - Aggregate ex-ante savings per unit – We obtained account- and measure-level downstream program savings and measure- and location-specific upstream data from the PA tracking data
- Due to data limitations, we were not able to include measure of the lag in program activity. We did attempt to construct lagged variables based on data provided within the PAs' annual reports. However, we concluded that the data contained in the PA annual reports and in the program tracking data were too dissimilar. We could not, with confidence, use the allocation of program tracking data by geography to allocate the data contained in the annual reports without making arbitrary assumptions regarding the differences between these two series:
- Separate program activity by program type – We examined the impact of separating program expenditures and ex-ante savings into upstream and downstream activity, and examined the impact of lighting and non-lighting program activity on NAC.
- Estimate first-difference form of each model – We estimated each of the models specified in their first differenced form (i.e., the year over year change in the dependent variable and corresponding independent variables).

We employed these steps to estimate both county- and town-level non-differenced and differenced forms of each of the models identified in Table 1.

Table 1. Alternative Model Descriptions for PA Data C&I Models

Model	Model Name	Model Description
Model 1	Employment Only	NAC is a function of employment plus time and geography fixed effects only
Model 2	Employment Plus Ex Ante Savings	NAC is a function of employment plus ex ante savings and time and geography fixed effects
Model 3	Employment Plus Total Expenditures	NAC is a function of employment plus time total program expenditures and geography fixed effects
Model 4	Upstream Plus Total Downstream Expenditures	NAC is a function of employment plus upstream and downstream program expenditures and time and geography fixed effects
Model 5	Upstream plus Lighting and Non-lighting Downstream Expendutures	NAC is a function of employment plus upstream and downstream expenditures and time and geography fixed effects. Downstream expenditures not are separated into lighting and non-lighting.
Model 6	Upstream Plus Total Downstream Savings	NAC is a function of employment plus upstream and downstream program savings and time and geography fixed effects
Model 7	Upstream plus Lighting and Non-lighting Downstream Savings	NAC is a function of employment plus upstream and downstream ex ante savings and time and geography fixed effects. Downstream savings are separated into lighting and non-lighting.

Massachusetts PA-data Commercial and Industrial (C&I) Top-down Pilot Study Results

This section presents the results of the seven models discussed Section 4. We review the estimation results for the county-level models only, as the town level models did not provide any additional explanatory power over the county level models. We examine both the non-differenced and the differenced model results for small and large commercial and industrial sectors. Then, we examine results for these models estimated using the first difference in dependent and independent variables. Table 2a shows the results for the seven non-differenced county-level models. The first model estimated (Model 1) used only employment to predict the estimated level of NAC. The results show that for small businesses, employment-only was a good predictor of NAC, as the parameter estimate was both statistically significant and positive. This model also was significant for large commercial and total commercial but not for the industrial sector. Model 2 adds in the total ex ante savings. The Model 2 results show that it was not possible to have statistically significant terms for both savings and NAC. The same is true for Model 3 (employment plus total program expenditure). As program expenditures are further split into upstream and downstream expenditures (Model 4), the parameter estimates on expenditures are negative and significant; however, the estimated parameter for employment is still not significant. When downstream expenditures are separated into lighting and non-lighting expenditures (Model 5), the parameters are significant, but the sign on downstream non-lighting expenditures is positive.

None of the savings models show statistically significant results for the program variables for any sector. This is likely due to the sharp increase in programmatic activity that coincides with the recovering economy, and the corresponding increases in employment and consumption. While separating programmatic activity into program types does provide some degree of variation across observational units to allow the model to separate some of the programmatic impacts from changes in the economy, there is insufficient data to separate changes due to program activity from the general trend in the limited time series.

Table 2b shows the results for the same set of models using the first-difference series. This table shows that differencing does improve the statistical significance of some of the commercial models. This is likely because differencing reduces the impact associated with year-over-year changes in the economy to allow the model to isolate impacts that result from changes in the program variables relative to NAC. The county level PA Data model was not able to detect a statistically significant relationship between programmatic activity and consumption. However, based on the results of the second pilot study (the PA-Muni model) as well as previous top-down research, we speculate that it is necessary to have a sufficiently long time-series to account for the cumulative effects of program expenditures over time.[ii] A longer time series would allow the evaluation team to test for what the PA-Muni model found to be a lagged effect of program spending and energy savings. In addition to expanding the time series, it may be desirable to explore sector level models such as retail, manufacturing, and public sector. Another possibility would be to expand the scope to a regional analysis as opposed to just Massachusetts. This is because the introduction of variables that provides greater differentiation of programmatic activity across units (i.e., upstream and downstream lighting and non-lighting activity) shows greater significance of program variables. However, we note that the timeframe being studied would be particularly challenging given the period of economic decline and recent growth, as well as the recent escalation in programmatic activity.

Table 2a. Non-differenced Commercial and Industrial Model - County Level

	Model Specification 1 Employment Only		Model Specification 2 Employment Plus		Model Specification 3 Employment Plus		Model Specification 4 Upstream Plus		Model Specification 5 Upstream plus		Model Specification 6 Upstream Plus		Model Specification 7 Upstream plus	
Economic Sector	Estimate	P-Value	Estimate	P-Value	Estimate	P-Value	Estimate	P-Value	Estimate	P-Value	Estimate	P-Value	Estimate	P-Value
Small Commercial														
Total kWh Savings			0.074	0.231										
Downstream Total Savings													0.014	0.860
Downstream Lighing Savings											-0.026	0.722		
Upstream Total/Lighting Savings											-0.004	0.916	-0.016	0.570
Downstream Non-lighting Savings											0.021	0.481		
Total kWh Expenditures					**0.067**	**0.025**								
Downstream Total Expenditures									-0.005	0.890				
Downstream Lighing Expenditures							**-0.077**	**0.084**						
Upstream Total/Lighting Expenditures							-0.011	0.563	-0.029	0.209				
Downstream Non-lighting Expenditures							**0.035**	**0.071**						
Total Employment (Economic Condition)	**0.589**	**0.065**	-0.248	0.772	-0.259	0.749	0.484	0.523	-0.450	0.570	-0.043	0.966	-0.355	0.684
Large Commercial														
Total kWh Savings			0.298	0.102										
Downstream Total Savings													0.191	0.399
Downstream Lighing Savings											-0.141	0.176		
Upstream Total/Lighting Savings											0.030	0.741	0.040	0.721
Downstream Non-lighting Savings											0.079	0.512		
Total kWh Expenditures					0.305	0.138								
Downstream Total Expenditures									0.005	0.974				
Downstream Lighing Expenditures							**-0.222**	**0.002**						
Upstream Total/Lighting Expenditures							0.029	0.456	-0.027	0.751				
Downstream Non-lighting Expenditures							**0.179**	**0.012**						
Total Employment (Economic Condition)	**2.164**	**0.020**	**3.845**	**0.006**	**3.355**	**0.013**	1.527	0.337	0.705	0.846	1.470	0.585	1.299	0.700
All Commercial														
Total kWh Savings			0.015	0.880										
Downstream Total Savings													-0.150	0.117
Downstream Lighing Savings											-0.170	0.110		
Upstream Total/Lighting Savings											0.022	0.856	0.003	0.978
Downstream Non-lighting Savings											-0.044	0.327		
Total kWh Expenditures					0.021	0.793								
Downstream Total Expenditures									**0.145**	**0.000**				
Downstream Lighing Expenditures							**-0.176**	**0.062**						
Upstream Total/Lighting Expenditures							0.041	0.698	**0.055**	**0.069**				
Downstream Non-lighting Expenditures							0.042	0.402						
Total Employment (Economic Condition)	**3.409**	**0.000**	**4.667**	**0.000**	**4.684**	**0.000**	**6.391**	**0.000**	**0.625**	**0.000**	**3.538**	**0.000**	**3.508**	**0.000**
Industrial														
Total kWh Savings			0.043	0.369										
Downstream Total Savings													0.060	0.450
Downstream Lighing Savings											0.014	0.862		
Upstream Total/Lighting Savings											**0.218**	**0.062**	**0.217**	**0.029**
Downstream Non-lighting Savings											0.055	0.431		
Total kWh Expenditures					-0.013	0.792								
Downstream Total Expenditures									0.205	0.307				
Downstream Lighing Expenditures							0.145	0.299						
Upstream Total/Lighting Expenditures							-0.092	0.707	0.093	0.430				
Downstream Non-lighting Expenditures							-0.066	0.792						
Total Employment (Economic Condition)	0.184	0.632	0.018	0.989	-0.341	0.780	-0.107	0.986	4.332	0.358	3.672	0.260	3.548	0.236

Table 2b. Differenced Commercial and Industrial Model—County Level

	Model Specification 1		Model Specification 2		Model Specification 3		Model Specification 4		Model Specification 5		Model Specification 6		Model Specification 7	
	Employment Only		Employment Plus Ex Ante Savings		Employment Plus Total Expenditures		Upstream Plus Total Downstream Expenditures		Upstream plus Lighting and Non lighting Downstream Expendutures		Upstream Plus Total Downstream Savings		Upstream plus Lighting and Non lighting Downstream Savings	
Economic Sector	Estimate	P-Value	Estimate	P-Value	Estimate	P-Value	Estimate	P-Value	Estimate	P-Value	Estimate	P-Value	Estimate	P-Value
Small Commercial														
Total kWh Savings			0.043	0.493										
Downstream Total Savings													0.014	0.860
Downstream Lighing Savings											-0.026	0.722		
Upstream Total/Lighting Savings											-0.004	0.916	-0.016	0.570
Downstream Non-lighting Savings											0.021	0.481		
Total kWh Expenditures					**0.057**	**0.063**								
Downstream Total Expenditures									-0.005	0.890				
Downstream Lighing Expenditures							**-0.077**	**0.084**						
Upstream Total/Lighting Expenditures							-0.011	0.563	-0.029	0.209				
Downstream Non-lighting Expenditures							**0.035**	**0.071**						
Total Employment (Economic Condition)	**0.658**	**0.007**	0.111	0.927	-0.159	0.887	0.484	0.523	-0.450	0.570	-0.043	0.966	-0.355	0.684
Large Commercial														
Total kWh Savings			**0.368**	**0.049**										
Downstream Total Savings													0.191	0.399
Downstream Lighing Savings											-0.141	0.176		
Upstream Total/Lighting Savings											0.030	0.741	0.040	0.721
Downstream Non-lighting Savings											0.079	0.512		
Total kWh Expenditures					0.218	0.183								
Downstream Total Expenditures									0.005	0.974				
Downstream Lighing Expenditures							**-0.222**	**0.002**						
Upstream Total/Lighting Expenditures							0.029	0.456	-0.027	0.751				
Downstream Non-lighting Expenditures							**0.179**	**0.012**						
Total Employment (Economic Condition)	**1.774**	**0.028**	**5.256**	**0.017**	**6.078**	**0.008**	1.527	0.337	0.705	0.846	1.470	0.585	1.299	0.700
All Commercial														
Total kWh Savings			**0.288**	**0.012**										
Downstream Total Savings													0.131	0.237
Downstream Lighing Savings											**-0.155**	**0.004**		
Upstream Total/Lighting Savings											0.018	0.647	-0.001	0.982
Downstream Non-lighting Savings											**0.033**	**0.401**		
Total kWh Expenditures					0.114	0.167								
Downstream Total Expenditures									-0.024	0.678				
Downstream Lighing Expenditures							**-0.192**	**0.000**						
Upstream Total/Lighting Expenditures							0.018	0.462	-0.030	0.467				
Downstream Non-lighting Expenditures							**0.091**	**0.000**						
Total Employment (Economic Condition)	1.608	0.002	**5.947**	**0.000**	**6.546**	**0.000**	**1.862**	**0.052**	0.223	0.886	0.946	0.452	0.407	0.790
Industrial														
Total kWh Savings			0.051	0.377										
Downstream Total Savings													0.060	0.450
Downstream Lighing Savings											0.014	0.862		
Upstream Total/Lighting Savings											**0.218**	**0.062**	**0.217**	**0.029**
Downstream Non-lighting Savings											0.055	0.431		
Total kWh Expenditures					-0.012	0.836								
Downstream Total Expenditures									0.205	0.307				
Downstream Lighing Expenditures							0.145	0.299						
Upstream Total/Lighting Expenditures							-0.092	0.707	0.093	0.430				
Downstream Non-lighting Expenditures							-0.066	0.792						
Total Employment (Economic Condition)	0.221	0.570	0.976	0.641	0.668	0.739	-0.107	0.986	4.332	0.358	3.672	0.260	3.548	0.236

Conclusions and Limitations

This study sought to determine whether top-down methods should play a role in the overall portfolio of attribution methods both in terms of the recommended role on an ongoing basis as well as the methodological approaches that are recommended. The methods review portion of this study concluded that top-down modeling may provide an additional tool in the set of tools used to evaluate the portfolio of programs. However, the top-down approach cannot replace bottom-up approaches, as bottom-up techniques provide much information that top-down techniques cannot provide. Information pertaining to the relative contribution of different activities to overall savings can assist in the allocation of resources across the portfolio of programs, or help with program design. Such information cannot be obtained from top-down approaches. Moreover, the review of data suggested that top-down techniques face a variety of challenges pertaining to the reporting and availability of data that limited the effectiveness of these techniques. The PA Data model confirms that data availability was a primary obstacle to successful estimation of the models presented in this report. The model results were limited to just three years of data. One factor that may lead to meaningful estimates is the availability of a longer time series. With only three years of data, the PA Data pilot study portion of this report shows inconclusive evidence that the approach we employed is able to detect programmatic impacts. Our analysis demonstrates the ability to construct the necessary variables at the desired levels of aggregation, and the ability to systematically test a variety of models. Some of the models showed statistically significant parameter estimates for measures of either programmatic and/or economic activity; however, these results were not consistent across model specifications or levels of geography.

One could conclude that the statistical significance of parameter estimates for some models is an indication that the models would perform well given a sufficiently long time series. However, one could also argue that the significance of terms measuring programmatic activity is the result of noise in the model, and the true models are ones in which there are no program effects. While our analysis does not provide sufficient information to make a determination that program effects can be detected with certainty, some model results do show statistically significant parameter estimates on the program variables. Further, our review of the available literature suggests that effective top-down modeling of energy impacts requires a sufficiently long time series to account for:

- Variation in the level of program data over time – Our time series included only three years of data, which all occur during a period of economic recovery and rapid increase in programmatic activity.
- Multiple lags in programmatic activity – Previous research, as well as the PA-Muni pilot study, illustrate the importance of using multiple lags in both the program variables and dependent variable.
- Use of first-difference in the dependent and independent variables – By including only three years of data in the model, the first-difference models included in this study contain only two years of data for unit of observation.
- Absent these measures, it is not surprising that the model results did not provide statistically significant parameter estimates, that the results were not consistent across levels of aggregation, and that the results were not stable in terms of the significance of variables or their sign. However, the PA Data model is able to capture variation across program and customer types that provides valuable information for program planning and implementation, and allows program evaluators to determine the effectiveness of differing program offerings and/or marketing strategies. Compiling a sufficiently long historical time series retrospectively would be costly, and may not be possible due to limitations in electronic record keeping.

We identify the following limitations of the pilot study presented in this paper.

- Fuel prices not reported at the same level of granularity as unit of analysis – The evaluation team did not identify any data for actual average electricity prices ($/kWh) at the county or town level. The DNV GL billing data set contained rate codes, and billing amounts could therefore be imputed, but there were many missing values and other data quality concerns.
- Absence of lagged program activity and consumption variables – The literature review identified the importance of incorporating lagged program and consumption variables into the models. Because the existing time series was limited to just three years of data, we were not able to construct lagged variables using the consumption and program tracking data. The evaluation team attempted to construct lagged series based on data available through the PAs' annual reports; however, we were unable to construct a series that did not introduce bias into the model.
- Absence of building codes – The evaluation team attempted to construct variables to account for

the impact of building codes on consumption; however, due to the limited time series, there was insufficient variation in the building code data in include in the model.

- Limited time series during periods of rapid expansion of both economic and programmatic activity – Our review of the available consumption, program tracking, and economic activity variables revealed a fundamental limitation of the present analysis. During the three years of observation for this study, the three critical series for the analysis underwent a period of rapid expansion of employment, program expenditures, and ex ante savings. Given the limited time series, it is likely that the model results will be impacted by the corresponding increase in these three series. Without a longer time series, or substantial variation between observational units, it is likely that the model will not be able to differentiate between increases in programmatic activity and reductions in consumption.
- Limitations in use of per capita income – Income is not very indicative of C&I economic conditions, as it can be skewed by individuals with relatively high salaries, such as CEOs. Consequently, this variable is not used by the load forecasters at the PAs.
- Isolating industry- or sector-level differences – There may be considerable variation in the savings and consumption by industry sector. However, economic series by sector are only available at the county level or for the major metropolitan areas. Population, a variable that is available at all levels of aggregation, could theoretically serve as a measure of market size, but it is more closely associated with residential consumption than commercial or industrial, which is also true of per capita income. The evaluation team believes that per capita income and population is likely to be correlated to employment if they are both included jointly in a statistical model. Population is available at the town level or census-track level from the American Community Survey data set.

References

[i] DNV GL and NMR March 31, 2015. *Top-down Modeling Methods Study—Final Report*. Prepared for the Massachusetts Electric and Gas Program Administrators and EEAC Consultants. .   http://ma-eeac.org/wordpress/wp-content/uploads/Top-down-Modeling-Methods-Study-Final-Report.pdf

[ii] Ferit Ucar, Ph.D., Lawrence Masland, Noel Stevens, and Monica Cohen. August, 2015. *The View from the Top: Top-Down Estimation of Program Savings Using Utility-Level Data in Massachusetts*. IEPEC.

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

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

vi. Demand Research, L. (2012). *Macro Consumption Metrics Pilot study Technical Memorandum – Preliminary Findings. Prepared for the California Public Utilities Commission.*

v. Horowitz, M. J. (2001). *Economic Indicators of Market Transformation: Energy Efficient Lighting and EPA's Green Lights. The Energy Journal, Vol. 22, No. 4.*

vi. Horowitz, M. J. (2004). *Electriciy Intensity in the Commerical Sector: Market and Public Program Effects.* The Energy Journal, Vol. 25, No. 2.

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

viii. Horowitz, M. J. (2010). *Measuring the savings from energy efficiency policies: a step beyond program evaluation.* Springer Science & Business Media B.V. 2010.

ix. Jacobsen, G. D., & Kotchen, M. J. (2009). *Are Building Codes Effective at Saving Energy? Evidence From Residential Billing Data in Florida.* Conference on Green Building, the Economy, and Public Policy.

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