Top Down Analyses of State Level Energy Savings

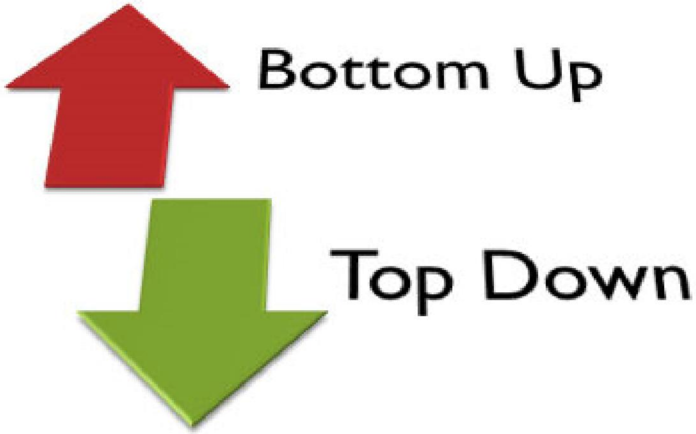
Noel Stevens, DNV GL
Chris Russell, NMR

2015 IEPEC Conference — Long Beach, California

Agenda

- **Definitions top-down and bottom-up modeling**
- **Motivations in pursuing top-down**
- **Expectations for top-down modeling**
- **Contributions of MA top-down research**

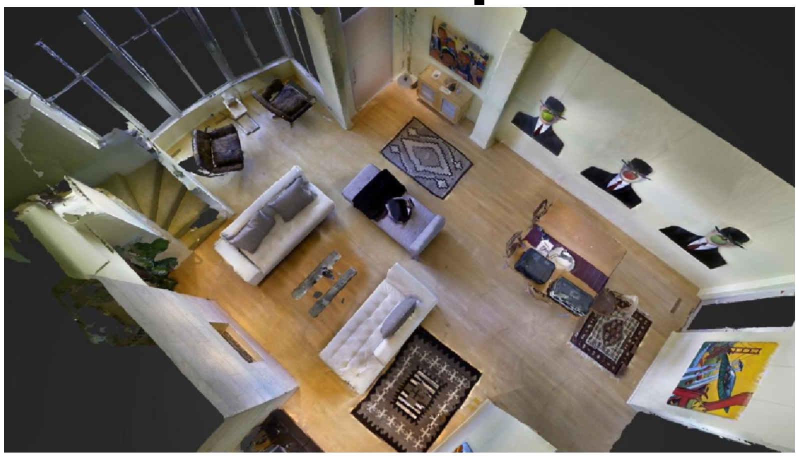
What are we talking about?



Goal of top-down modeling:

To isolate the effect of program activity from other natural changes and policy variables, in regional consumption over time.

Comparison of top-down and bottom-up?



Top-down modeling

Holistic approach – Estimates program impacts across all energy-efficiency programs or initiatives in a region

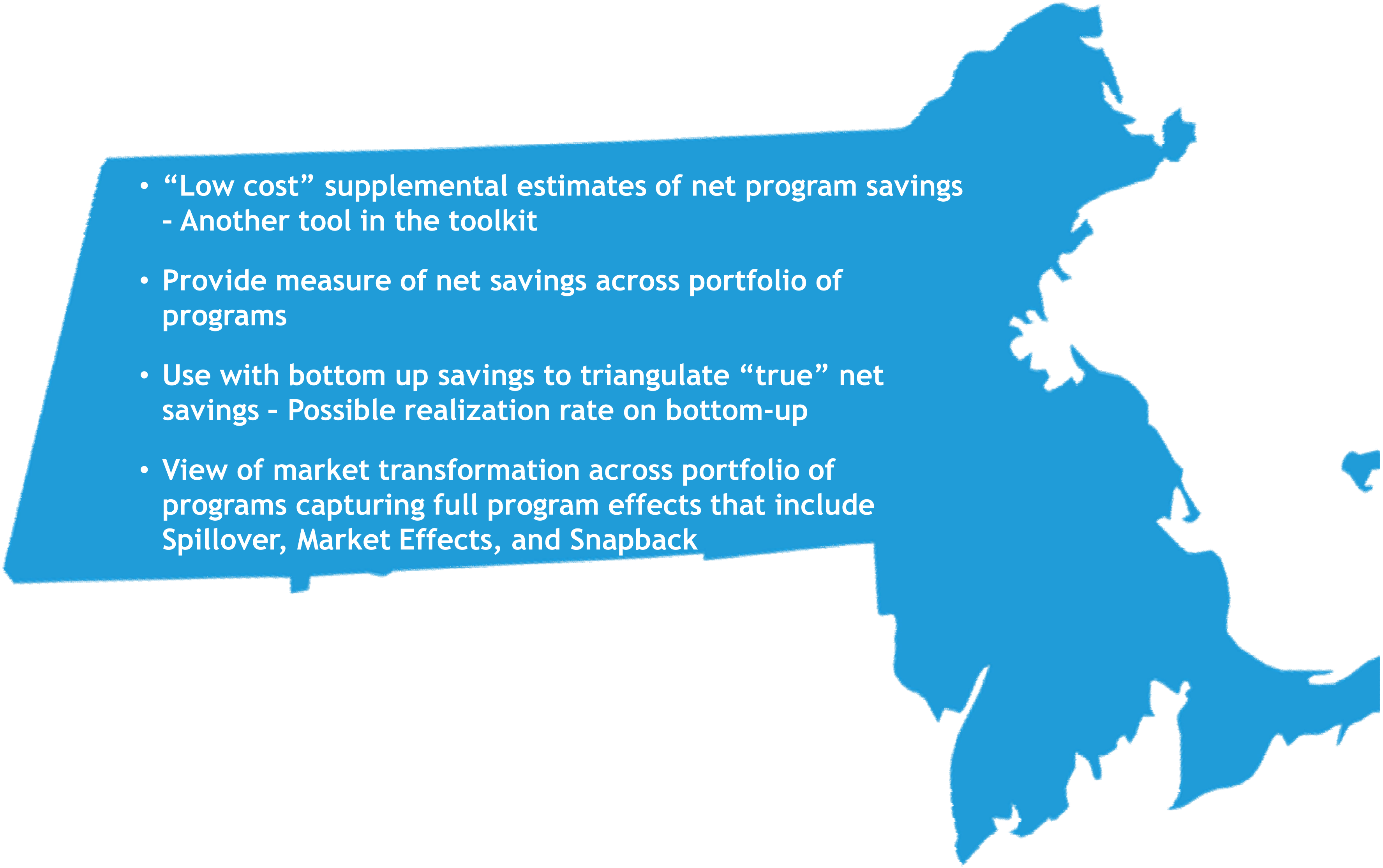
- Econometric (regression based) approach
- Changes to aggregate energy consumption
- Program activity is an explanatory variable
- Controls for economic factors at aggregate level (e.g. county, IOU territory)
- Ideally accounts for free-ridership, combined impact of programs, and market effects

Bottom-up modeling

Disaggregate approach – Measures impacts program by program within a utility territory;

- Add up changes from units to programs to portfolios to determine aggregate energy consumption change
- Program activity is an explanatory variable
- Economic factors at disaggregate level (e.g. account level)
- Separate free-ridership and market effects studies used to capture net savings

Motivations in pursuing top-down



- “Low cost” supplemental estimates of net program savings – Another tool in the toolkit
- Provide measure of net savings across portfolio of programs
- Use with bottom up savings to triangulate “true” net savings – Possible realization rate on bottom-up
- View of market transformation across portfolio of programs capturing full program effects that include Spillover, Market Effects, and Snapback

Setting expectations

What are we trying to do?

Construct regression model to isolate 2% of total consumption attributable to energy efficiency – This is difficult given modeling challenges in even the best of circumstances.

Setting expectations – What do we need to properly specify models

Econometric considerations for successful top-down models	Loughran and Kulick	Rivers and Jaccard	Horowitz 2004	Horowitz 2007	Parfomak and Lave	Horowitz 2010	Demand Research	Cadmus	Jocobsen and Kotchen
Use of Differences Dependent	√	√			√		√		
Use of Differences in Explanatory Variables									
Program Variable	√	√			√		√	√	
Other Explanatory Variables	√				√		√	√	
Account for heteroscedasticity									
Natural Log of Dependent Variable	√				√	√	√		
Natural Log of Program Variable	√				√	√	√		
Natural Log of Other Explanitory Variables					√	√	√	√	
Estimated Using Weighted Least Squares			√		√	√	√	√	
Lag of Depdent Variable		√	√						
Account for Fixed Effects									
Annual Fixed Effects	√		√		√		√	√	√
Geographic Fixed Effects	√		√		√		√	√	
Difference in Differences Approach				√		√			√
Allow for Differences in Types of Programmatic Activity			√				√		
Multiple Measures of Program Influence							√	√	
Account for Changes in Consumption Resulting from Building Code Changes							√		√
Account for Energy Prices									
Electricity Prices	√	√	√	√	√	√	√	√	
Substiture Fuel Prices	√	√	√	√	√	√	√	√	

Setting expectations – Criteria for success

Elements that increase signal

- **Diversity of program activity** – Programs have to vary over time (year over year) and across geography (towns, counties or states have different offerings)
- **Minimal effect of one area on another (cross-area spillover) -** Information and experience from one area influencing behavior in another area
- **Long enough time series to detect and isolate program impacts** – Research shows successful models have more than 10 years of program and consumption data
- **Account for the lag structure of program impacts** – Program expenditures 3 to 5 years ago may result in savings today through equipment survival and spillover

Elements that reduce noise in estimates

- **Consistent reporting of data -** Aggregate data compiled the same across geographies and may report the same phenomena
- **Consistent relationship between program activity metric and savings** – The influence of program variables and consumption must be consistent across units of observation

Setting expectations: What it can and cannot do?

Can do if successful

- Inexpensive estimates of net impacts (Data permitting)
- Combined effects of cumulative activity over programs and over time
- Net savings including spillover, market effects, and snapback
- Provide confidence intervals and precision levels for net energy savings estimates (WHY)

Cannot do

- Obtain savings estimates net of free riders only
- Separate free-ridership, spillover, and market effects estimates
- Isolate effect of a particular program and year
- Identify which groups of measures or customers are performing better, or worse

Top-down is an on-going research effort…

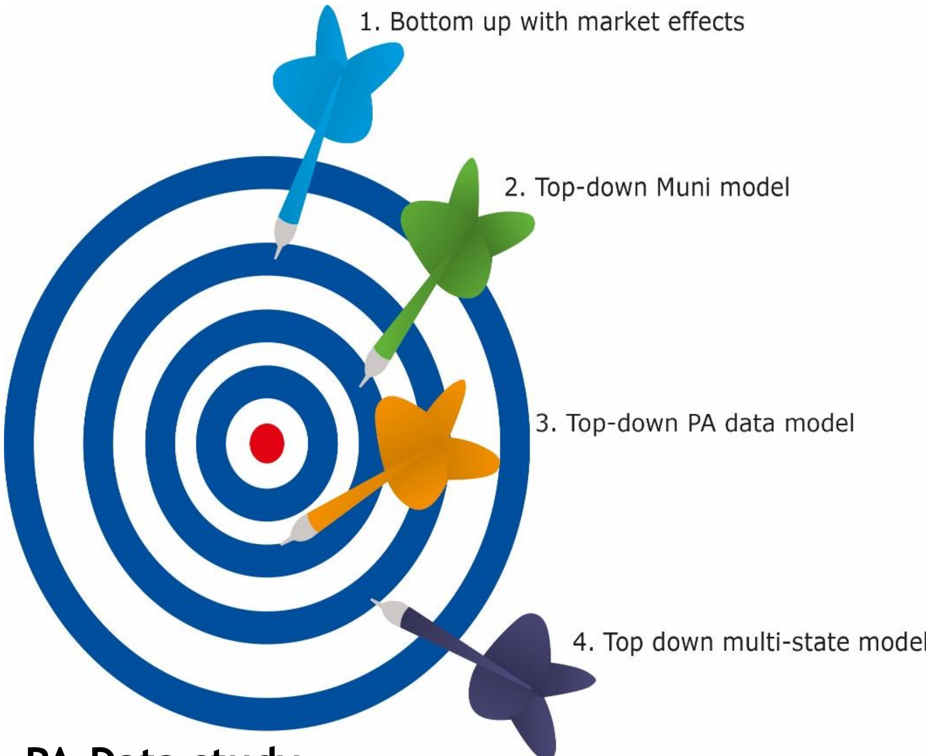
You get an answer **BUT** these are still estimates

Caution about regression results

- **Easily summarized and explained with R^2 and statistically significance**
- **Run a different model and you may get a different result**
- **With shorter time series, results may be very sensitive to data points included or not**
- **With longer time series the factors being estimated may not be constant**

- Results are highly dependent upon model specification and the availability of data
- Many familiar challenges apply - Really an extension of typical "billing analysis" techniques used in bottom-up evaluation apply
- Broader geography and timeframe imposes additional challenges

Top-down modeling: MA PA- 2014 pilot studies



PA-Muni

- Data aggregated at PA/municipal utility level
- Residential and Commercial & Industrial models
- Changes in total PA PA/municipal utility level consumption over 22 years
- Program activity measured as total program expenditures
- Use variables to control for economic conditions and weather
- 1 to 6 previous year's expenditures included to account for cumulative impacts
- Municipal utilities served a no-program comparison area

PA-Data study

- Used account level consumption and program data – Aggregated to town and county level
- Estimated separate models for large and small commercial, and industrial separately
- Normalized annual consumption as DV; Data series limited to just 3 years; and No comparison area
- Program activity measured as program expenditures and ex ante savings; lighting and non-lighting, upstream and downstream measured separately

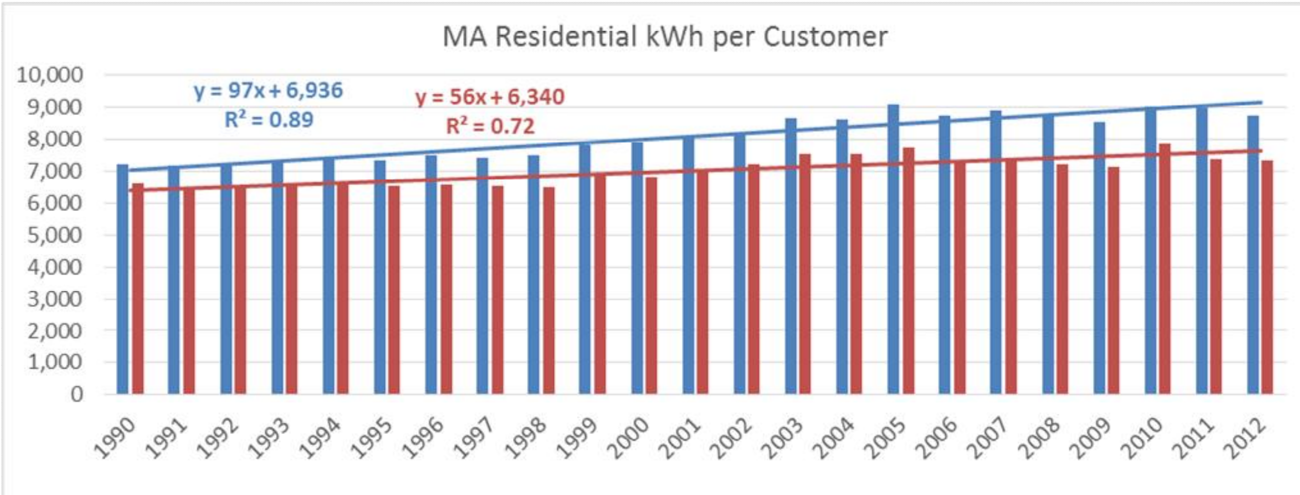
(This is a multi-year methods review study – we continue to explore and refine modeling approaches)

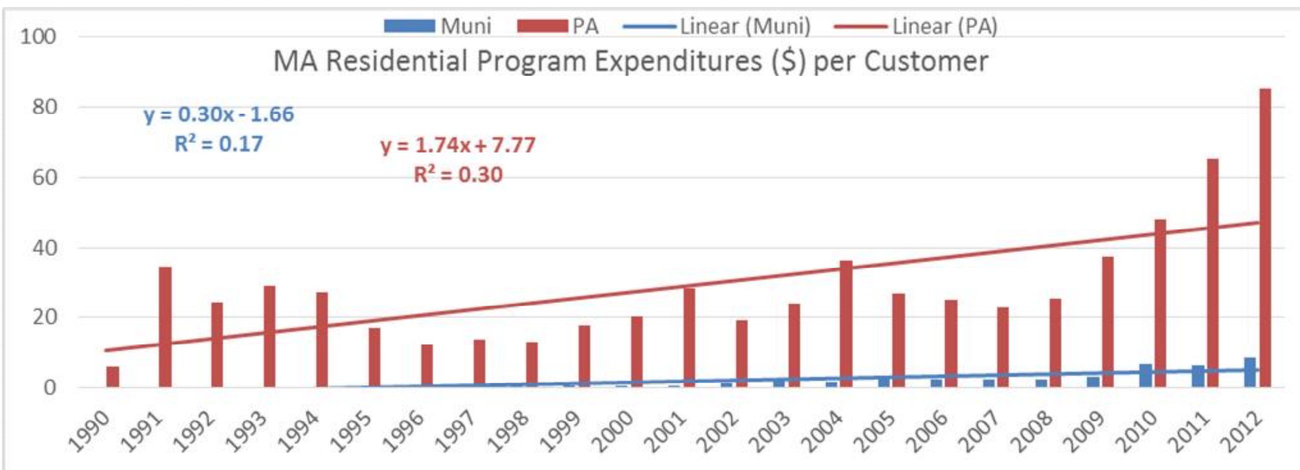
Pros and Cons of MA PA's - 2014 pilot studies

- **Both models seek to estimate net savings**
- **PA data model**
  - ✓ Uses differences in time period and geography to simulate net conditions
  - ✓ Use of account level data provides for investigation of program, measure, and customer type differences
- **PA-muni model**
  - ✓ Uses longer time series and activity in muni territories to simulate net conditions
  - ✓ Aggregate PA and muni level data limits investigation of program, measure, and customer type differences

Factor for successful models	PA-muni model	PA-data model
Diversity of program activity	*****	**
Minimal effect of one area on another	*	*
Consistent reporting of data	**	*****
Long enough time series to detect and isolate program impacts	*****	*
Account for the lag structure of program impacts	*****	*
Consistent relationship between program activity and savings	***	****
Data… data… data	*	*

PA/Muni model Motivation





PA/Muni Residential Model

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

$\log(\mathrm{EC}_{\mathrm{it}})$	Log of annual consumption per customer
$\log(P_{it})$	Log of electricity price in 2012 dollars
$\log(HDD_{it})$	Log of HDD
$\log(CDD_{it})$	Log of CDD
$\log(I_{it})$	Log of median household income
EH_{it}	% electricity as primary heating fuel
VAL_{it}	Median home value
NC_{it}	% new construction residential homes
SF_{it}	% single-family homes
$RENT_{it}$	% renters
EMP_{it}	Employment rate
EE_{it-j}	EE program dollars per residential customer
τ_t	Time trend variable

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
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

PA/Muni Residential Model

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	-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%

PA/Muni C&I Model

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

$\log(\mathrm{EC}_{\mathrm{it}})$	Log of annual consumption per employee
$\log(P_{it})$	Log of electricity price in 2012 dollars
$\log(HDD_{it})$	Log of HDD
$\log(CDD_{it})$	Log of CDD
$\log(EINC_{it})$	Log of mean annual employment income per employee
NC_{it}	Square footage of new C&I construction per employee
$\mathrm{NAICS}_{k,it}$	% of establishments in NAICS industry code
EMP_{it}	Employment rate
EE_{it-j}	C&I program expenditures per employee
τ_t	Time trend variable

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

PA/Muni C&I Model

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%

PA/Muni Model Discussion

Residential model

- Statistically significant estimates
- Estimates had wide confidence intervals and varied substantially between model specifications
- Realization rate on bottom up savings estimates ranged from 85% to 190% (point estimates)
- Further analysis of results showed results sensitive to the recession period

C&I model

- Fewer models statistically significant estimates
- Realization rate on bottom up savings estimates ranged from 65% to 101% (point estimate)
- Results sensitive to outliers and recession period

PA Data Model Specification (C&I only)

Access to account level information allowed for model separate models by sector

- Small Commercial;
- Large Commercial;
- All Commercial; and
- Industrial

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

Where each variable in Equation 2 is defined as follows:

- **Dependent Variable =** = Normalized (C&I) Annual Energy Consumption **(NAC)** $_{tsgf}$ in year (t), sector (s), within geographic region (g).
- **Scaling variable** = County-level models = GDP; Town level models = Population
- **Employment** $_{tsg}$ = Economic activity measured as the total employment per GDP or population
- **Program activity =** We considered two separate measures of programmatic activity separately:
  - **EE $ Program Expenditure** Vbl(s) $_{tsgf}$ = Program expenditures as reported in the PA program tracking data
  - **EE Program Energy Savings Vbl** $_{tsgf}$ = Ex Ante savings, as reported in the PA program tracking data
  - Program variables separated into Upstream / Downstream; Lighting /Non-Lighting
- **Fixed Effects**
  - *£ $_{sgf}$ = Region - Parameter for geographic fixed effects for county or town g in sector s, and fuel type f.
  - ¥ $_{tsf}$ = Year - Parameter for annual fixed effects for year t in sector s, and fuel type

PA Data Model Results

PA/data models

- **No models were statistically significant** – Data availability a primary obstacle to successful estimation
- **Segmentation of top-down models is possible given enough data** – It is possible to use top-down techniques to examine differences in program types and customer groups.
- **Effective top-down models require 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

Conclusions and takeaways from MA Pilots

- Top-down can provide a set ADDITIONAL ESTIMATES of net savings
  - Each approach has pros and cons
  - Each contributes to the overall picture of net impacts
- Top-down models face a range of data concerns that complicate estimation and can add substantially to costs
- Consistent reporting of consumption and program tracking data across geographies and time periods is a major obstacle
- It is challenging to properly define a model or set of models that truly isolate program impacts
  - Models can be sensitive to individual observations or particular time periods
  - Model specification is a work in progress – Variables needed to control for non-program variations are not necessarily available, or don't have enough data to separate effects
  - Just because you have a model with some control variables and a coefficient on program activity with nominally good precision does not mean it's a good estimate.
- PA data model allows for examination of program and measure level impacts, exploration of savings by customer segment, but requires account level data
- Muni model has enough observations and comparison area but results have wide confidence interval and are sensitive to model specification and observations.

Summary: What's a utility to do?

- Utilities and stakeholders interested in top-down should develop a data collection strategy that:
  - Accumulate the necessary data as it becomes available
  - Refines models over time
- Continue to support exploratory top down work
- Include plenty of tire-kicking in any top-down analysis
- Do not be determined to get "the answer" from top-down analysis
  - Use the results under alternative specifications and data restrictions along with bottom up results as informative

Contact us?

- MA Top-down research:

http://ma-eeac.org/wordpress/wp-content/uploads/Top-down-Modeling-Methods-Study-Final-Report.pdf

- **Noel Stevens**, Senior Consultant

  DNV GL - Energy

  Noel.Stevens@DNVGL.com

  Phone (781)418-5806

- **Chris Russell, Ph.D.,** Senior Quantitative Analyst

  NMR Group, Inc.

  crussell@nmrgroupinc.com

  Phone: 617-284-6230 ext. 11