Assessing Bottom-Up and Top-Down Approaches for Assessing DSM Programs and Efforts

Daniel M. Violette, Ph.D., Navigant Consulting, Inc., Boulder, CO;
Bill Provencher, Ph.D, Navigant Consulting, Inc., Madison, WI; and Iris Sulyma, BC Hydro, Burnaby, B.C.

Abstract

This paper examines two approaches for developing estimates of energy savings from Demand-Side Management (DSM) efforts. these approaches are characterized as: 1) bottom-up approaches and 2) top-down approaches. Both approaches address important questions, but with different data and different methods. There has been some controversy created by proponents of the top-down approaches in North America due to findings that have been judged as inconsistent with bottom-up approaches often used in evaluations conducted at the program level by utilities and regional program implementers in the United States and Canada. This paper presents examples of bottom-up and top-down evaluation approaches, focusing on the challenges inherent in implementing each approach. The paper assesses the strengths and weaknesses of each approach and the role they might play in future DSM evaluation efforts.

Introduction

There has been an ongoing debate regarding two different classes of methods for assessing energy savings resulting from Demand-Side Management (DSM) programs in North America. DSM programs are implemented by utilities (private or public) or by government authorities[1] (state or provincial). As part of their responsibility for implementing these DSM programs, they also take on responsibilities related to quality control and program tracking systems to provide initial estimates of energy savings. Final program-wide estimates of savings are developed working with an evaluator that is independent of the program implementor. Program implementors seeking an evaluation of their program with its unique characteristics are often forced into using program-specific micro-data in a bottom-up estimation approach. Top-down approaches[2] have been employed to assess DSM savings using macro-data across programs at the state or province level to see if a relationship between DSM activities represented by expenditures at a broader regional level can be associated with reduced energy use.

Both approaches are addressing questions that should be asked, but with different data and different methods. However, controversy has been created by proponents of the top-down approaches that claim this method produces estimates of energy efficiency (EE) savings that are inconsistent with the bottom-up approaches.; specifically, that the estimates produced by utilities and other program implementors overestimate savings. Proponents of the top-down approach view these methods as more accurate, and better able to address key issues in assessing the energy savings from DSM programs. This claim needs to be grounded in an assessment of the strengths and weaknesses in each approach. This assessment needs to include: 1) the issues addressed by each approach, 2) the role of judgment across both methods, and 3) the potential for bias in each method.

[1] Governmental entities play an important role in implementing DSM programs in a number of regions. Examples include the Ontario Power Authority (OPA) in Canada, the New York State Energy Research and Development Authority (NYSERDA), and federal agencies such as the Bonneville Power Administration (BPA) and the Tennessee Valley Authority (TVA) that serve multistate regions in the United States.

[2] The top-down approaches in this paper refer to the growing literature that have used econometric techniques to examining the relationship between DSM activity usually measured by DSM expenditures or expenditures per capita, and effect as measured by reductions in energy use or energy use per capita. One reviewer indicated that Europe may have a broader definition of Top Down approaches: “in Europe the most common top-down approach is that of energy efficiency (EE) indicators.”

There are many choices an evaluator has to make when evaluating any policy or program whether it is focused on, for example, energy efficiency, health care, or education. One purpose of evaluation is to provide information that can help make good decisions regarding investments[3] and activities designed to increase EE. One important question is whether current activities and investments in DSM are reaching the targets, i.e., providing the expected returns. Other important questions relate to the specific portfolio of DSM activities that have been selected and whether adjustments to this portfolio of activities are warrented based on evidence collected. Additionally, there may be ways to improve the implementation of specific DSM activities.

This paper is organized into three sections:

1) Overview and examples of bottom-up approaches with subsections on estimation issues and views on the application of these methods;
2) Overview and examples of top-down approaches with subsections on recent examples, estimation issues, the re-estimation of a recent study to illustrate these estimation issues; and views on the application of these methods.
3) Conclusions that focus on strenths and weakness of each method, the issues they are best designed to address, and the role of the two approaches going forward.

Overview and Example of Bottom-Up Evaluation Approaches

This section discusses the general bottom-up approach, presents an example and discusses key estimation issues. Views on this approach are presented at the end of this section.

Many evaluations of DSM efforts performed to date have focused on the efforts of a single program implementor (utility or agency) or a group of implementors within a region to assess the overall economics of their activities. For example, statewide evaluations that address the programs implemented by all the utilities within that state have become more common in the United States. These evaluations have generally been program focused (i.e., they take a specific program or set of programs and employ an approach where a sample of participants are taken in subsequent time periods). A measurement and verification (M&V) approach is developed for this sample of participants that gathers and verifies appropriate energy efficiency measure installation and operation, and examines other factors that impact the savings for that particular participant.[4] These in-field verified estimates of energy savings are then compared to the tracking system estimates to create a realization rate (i.e., how much of the expected savings for each site or participant group can be verified through the more detailed M&V studies on a sample of program participants). A realization rate of .90 would imply that 90% of the expected savings based on the program tracking system were able to be verified in the field by the M&V.

This realization rate evaluation approach, like any estimation method, has strengths and weakenesses. The micro-data approach used in this bottom-up approach provides advantages for assessing and improving DSM programs that other more aggregate approaches cannot provide. Over the past decade, these realization rate approaches have evolved to encompass a number of sophisticated analyses as part of the M&V effort.[5] Still, there are choices that are required by this

[3] The term investment in DSM is use broadly and includes any activity that might encourage energy efficiency from rebates to behavioral programs. They all need resources.

[4] This is a simplistic view of what has become a complex M&V process of DSM programs. These methods have been adapted for programs that are focused on programs that have specific sites as participants (e.g., commercial and industrial facilities) and programs that focus on trade ally activities to increase the sales of energy-efficient equipment. A detailed exposition of these methods is beyond this paper, but applications of these methods can be found in the evaluation of the EmPOWER programs – a statewide evaluation of Maryland's programs spanning five utilities. (EmPOWER, 2012).

[5] See the International Performance Measurement and Verification Protocol (IPMVP) for M&V practices. (www.evo-world.org)

method and the final estimates of energy savings not only incorporate technical M&V issues but behavioral issues such as program free ridership, spillover, and the potential for rebound.[6]

Simplified Example of a Bottom-Up Approach

A program targeted at large commercial buildings with substantial lighting requirements and air conditioning needs can serve as the basis for a simplified example of the bottom-up approach. For this example, 200 of the largest buildings in a utility service territory are targeted by the DSM program. Each of the 200 buildings is undergoing an energy assessment that identifies cost-effective investments in energy efficiency that may include:

1. Identifying investments in lighting retrofits and lighting designs to reduce the electricity used in lighting. Lighting cabs represent as much as 30% of a large building's electricity use.
2. Assessing major energy-consuming equipment such as boilers and AC, to determine if there are cost-effective investments in equipment, or changes in how the equipment is maintained and operated, or address issues with the ventilation system that would reduce energy use.
3. Examining the building's energy management and controls system (EMCS) to see if it is operating properly and if a new system would produce gains in efficiency.

Most energy assessments of existing commercial buildings find a wide range of cost-effective energy efficiency investments. The energy assessment documents these EE investment opportunities. A program implementation phase uses approved lighting, HVAC or controls contractors to implement these cost-effective investments.[7] As the implementation of these EE investments (or a subset of these investments) is undertaken, a tracking system is developed that documents the actions taken. Initial estimates of expected energy savings are based on the information available at the time of installation. An important part of program implementation is the development of a tracking sytem that records the activities undertaken as part of the program and, based on the information available, developes intial estimates of energy savings.

Estimation Issue - Gross Program Impacts.

A simplified bottom-up estimation approach would involve drawing a sample from the 200 buildings that participated in this DSM program. This sample may consist of 30 buildings that are then scheduled for an M&V assessment. The M&V effort examines the energy efficiency actions taken at the sampled buildings and through a set of more detailed assessments (often involving kW metering and logging of run-time hours) determines if the expected energy savings is in fact being achieved. The more detailed M&V process produces higher quality estimates based on more granular information than the estimates in the program tracking system. Using ratio estimation and the sample data, statistical estimates of the realization rate are developed for the entire group of buildings that participated in the program. If the M&V determines that only 90% of the expected energy savings are being realized in the field, then the realization rate is .90, and this is used to produce the estimate of gross savings for the program (i.e., total savings in the tracking system multiplied by .90).

This example showed a reasonably high realization rate that offers a level of confidence that the equipment installed in these buildings is saving energy and that these investments are cost

[6] Rebound occurs where participants take back some of the savings by increasing their energy use. For example, the installation of a more efficient air conditioner may result in a building or household setting their thermostats at a cooler temperature due to the lower cost of cooling.

[7] These investments may be paid for by the building owner with the free audit/assessment being the inducement to participate in the program or there may be subsidies for certain investments provided by the program implementer. The incentives for the building owner to undertake the identified cost-effective EE investments are constructed differently across different programs, and the process of selecting appropriately qualified contractors also varies across programs.

effective. Rebound or take-back can also be addressed in these bottom-up estimates if information on the pre- and post-operation of the building is available. If this is all that were involved in the bottom-up approach, then the method may not be subject to the high degree of controversy.

The information gathered by the M&V approach so far shows that the monies spent by the program have resulted in cost-effective energy savings. For some proponents of DSM, knowing that the program resulted in cost-effective investments in energy efficiency and that money has not been wasted is enough to justify the investment.[8] However, to make the case that DSM should be invested in as a resource that can be compared to other supply-side resource investments, additional factors need to be considered.

Estimation Issue - Net-to-Gross (NTG)

The resource value of a DSM program needs to take into account a baseline that represents what would have happened in the absence of the program. For example, if 20 out of the 200 buildings would have made the same energy efficiency investments in the absence of the program, then the savings from those participating buildings should not be counted as energy savings that were the result of the program. A number of factors are important in addressing net-to-gross:

1. **Free Riders.** This occurs when some of the participants in a program would have taken some or part of the actions promoted by the program even if the program had not existed.
2. **Spillover.** Spillover works in the opposite direction of free riders. There is strong evidence that some spillover energy savings and related benefits from DSM programs exist and may not be counted. For example, one of the building owners might be impressed by the savings attained through the program such that they install the same energy efficiency measures in another building outside of the target area of the program, or they pursue other energy efficiency investments at the participating location on their own because of the experience gained through the program. Similarly, other non-participating customers may observe the benefits attained in the buildings that did participate and decided to take action on their own.
3. **Rebound.** Lowering the cost of providing energy services (e.g., cool air or lighting) may result in the customer taking back some saving by increasing their use of a service.

Estimation Issue - Methods for Estimating NTG Factors

Estimating gross impacts has its complexities, but there are generally a set of supporting quantitative estimation methods based on building operation and thermodynamics. The three NTG factors cited above are more behavorial based, but approaches have been developed to assess free riders, rebound, and spillover impacts. Free rider estimation probably best illustrates the issue that most top-down proponents have with bottom-up approaches (i.e., how to estimate what would have occured among the program participants in the abcense of the program). There are several approaches to address this issue. Approaches that use a control group of comparable non-participants can address what would likely have happened among participants if the program had not been offered. In fact, this approach is often used for program types where a representative control group can be constructed at reasonable cost. A common approach is to conduct structured surveys of the participants and attempt to elicit views on what actions they would have taken if the program had not been offered. Like any statistical approach, these free rider estimation approaches produce range

[8] This view of gross savings being the key variable to look at when assessing the appropriateness of investments in energy efficiency implicitly takes a position on equity and cross-subsidies. If some of the owners of the 200 buildings would have installed the energy efficiency measures even if the program had not existed, then it is argued that they are being subsidized by the program since they would have made the energy efficiency investment anyway and it cannot really be counted as a program-induced impact. However, if “who pays” is not a policy issue, then efficient investments in energy savings are being made and it is argued that only gross savings estimates need to be estimated (See Titus and Michals, 2008).

estimates (i.e., constructed confidence intervals) for the free rider estimate. Survey methods and questions have been developed using the participants themselves, discussions with trade allies that work with participants, observations of the contractors that installed the equipment, and other potentially useful sources of data. The result is a range estimate of free riders that is based on the weight of the evidence from these survey activities.[9]

Views on Bottom-Up Approaches

Proponents of bottom-up approaches tend to believe you can obtain reasonable estimates of free riders (as well as spillover and rebound) that can provide a preponderance of evidence that the investments in energy savings due to the program are cost effective even taking into account NTG effects

The proponents of the top-down approach are skeptical about whether meaningful estimates of the NTG factors can be estimated. For example, in an application of a top-down method applied to Canada, Rivers and Jaccard (2011) state that ”micro-data ... program level evaluations must make difficult judgments about important factors that are key to program effectiveness: the free ridership rate, the spillover rate, and the rebound effect.” Rivers and Jaccard used a top-down approach based on aggregate cross-sectional data for10 provinces in Canada across 16 years. Their study examined the relationship between DSM expenditures per capita[10] and sales of electricity per capita. Using this aggregate top-down approach, Rivers and Jaccard conclude that ”Demand-side management expenditures have had a minimal impact on electricity demand” with estimated savings being considerably less than those estimated by utilities using program-level estimation approaches.

Several comments are important here and will be discussed in greater detail in the example of top-down approaches below.

1. Bottom-up approaches provided information on specific program and portfolio investment issues that cannot be addressed by top-down approaches.
2. Developing a data set that contains appropriate and accurate information at aggregate levels for the top-down approach as well as enough data points to provide a robust estimation of the DSM expenditure/electricity sales relationship poses challanges.
3. Estimating NTG factors in bottom-up models requires judgments and assumptions, but the top-down approaches also require judgments and assumptions that are no less broad.
4. Recent work using top-down approaches has now started to confirm the energy savings estimates by utilities rather than produce the conflicting results of no energy savings.

---

[9] Point estimates of free riders and other NTG factors are often needed for DSM program implementers that receive incentives for the program savings achieved, or for estimating GHG reductions. In these cases, a more specific survey protocol is usually agreed upon and used to produce this estimate. Most NTG survey protocols have four steps – 1) establish existence of the effect, i.e., is there evidence that the effect exits at all; 2) once existence has been established, the lower and upper bounds of the range of the effect are addressed by the survey; 3) after the lower and upper bounds of the range have been established, the survey then asks questions about where within this range might the free rider number be most likely to fall; and 4) most analyses of free riders try to apply other data (e.g., data from both participants, non-participants, and trade allies to triangulate and provide additional validation of the estimate. This process does contain uncertainty, but the real question is whether it produces information that can be made to make good decisions regarding investments in DSM as a resource. (See S. Schare, 2007)

[10] Obtaining good data on expenditures on energy efficiency activities can be difficult whether by utility, province (in Canada), state (in United States) or other cross-sectional units (e.g., an attempt to examine energy use using cross-sectional data at the county level is being developed in California under the guidance of the California Public Utilities Commission (CPUC). Data on DSM expenditures can be accounted for differently by different program implementers (e.g., expenditures on market efforts that span different objectives), and most DSM data include expenditures peak demand/load-shifting programs as well as on energy conservation. Rivers and Jaccard (2011) state that load management expenditures amounted to less than 25% of the total (p. 113), but U.S. data shows that for some utilities, in some years

5. The Rivers and Jaccard study has attracted attention in North America with its finding of no relationship between DSM expenditures and electricity sales using utility and province-wide data. A re-estimation of this model by Provencher, Violette, and Sulyma (2012) used the same data set, but less restrictive assumptions, which produced estimates of DSM energy savings and cost-effectiveness (costs per kWh saved) in line with utility-based, program-level evaluation approaches.

   The comments above preview the discussion on top-down approaches presented below.

Overview and Example of Top-Down Approaches

This section presents the basics of the top-down approaches followed by an example. Views on these methods are presented with a focus on estimation issues. A recent top-down model is re-estimated and updated to address several of the estimation issues identified and the results of recent top-down applications are updated.

Top-down approaches have attempted to address the credibility of bottom-up estimates by using econometric models to relate top-down aggregate DSM expenditures for an energy provider (utility or province) to aggregate energy use by the provider's customers (see Rivers and Jaccard, 2011 for a review of applications). The logic of this approach is that by modeling aggregate energy use as a function of a number of observable variables, including current and past DSM expenditures, one can identify the incremental effect of DSM on aggregate energy use. This approach allows the analyst to obtain an estimate of average cost-effectiveness of DSM while bypassing utility estimates of program-level energy savings, and, by extension, the conceptual, practical, and statistical issues that may attend these estimates.

Example of a Top-Down Approach

The top-down approaches generally estimate the effect of DSM expenditures on electric consumption at the aggregate level using an econometric model. Rivers and Jaccard (2011) use data at the provice level for 10 Canadian provinces across 16 years (1990 to 2005). Loughran and Kulick (2004) use panel data at the utility level for the period 1990 to 2002. Two other studies, one by Aufhammer, Blumstein, and Fowlie (2007) and one by Arimura, Li, Newell, and Palmer, use similar data, but for time periods 1990-2006, and 1992-2006, respectively.

The recent Rivers and Jaccard study has generated some controversy in Canada and the authors of this paper recently reworked this analysis using the same data. As a result, the Rivers and Jaccard (2011) is used to illustrate the basic top-down approach. The top-down econometric model involved the following key features:

1. DSM expenditures were obtained from various sources, including structured surveys of utilities. It was not possible to distinguish between DSM expenditures for load-shifting (e.g., AC load control programs which can require sizeable expenditures) vs. energy efficiency.
2. The sample size was 160 province-years, including 16 observations for Alberta, which had no DSM expenditures during the study period.
3. The dependent variable is the natural log of energy use per capita for a province in year t.
4. The model is a partial adjustment model in which the lag of the dependent variable (lag of natural log MWh per capita) appears on the right-hand side of the equation.[11]
5. The model accounts for province-level fixed effects.

---

load management programs can represent nearly100% of the DSM expenditures and, thus, would not be expected to produce energy savings.

[11] The rationale for including the lagged dependent variable as a regressor is that energy users take time to adjust to new economic circumstances, and this is revealed in this specification as a relatively simple reduced form model as inertia in the adjustment process.

The advantage of this approach, as summarized by Rivers and Jaccard, is that aggregation avoids many of the statistical issues that arise in evaluations of the program level. These statistical issues arise due to the difficulty of controlling free ridership, spillover, and rebound effects.

The estimated model uses per capita energy use in province *k* in year *t* regressed against letting denote per capita expenditure on DSM, and letting X_{kt} denote a vector of other regressors, the model takes the general form (simplified write out of the equation where “LN” denotes the natural log)[12]:

LN Per Capita Energy Use = B_1 (DSM expenditures per capita) + B_2 (LN energy use per capita lagged one time period) + B_i (other X_i independent variables such as electricity price) + Intercept specific to province k + error term$_{it}$

[The B_i’s are regression coefficients in the equation above.]

Allowing the value of intercept to vary across each province makes this a fixed effects model. A fixed effects model allows for factors that vary across provinces, but are constant over the analysis time period, to be addressed in the separately estimated intercept for each province. This might, for example, address the effect of the mix of industrial, agricultural, and urban energy uses on per capita energy use in each province. However, the longer the time frame being addressed, the less likely that a unique intercept for each province appropriately controls for these province-specific factors. Annual fixed effects may also be necessary. Annual fixed effects have been used in other top-down econometric models.

The key parameter being estimated is B_1, the coefficient on the DSM expenditures per capita. If DSM expenditures influence energy use, then, this term is expected to be negetive and significant. Among the factors that influence total electricty consumption, the effect of DSM expenditures may be small relative to other factors. As a result, for this relationship to be accurately estimated, a large data set with adequete variability can be required for these top-down models to parse out this effect from other factors. In this case of Rivers and Jaccard, there were 160 observations available to estimate the regression equation.

Rivers and Jaccard estimate three versions of the model, with the versions varying by the treatment of the equation error. The first version assumes errors are uncorrelated over time, and the other two versions attempt to correct for possible bias arising in the event of serial correlation in the errors. All the models generate the same qualitative result with respect to B_1, the effect of DSM expenditures on energy use. In each case, the estimated effect of DSM expenditures on energy use is nonsignificant and generally produces positive rather thannegative coefficients.

Views on Top-Down Approaches

The regression model estimated by Rivers and Jaccard above is a version of the general top-down approach with Energy Use regressed against DSM expenditures and other variables. Econometric models require judgments to be made regarding the specification of the model and underlying assumptions related to relationships between variables and assumptions about the error term. Potential issues in regression models include:

1. Endogeneity	One of the regressors is causally related to other indepenent regressors in the model.
2. Autocorrelation	The error terms for different observations are correlated.
3. Heteroskedasticity	The variance of the error term is not constant across observations.

[12] A more technical representation of this regression equation can be found in Provencher et al. (2012).

4.	Multicollinearity	Some regressors are highly correlated with each other.
5.	Specification error	Relevent explanatory variables are omitted from the equation.
6.	Inclusion of irrelevant variables	If correlated with included key variables, the estimated variance on the coefficients of these variables will be biased upwards and statistifcal significance biased downwards.
7.	Errors in variables	This occurrs when there are errors in measuring the independent variables (e.g., the development of DSM expenditure data for a unit of time and geographic area that appropriately correlates to the enegy use data for that area).

Processes and procedures have been developed to address these and other issues in the econometrics literature. Still, judgment is required and often alternative sets of reasonable assumptions are used to test the robustness of the models being estimated. These sensitivity analyses are common in applied econometrics and may produce different findings. Peter Kennedy (2003) discusses these basic model assumptions and threats to their validity in a concise chapter in his "Guide to Econometrics."[13] As a result, these top-down econometric models may require broad judgements to be made that can be as impactful as are assumptions made in the bottom-up approaches discussed above. There are a wide number of issues that need to be addressed to credibly apply top-down econometric methods to estimate the relationship between DSM efforts (expenditures on DSM is the ususal proxy for effort in a year) and changes in energy use. Two are discussed below:

<u>Top-Down Issue One</u> – Errors in variables.

The example of errors-in-variable of most relevance to top-down models may be related to the estimation of DSM expenditures. The data on DSM expenditures, both in Canada and in the U.S., generally include expenditures on all demand-side activities including both load-shifting expenditures and energy efficiency expenditures. Rivers and Jaccard indicate that the subset of utilities that were able to break out the data into load management and energy efficiency suggests that 'load management expenditures amounted to less than 25 percent of the total." This led to their conclusion that the "error in our estimates should not be too severe." However, this errors in variables problem has an unknown impact on the results. If there are systematic relationships with high expenditures on load management being correlated with utilities/provinces that have low expenditures on energy efficiency, then the influence of this errors in variables problem could be more pronounced.

Other regions are working on the application of these top-down approaches to smaller regions. For example, California is considering this option to assess investments in energy efficiency. With only three IOUs in California, there are not enough cross-sectional data available to estimate the top-down models that use variation in DSM expenditures across entities (utilities, provinces, or states). To create a data set that would allow for a top-down approach to be used, alternative areas are being given consideration to form the cross sections (e.g., one alternative is the use of counties as the geographic unit). While economic data and energy use data may be available from advance metering, appropriately allocating the DSM effort in terms of expenditures or other measures effort will be difficult.[14]

---

[13] Kennedy (2003) presents the key underlying assumptions common to regression models in Section 3.2 "The Five Assumptions" with a listing of potential violations of these assumptions and in Table 3.1 which summarizes these assumptions and implications.

[14] The implications of the errors-in-variables problem are well set out in Section 9.3 in Kennedy (2003). Kennedy mentions that a key drawback to econometrics is often the quality of the data as reflected in incorrectly measured variables (p. 160). Measurement errors may also exist in bottom-up methods, but the micro-data methods on which the bottom-up approaches are based, allow for examination of key data in program tracking systems through M&V methods.

<u>Top-Down Issue Two</u> – Temporal effect of DSM expenditures on energy use.

DSM Expenditures in one year may or may not have an impact on energy use in that same year and, regardless, the effects of DSM expenditures in one year will influence energy use in a number of subsequent years (e.g., possible 3, 5 or even 10 years out in time). Energy savings may begin during the year the DSM expenditure is made (for a residential AC equipment retrofit), beginning one year after the DSM expenditure (in the case of basic actions taken as the result of an energy audit), or possibly beginning two or more years after some of the expenditures are made (in the case of large commercial design projects or industrial process changes). As a result, the econometric models require assumptions about the structure of the temporal relationship between DSM expenditures (and other efforts) for the top-down approaches to estimate accurate relationships. Miss-specification of the temporal relationship in the model can have large effects on the estimates. The difficulty of accurately capturing the temporal relationship in the econometric model results may result in the need for complex lag structures, which requires larger data sets to produce reliable estimates. The re-estimation of the top-down model developed by Rivers and Jaccard (2011) in Provencher et al. (2012) shows the significance of the temporal assumptions used in these top-down models.

Re-Estimating a Top-Down Model and Update on Results

This section is based on Provencher et al. (2012), which examined and re-estimated the model developed in Rivers and Jaccard (2011) using alternative and updated assumptions from recent literature, and examined three other recent papers.

The three other recent papers that examine the effect of DSM expenditures on electricity consumption at the aggregate level are:

1. Loughran and Kulick (2004) examined utility-level panel data for the period 1989-1999.
2. Aufhammer, Blumstein, and Fowlie (2007) provides amended estimates of DSM savings and cost-effectiveness and estimating confidence intervals on these statistics, which Loughran and Kulick did not do.
3. Most recently, Arimura, Li, Newell, and Palmer (2011) extended the utility-level panel data of Loughran and Kulick to cover the period 1992-2006.

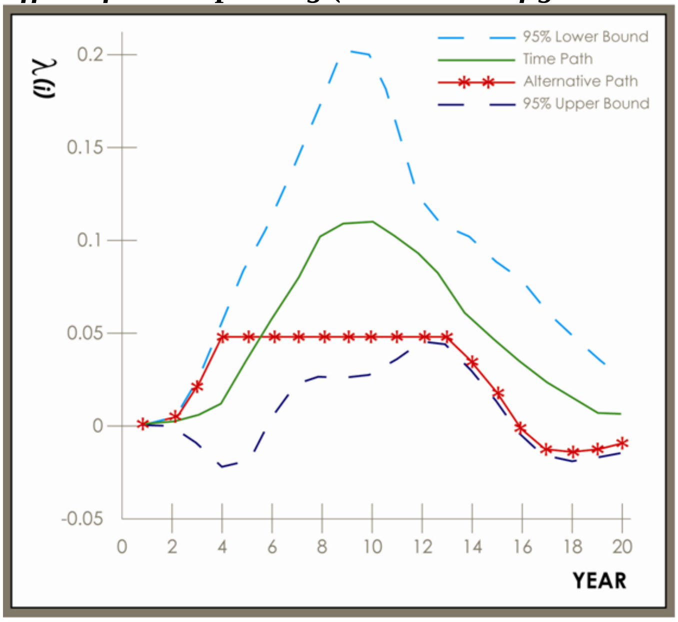
The results of both the Loughran and Kulick analysis (as amended by Aufhammer et al.) and the Arimura et al. analysis find that, using aggregate level panel data, it is **<u>not</u>** possible to conclude with high confidence that the savings and cost-effectiveness ($/kWh saved) of DSM programs differs from the average estimated by utilities. A fundamental difference between the models of the three studies refered to above and the Rivers and Jaccard study that has been highly visible in Canada is the treatment of DSM dynamics.

Focusing in on the most recent study, the dataset used in Arimura et al. includes 3,326 observations (utility-years) for 307 utilities. The DSM variable enters this model in a complicated 3-parameter nonlinear form. Importantly, this specification incorporates in year *t* the DSM expenditure per capita by the utility from all previous sample years. Arimura et al. estimate a primary model and they conducted a number of robustness tests that essentially support the qualitative conclusions from their primary modeling results. In general results were as follows:

1. DSM expenditures generate a nonlinear dynamic effect on energy savings. Figure 1 shows that the average savings effect of DSM expenditures peaks about 9 years after the expenditure. The 95% confidence bounds on this effect indicate that there is considerable uncertainty about when the peak actually occurs, but it is reasonable to conclude that on average the savings effect of a dollar spent in year *t* peaks in year *t+4* or later.

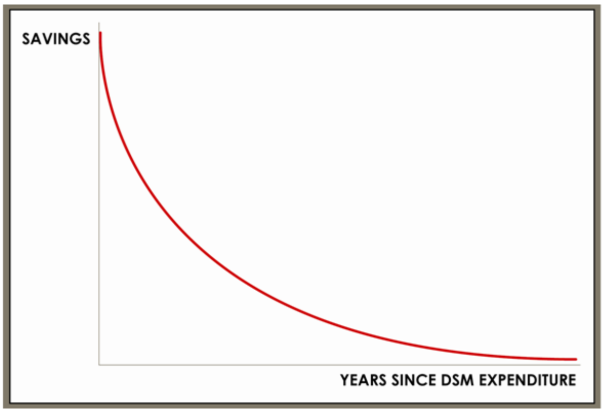
2. Using a real discount rate of 3%, program cost-effectiveness is about $0.042/kWh, with standard error about $0.025/kWh. This places the estimated cost-effectiveness of DSM well within the range of the average estimated by utilities.

***Figure 1*. Long-term Effect of DSM Spending (Note: This is figure 3 in Arimura et al.)**



The temporal flexibility shown in the Arimura et al. model above is absent in the Rivers and Jaccard model formulation. The Rivers and Jaccard model imposes a strong restriction on the structural relationship between DSM expenditures and energy use over time. The energy savings resulting from a dollar of DSM spent in year t decays exponentially over time, as shown in Figure 2.[15] This structural assumption was likely motivated, at least in part, by the relatively small sample size, but imposes a strong assumption about the relationship between DSM expenditures and energy use over time.

***Figure 2*. Exponential Decay in Energy Savings from DSM expenditures in Year 1 as assumed in the Rivers and Jaccard Model**



[15] This temporal relationship is based on the formulation of model as a partial adjustment model where the lag of the dependent variable (energy use per capita) appears on the right hand side of the equation. Since the model dynamics arise only through the lagged dependent variable, this rate of decay applies to all variables in the model (e.g., changes in price, weather variables and other independent variables included in the equation).

There is a clear difference in the assumed temporal effects between the Arimura et al. model and the Rivers and Jaccard model. This structural assumption has a strong influence on the results of the two models.

Review and Re-Estimation of the Rivers-Jaccard Model

Three features distinguish the Rivers and Jaccard analysis from the other three top-down studies reviewed above. The first is a much smaller data set, arising mostly due to the fact that the latter two analyses involve panel data at the utility level in the U.S., whereas the Rivers-Jaccard analysis uses panel data at the provincial level for Canada. The second is the use of annual dummy variables to control for annual fixed effects – Rivers and Jaccard did not use annual fixed effects where the other studies do. The third is the treatment of the dynamic effect of DSM expenditures on energy use. The Rivers and Jaccard model constrains the effect of DSM expenditures to diminish exponentially over time; the other analyses allow the effect of DSM expenditures to increase in the years following the expenditure.

To examine the impact of these latter two constraints imposed in the Rivers and Jaccard model, Provencher et al. used the same data[16] to estimate two models. The two models differ in the specification of the DSM variable. Model 1 uses the natural log of DSM expenditures as the DSM variable of interest (following Loughran and Kulick, and the update by Aufhammer et al.). Model 2 uses DSM per capita, the DSM variable used by Rivers and Jaccard as well as Arimura et al. In both models, the specification includes:

1. The dependent variable is the difference between natural log MWh at time t and natural log MWh at time t-1. This accounts for province-specific fixed effects.
2. The set of relevant DSM variables is composed of the current variable, the lagged variable, and the twice-lagged variable.
3. The model includes annual dummy variables to account for annual fixed effects.

The re-estimation and construction of this competing model using the same data indicates that the results obtained by Rivers and Jaccard are likely strongly influenced by two important features of their model specification:

<u>First</u> - The omission of annual fixed effects, which have been used in other “top-down” econometric models that do not rely on utility estimates of program savings.

<u>Second</u> – The use of the partial adjustment model formulation that effectively enforces the result. In the long run, the savings effect of contemporaneous DSM expenditures on future energy use is constrained to diminish exponentially over time.

Using the data from Rivers and Jaccard, estimation of two competing models produces estimates of DSM cost-effectiveness of $.035/kWh for one model and $0.047/kWh for the other. Both are strikingly similar to estimates obtained in the other recent econometric analyses of the cost-effectiveness of DSM, and consistent with the estimates obtained by BC Hydro,[17] as reported by Rivers and Jaccard. Moreover, because of the small number of observations, confidence intervals are large and so it is not possible to reject at a reasonable confidence level the typical estimates of cost-effectiveness derived by utilities.

---

[16] The data used in this review was graciously supplied by Nic Rivers. Using these data, Provencher et al. were able to replicate the original Rivers and Jaccard model prior to exploring the implications of model constraints.

[17] The BC Hydro reference is made as the Rivers and Jaccard model received attention in that province in terms of assessing the cost-effectiveness of DSM investments.

Conclusions – Bottom-up and Top-down Approaches

Bottom-up and top-down approaches both have their strengths and weaknesses. The view that top-down approaches avoid making broad-based judgments in the development of the model is not as appropriate as it first seems. Bottom-up approaches are likely the best method for developing estimates of gross impacts at the program level, but developing net-to-gross factors that address free riders, spillover, and rebound do require methods that are subject to judgment and uncertainty. However, top-down approaches using econometric approaches are also built upon judgment and uncertain assumptions, and have their own challenges. Top-down approaches do have the advantage of directly estimating net impacts and avoiding some of the issues inherent in the NTG process used by bottom-up approaches.

Top-down approaches are only useful for estimating the average effect across a number of utilities, provinces or other cross-sectional units. It is very difficult for these models to address the relative performance of one utility's programs versus another utility's programs, to offer suggestions on how to improve individual DSM programs, or the composition of a portfolio of interrelated DSM activities to help hit higher energy savings goals. The aggregate data used in top-down models produce aggregate results for a cross-section of entities engaged in DSM. The top-down models can potentially address whether, for example, there has been energy savings in Canada due to DSM expenditures, and whether the investments have been cost-effective. However, only average changes in energy savings from incremental DSM expenditures across the sample can be estimated and the effects for specific utilities, programs, and sectors are not produced by the aggregate top-down models. [18]

If carefully done, bottom-up models can provide program- and sector-level data at the utility/ province level. The development of methods to address NTG factors can allow for a preponderance of the evidence in the approach to producing reliable estimates of DSM effects and can serve the purpose of providing insights needed to make good investment decisions in energy efficiency.

Top-down models have to address issues concerning data quality, specifically the potential for measurement error in the DSM variable of interest (DSM expenditures or other measurement of DSM efforts). Top-down models also have to address the potentially complex issue of how the temporal relationships between DSM expenditures are made and when the resulting energy savings occurs. The models that have most robustly addressed the temporal issue have required large numbers of cross-sectional units and total observations.

Researchers that are looking to apply top-down approaches for assessing the effect of DSM investments at an aggregate level need to examine the data platform on which the model is to be estimated and make sure it is appropriate for the application of these econometric methods. If the number of needed observations are not available, and the data are not of sufficient quality (measurement accuracy and relevant explanatory variables), estimating a top-down econometric model will likely show no relationship between the DSM effort variable and energy savings. The large confidence intervals likely to be produced by these efforts (and have been characteristic of top-down approaches in general) are unlikely to support a broad and important general conclusion that "demand-side management expenditures have had a minimal impact on electricity demand"[19], with estimated savings being considerably less than those estimated by utilities using program-level estimation approaches. The top-down studies do list a number of potential issues in the estimation of

[18] If the number of observations and cross-sections are large enough, average effects across a subset of utilities might be able to be obtained.

[19] Rivers and Jaccard, p. 112, second paragraph. A reading of the Rivers and Jaccard study illustrates that they recognize a number of issues endogeniety of the price variables in the model, and potential endogeneity of the DSM variable, i.e., there may be more DSM in regions with robust growth in electricity demand.

the models including endogeniety, the hard to specify temporal relationship between DSM effort and savings, errors in variables, and potentially confounding effects from regulations, standards, and policies that may also influence energy use among other issues. These issues are not addressed within the model, but are judged by the study authors, to not significantly affect the results of the model. These are broad judgments and may or may not be correct. As a result, the findings of top-down models are based on the preponderance of the evidence as judged by the developers of the model in question. This is not unlike what is done in bottom-up DSM impact analyses.

In conclusion, both approaches have strengths and weaknesses, and they address different questions – average effects across a number of cross-sectional entities implementing DSM versus utility portfolio, sector, and program level analyses, which are the focus of bottom-up methods. Both methods are likely to be needed to make good decisions regarding investments in DSM.

References

Arimura, T.H., S. Li, R.G. Newell, and K. Palmer, 2011. "Cost-Effectiveness of Electricity Energy Efficiency Programs," National Bureau of Economic Research, NBER Working Paper No. 17556

Auffhammer, M., C. Blumstein, and M. Fowlie, 2007. "Demand-Side Management and Energy Efficiency Revisited," Center for the Study of Energy Markets, University of California Energy Institute, March.

EmPOWER Maryland, 2012. "Verification of Reported Energy and Peak Savings from the 2011 EmPOWER Maryland Energy Efficiency Programs" by Cory Welch et al., Navigant Consulting, for BG&E, PEPCO (and 3 other utilities) April.

Kennedy, Peter 2003. A Guide to Econometrics, MIT Press, 5th Edition.

Loughran, D.S., and J. Kulick, 2004. "Demand-Side Management and Energy Efficiency in the United States," *Energy Journal* 25(1): 19-43.

Rivers, N., and M. Jaccard, 2011. "Electric Utility Demand Side Management in Canada". *The Energy Journal* 32(4), pp. 93-116.

Violette, D and B. Provencher. 2012. "Review of a Top Down Evaluation Study: Rivers & Jaccard (2011)." Prepared for BC Hydro, Navigant Consulting, Inc., April.

Schare, Stuart 2007. "Advancing the "Science" of Free Ridership Estimation: An Evolution of the Self-Report Method for New York Energy $martSM Programs," Proceedings of the Association of Energy Services Professionals (AESP) 17th National Energy Services Conference, Las Vegas, NV, 2007.

Titus, E. and J. Michals 2008. "Debating Net Versus Gross Impacts in the Northeast: Policy and Program Perspectives," Northeast Energy Efficiency Partnerships, ACEEE Summer Study on Energy Efficiency in Buildings, Asilomar, CA.