A Meta-Analysis of Net to Gross Estimates in California

Jennifer Fagan, Itron, Inc.
Mike Messenger, Itron, Inc.
Mike Rufo, Itron, Inc.
Peter Lai, Energy Division: California Public Utility Commission

ABSTRACT

The objective of this paper is to provide an overview of estimates of the proportion of free riders participating in utility energy efficiency programs in California since 2001. The paper summarizes the results of a comprehensive review of recent evaluation studies focused on developing an updated set of ex ante net to gross ratios for program planning and other regulatory applications. Comparisons are made with NTG ratios estimated from previous literature reviews in the 1980s and early 1990s. The analysis includes a review of the pros and cons of different net estimation methods for the purpose of estimating free riders only and those methods used to estimate net market effects that include participant and non-participant effects. This results show that a wide range of NTG estimates can be estimated from the same baseline data depending on NTG definitions, analysis methods, and whether the results are intended to reflect the immediate past or a forecast of the near-term future. The results also highlight the need for jurisdictions to cooperate in the collection of sales and market share data for efficiency products such as CFLs to expand available data and reduce evaluation costs, as well as and the need for oversight agencies to plan for potential market effects of programs operated over along periods of time. Finally we present some lessons learned for those jurisdictions considering the issue of how and when to use estimates of net savings to adjust future program savings estimates.

Introduction

This paper presents a comparison of different methods used to estimate net savings for a variety of different energy efficiency program design strategies operated in California over the last decade. This paper also presents the results of a recent systematic review of net to gross ratios derived from empirical evaluation studies, primarily of California's program years 2004-2005, but also including studies of 2002-2003 programs (Itron 2008). This analysis was made possible by the long-term commitment of the California Public Utility Commission to systematically collect the results of program evaluations and support cross-study analysis of these results to support program planning and other ex ante[1] regulatory applications. In addition to the studies' findings, we summarize lessons learned in applying this information to estimate savings from future programs.

Overview of 2008 DEER NTG Update Process

California recently updated many of the values in its Database for Energy Efficiency Resources (DEER), including the Net-to-Gross (NTG) ratios assigned to specific types of program designs and energy efficiency measures. This was the first time the DEER project had undertaken this type of effort. California's investor-owned utilities are expected to use the NTG ratios in the DEER database when developing ex-ante savings estimates of savings for future energy efficiency plan filings. (NEED

[1] Ex-ante, from the Latin for "beforehand" are *forecasted* values.

citation for two CPUC decisions: the 2009-2011 directive and the 2006-2007 February Petition to Modify decision)

Our team was responsible for updating the NTG ratio values in the DEER, since these values had not been updated since 2001. In January of 2008, the California Public Utilities Commission (CPUC) by Decision 08-01-042 directed that the DEER team develop new NTG ratios based on more recent evaluation results, in order to increase confidence in their application to current and future programs. To meet this objective, a comprehensive literature review was conducted in order to develop more up-to-date values for DEER. Evaluation studies recently completed in California covering program years 2002 through 2005 were reviewed, along with relevant evaluations conducted in other states for programs offered during a similar time frame. The California evaluations were primarily for Statewide and Local programs offered in PY2002-2005, and included NTG ratios by technology, measure and delivery method.

The review focused on the methods used to estimate the fraction of program participants likely to be free riders, e.g., those participants that are likely to have adopted the energy efficiency measures in the absence of any program. Current CPUC policy (CPUC, 2007) uses a definition of net savings that incorporates free ridership but excludes spillover[2] and market effects. The CPUC's definition of NTG is based on the belief that spillover estimates are still too uncertain for use in estimating the net benefits from programs within the financial context of the CPUC's risk-reward mechanism (which can result in significant monetary shifts between shareholders and ratepayers). The NTG values reflected in the DEER 2008 update thus include the effects of free ridership only. This perspective can be viewed as an indication of marginal program efficacy. That is, how effective is the current program in influencing current participants to adopt new efficiency measures, regardless of the effects of previous program years. It should be noted that California may consider spillover for some purposes in the future and, therefore, evaluation efforts are underway to estimate participant spillover and market effects associated with some energy efficiency projects implemented during the current 2006-2008 program cycle. In anticipation of this potential change in NTG policy, the DEER team also captured any estimates of participant and nonparticipant spillover in its review of the 2002-2005 evaluation studies. Approximately two-thirds of the evaluations contained some analysis of free ridership and/or participant spillover or both.

The DEER team reviewed a total of 52 evaluation studies, representing 14 separate program categories. For each study, the following information was summarized:

- the methods used to estimate NTG ratios;
- the program context pertinent to these values, i.e., target markets, delivery strategies and eligible measures; and
- estimates of Free Ridership, and if available, Participant and Nonparticipant Spillover at the end use and/or measure level.

Data gathered during the research included the name of the study, target market, number of measures, sample size, precision level, and ranges for participant and non-participant spillover. In some cases, information was also collected on the market share, rebate level, and incremental cost to provide

---

[2] Spillover refers to energy savings that are attributed to the presence of energy efficiency programs, but occur outside of the programs. There are two types of spillover: (1) Participant spillover, referring to actions undertaken by participants which were significantly influenced by their prior participation in the program, but which did not receive rebates. (2) Nonparticipant spillover, referring to general market activity by nonparticipants resulting in the installation of program-promoted measures but occurring outside of the program.

context for the NTG ratios that were developed. An overall confidence level was assigned to each estimate of the NTG ratio, based on the survey design, question wording, methods used, sample size and consistency/reasonableness of results. Initial NTG ratios by measure and method were recommended by the DEER team and released by the CPUC for review by utilities and stakeholders in early April 2008. After an extensive review and comments process, final NTG values were developed and published. Final values to be used for 2009-2011 program planning were released in May 2008, and a second set of values to be applied to 2006-2007 program performance was published in October 2008 to be used to assess the utilities' earnings claim.

Why is Net Savings Measurement Important?

The focus of net savings measurement is on understanding the net effects of program interventions (such as rebates or technical assistance) on customer decisions to install efficiency measures. We say net because it is important to isolate the program's effect compared to other factors that may have affected the decision to install energy efficiency measures, either by participants in a program, nonparticipants, or both.

However, measuring the program's impact on individual decision making is not always straightforward. Some have recently questioned the legitimacy and accuracy of NTG analysis methods and results, (Peters/McRae, 2008). These assertions are based on their expressed concerns over a customer's ability to recall the reasons for a past purchase decision, sensitivity to how customer responses are weighted, imperfections associated with comparison groups and time series data, and other factors. We strongly disagree with this perspective and note that challenges associated with measuring the net effects of efficiency programs are no more daunting than those facing other professions charged with evaluating education, public health, pharmaceutical, and other policy and medical interventions. That is not to say that continued improvement in measurement techniques are not needed to help increase confidence in results, but the conclusion should be that more and improved measurement is needed, not less. There are a number of reasons why understanding the net effects of programs has been and will be increasingly important and necessary:

- **Understanding program and portfolio cost-effectiveness.** Net savings measurement is needed in order to understand the net value of the program and the portfolio to utility ratepayers, and to society. For publicly funded programs, net estimation of savings and program benefits provides legislators and the general public with important feedback on how this use of public funds compares with other potential uses.
- **Improving portfolio design and resource allocation.** Having information on program-level net savings allows administrators to determine which programs are most and least effective in terms of impact yielded per dollar spent (i.e., "bang for the buck"). This information is needed to support their decisions regarding budget/resource allocations between programs and among measures and market segments**.**
- **Refining program design and tactics.** Measurement of program free ridership and spillover provides program administrators with valuable feedback they can use to determine *if* the program needs to be improved and, if so, *how* it can be designed to be more effective in terms of improving net savings. Administrators have a number of different strategies they can use to cut down of free ridership, if warranted.
- **Understanding market transformation.** NTG ratios are a key metric in determining the rate at which a technology market is being transformed, which is a primary goal of many programs. In general, progressively lower NTG ratios over time are an indication of this transformation (that is, for NTGs defined to be exclusive of spillover and market effects).

- **Aligning program administrators' financial interests with societal interests.** Making program administrators accountable for net, rather than gross savings aligns their interest with those of society. In the absence of a mechanism which values their performance and provides compensation based on net savings, program administrators have a powerful incentive to allow NTG ratios to fall to unnecessarily low levels. This is true because meeting savings goals is challenging and it is easier to convince someone who is already planning to adopt an efficiency measure to participate in a program than to persuade someone who has no such plans to change their mind.
- **Understanding how EE programs affect baseline load forecasts and short-term power procurement decisions.** Many jurisdictions view energy efficiency as an integral part of their power supply. Net program savings estimates from independent evaluations are needed to support energy resource procurement decisions and estimate efficiency effects on reference load forecasts. Of increasing importance around the country is the issue of how increased goals for energy efficiency affect baseline load forecasts. All or a portion of goals may already be included in baseline energy load forecasts, depending on how the forecasts are developed. Net to gross estimation provides critical information to the forecasting community to help improve projections and resolve this question of overlap.

Given the importance of this topic, it is important for policy makers and program administrators to understand the strengths and weaknesses of available methods as discussed below.

Review of Methods Used to Estimate NTG Ratios

A number of approaches are available to estimate free ridership and spillover effects. Some of these methods can be applied to virtually all types of programs and measures, while others are somewhat limited in their application. As an example, CFL rebate program evaluations have involved the greatest number and variety of methods, several of which can be applied to other types of programs. Our literature review revealed seven basic methods of estimating the net and peak savings attributed to CFL rebate programs. Each of these methods is described below along with some general observations on the challenges in using each method.

- *Self-report methods*
  - *Customer self-reports* (Method 1), relying on interviews of participating customers. A challenge presented for upstream programs, in particular, is that the use of product buy downs or instant discounts at the cash register makes the program invisible by design to many customers. As a result, customer responses about what they might have done in the absence of the program may not be reliable. For more traditional downstream programs, the customer self-report approach is one of the most commonly used methods to assess free ridership, particularly for programs where it is difficult to obtain reliable comparison groups. There is a wide range within this method with respect to level of detail and approaches to proactively address and mitigation potential biases (Ridge, et. al., 2009).
  - *Supplier self-reports* (Method 2). Manufacturers/Retailers' predictions of product sales with and without the program rebates have also been used to provide the basis for estimating NTG ratios. This method is sometimes used for upstream rebate programs. One concern is that supplier responses may be biased because suppliers may realize that giving the "right" answers can have an effect on both the continuation of the rebates themselves and their company's position in the competitive market. As with customer self reports, evaluators must anticipate potential biases and design approaches to mitigate them.

- *Sales Based Methods*
  - *Per-capita sales comparisons* (Method 3) with a comparable state(s) that does not have a program. Sales in states without programs represent the level of baseline sales in the absence of the program. This method can, in theory, provide an estimate of net program effects if the comparison areas are very similar to the program area. If not, it should be combined with appropriate statistical normalization for differences in customer and market characteristics that affect adoption. Unfortunately, there are numerous limitations to the data available for this approach that restricts its overall usefulness. In particular, the available adoption data is not always reliable, nor are data on the necessary normalizing variables (i.e., limiting comparison of program areas from non-program areas).
  - The so-called *"paired comparison"* approach (Method 4), has also been used on a limited basis. This method is currently being used in Wisconsin. This method involves comparison of EE product sales data for a leading big box retailer in a state with rebates versus a similar nearby state without EE product rebates. The difference in in-store EE product sales is attributed to the state with a rebate program. The big challenge with this method is finding a representative retailer operating in jurisdictions with and without rebates and making sure it is possible to control for differences in demographics between the control and experimental markets.
- *Econometric Methods:*
  - *Discrete choice analysis* (Method 5), an econometric method used to estimate EE product purchases made by customers as a function of factors that influence EE demand such as product awareness, prices, and other factors. However, in rapidly transforming markets where prices and product content are very dynamic, it may be difficult to construct a set of known price and energy savings tradeoffs that can be presented to a sample of customers as part of a discrete choice analysis. In practice, discrete choice methods have not been used often to estimate net impacts due to their complexity and expense. This method relies on a large body of nonparticipant survey data (usually several thousand surveys or more), which evaluation budgets often have not supported. However, if a measure accounts for a significant share of overall program and portfolio savings impacts, it may be well worth the expense for these important measure categories.
  - *Estimating a demand model* (Method 6) to predict the relationship between changes in EE product price, different levels of customer awareness generated by the mass media and incremental EE product sales in different regions of the country. Theoretically this model can be used to predict the net energy savings produced by reducing the price of an efficiency measure via rebates or increasing the level of awareness through program expenditures. This method is currently under development in California for screw-in CFLs.
  - *Net billing analysis* (Method 7) can also be used for measures that account for a minimum of 5 to 10 percent of end-use consumption[3]**.** There are a number of billing analysis approaches that can be employed to estimate the net effects of a program. One method developed is the inclusion of the Double Inverse Mills Ratio in the standard SAE billing analysis approach. This approach corrects for self-selection bias and can be used to estimate free riders.

[3] This method is not feasible for smaller measures such as CFLs, since such impacts have a very low "signal to noise" ratio and are difficult to detect.

Methods 1, 2, and 5 are primarily used by evaluators interested in estimating the fraction of program participants that can be identified as free riders and thus used to adjust gross savings downward. Methods 1 and 2 can also be adapted to estimate spillover by administering surveys to nonparticipants as well as participants. Methods 4, 5 and 6 are primarily used to estimate the net effect of the programs and may include both spillover and free rider impacts. Method 7 has routinely been used in conjunction with both the self report and discrete choice analyses to cross-validate the NTG results. However, if the relationship of the savings to the normal monthly variation in energy use is too small, or if the measures and savings in question of very heterogeneous (i.e., site specific), then a billing analysis should not be used. For this reason, billing analysis is rarely successful en mass for industrial customer projects.

Summary of Studies Reviewed and Resulting NTG Recommendations

Table 1 below provides a snapshot of the evaluation studies reviewed for the DEER NTG update. Roughly 67% of these studies estimated a NTG ratio, while the remainder used deemed NTG estimates. Of the former, nearly three-fourths used the self-report method to estimate the NTG ratio. In some cases, there were multiple methods used including the self-report approach plus others such as billing analysis, and discrete choice.

Table 1: Review of NTG Estimation Methods Used in Studies Reviewed for DEER NTG Update

Program Category	# of Studies Reviewed	# of Studies with NTG analysis	# of discrete measures with NTG Analysis
Residential Lighting	3	2	3
Residential HVAC	3	3	10
Residential Energy Star Appliances	2	1	5
Residential New Construction	2	2	5
Residential Multi Family	3	3	7
Residential Appliance Recycling	2	2	2
Residential Direct Install (Third Party Program)	2	1	6
Residential Audit Program	4	3	Many
Nonresidential Prescriptive Rebates	5	2	11
Nonresidential Audit	4	4	5
Nonresidential Direct Install and Other (Third Party Program)	2	2	7
Nonresidential New Construction	2	1	10
Large Nonresidential Custom Rebate	2	2	5
Agricultural Rebates	1	1	1
Retrocommissioning	2	2	Many
Local Government/University Partnership Evaluations	13	4	Many
Total Evaluations Reviewed	***52***	***35***	***77***

Table 2 below provides a comparison of NTG ratio values from the 2001 NTG summary effort (CALMAC 2000) and the updated 2008 DEER values for selected measures (those with the most significant change in values). Many of the differences shown in NTG values between the 2 research

efforts resulted from either a major change in markets or a change in overarching CPUC NTG policies. For example:

- Values suggested by CALMAC for 2001 primarily reflected the application of **default NTG values** of 0.80 for residential measures and 0.96 for nonresidential measures.[4] The 2001 analysis produced very few program or measure-specific values, but instead averaged the results of dozens of evaluation studies addressing a wide range of program/measure types. These averages were then used to set default values which were applied whenever program or measure-specific values were not available.
- The more recent DEER NTG ratios for **CFL screw-in bulbs** incorporate the effects of the rapidly transforming CFL market, reflecting dramatically lower bulb prices, increased acceptance and adoptions by the general population, and resulting higher free ridership levels. To develop the updated values, the DEER team reviewed the PY 2004-05 Single Family Rebate report, along with other CFL evaluation, measure cost, and market share studies, and performed its own sales-based NTG analysis. The documentation for this analysis (Itron, 2008) contains an extensive discussion of issues and analysis of sales data and trends, on a state, regional and national level.
- The increase in the NTG ratio for the **residential appliance recycling** program category reflects certain changes in the NTG methodology. These changes resulted in the inclusion in gross savings estimates of certain adjustments formerly included in the NTG ratio.
- Declining NTG ratios for various **residential appliances** and **nonresidential lighting, refrigeration and water heating technologies** reflect the move away from default values to NTG ratios based on a measure-specific NTG analysis in the 2008 study, as well as the likely effect of increased awareness and acceptance of these technologies by the general population. The **nonresidential lighting** reduction also reflects the exclusion of spillover in the 2008 NTG update (the 2001 value of 0.96 included spillover).
- The reduction in the NTG ratio for the **Nonresidential Custom measure** is due to the CPUC Energy Division's elimination of +0.10 and +0.05 adjusters for self-report bias, and participant spillover, respectively. These adjustments were permitted in the 2001 NTG update process.

[4] Note that these previous default values were not a DEER work product.

Table 2: Comparison of 2008 and 2001 NTG Ratios for Selected Measures

Target Market(s)	Measure Name(s)	Program Delivery Method/ Channel	NTG Values Based on 2001 Analysis	NTG Values Based on 2008 DEER Analysis
Residential	CFL-screw in	Upstream prescriptive rebate	80%	74% for PY2004-05, 60% for 2006-2007 and 2009-11
	Refrigerator/Freezer	Downstream prescriptive rebate	80%	57%
	Refrigerator	Turn-in/recycling	35%	52%
	Pool pump-2 Speed	Downstream prescriptive rebate	80%	69%
	Dishwasher	Downstream prescriptive rebate	80%	41%
	Central AC>14 SEER	Downstream prescriptive rebate	80%	67%
	Programmable thermostats	Downstream prescriptive rebate	80%	49%
	New single family home	Custom rebate, based on performance vs. code	80%	48%
	New multi family building	Custom rebate, based on performance vs. code	80%	50%
	Default value	For new measures or delivery methods where no previous NTG results are available.	80%	80%
Nonresidential	CFL-screw in	Upstream prescriptive rebate	96%	74% for 2006-07, 60% for 2009-11
	CFL-screw in	Downstream prescriptive rebate	96%	81%
	T5s and T8s, 4-foot and 8-foot lamps	Downstream prescriptive rebate	96%	78%
	Occupancy Sensors	Downstream prescriptive rebate	96%	77%
	Refrigeration Strip Door Curtains	Downstream prescriptive rebate	96%	46%
	Clothes washers, boilers, process boilers and insulation	Downstream prescriptive rebate	96%	46%
	New Construction Whole Building Measures	Building Design Incentive	82%	70%
	Custom Electric and Gas Measures for Large Customers	Custom incentive	69%	54%
	Medium/Large Customers (>100 kW) - All Measures	On-Site Audit	83%	46%
	Default value	For new measures or delivery methods where no previous NTG results are available.	96%	80%

Our team was also interested in assessing whether net to gross ratios change over time or can be used to generalize the likely net savings from any given portfolio. We decided to examine net to gross ratios from the most recent generation of programs and compare this to analyses of NTG values for California energy efficiency programs operated in the early 1990's and 1980s.

Table 3 below provides an approximate estimate of the savings weighted NTG ratio for PY2004-2005, based on the DEER NTG update results for the residential and commercial sectors. In the

calculation of this estimate, NTG ratios for each program were weighted by the claimed savings for each program. The resulting savings weighted NTG estimate is 0.72. There is a fairly wide range of NTG values across the different end-use categories both within and across residential and nonresidential programs. Values range from a low of 0.49 for Residential New Construction programs to a high of 0.87 for Commercial New Construction programs.

Table 3: Savings Weighted Average NTG Ratio for 2004-2005 Using Updated DEER NTG Values

End Uses	2004 GWh/Year	2005 GWh/Year	2004/05 End-Use Share of Savings by Sector	End-Use Share of Total Savings	NTG ratio for 2004/05 from Studies	Savings Weighted NTG ratio
Residential						
Lighting	624	799	82.3%	40.3%	0.74	0.30
Appliances	65	92	9.0%	4.4%	0.65	0.03
Central AC	54	81	7.8%	3.8%	0.56	0.02
New Construction	6	8	0.8%	0.4%	0.49	0.00
Subtotal	**750**	**980**		**49.0%**		**0.35**
Commercial						
Lighting	250	765	56.4%	28.8%	0.74	0.21
HVAC	69	272	19.0%	9.7%	0.60	0.06
Refrigeration	34	132	9.2%	4.7%	0.74	0.03
New Construction	53	99	8.4%	4.3%	0.87	0.04
Other	30	96	7.0%	3.6%	0.70	0.02
Subtotal	**436**	**1,363**		**51.0%**		**0.37**
Total	**1,186**	**2,343**			**weighted NTG ratio**	**0.72**

Sources: Utility Annual Savings reports used to estimate end use splits and to allocate savings between delivery mechanisms to derive the savings weighted NTFR for each end use in the next to last column. Commercial Other includes miscellaneous measures which did not have sufficient documentation to place in other commercial end uses.

A review of program evaluation methods and reported net to gross ratios performed by the California Energy Commission (CEC) staff in 1994 found an average NTG ratio of 0.70 across all utilities and over 20 evaluations - maximum NTG of 0.97 and minimum of 0.30 NTG (CEC 1994). A similar review performed for the CEC in 1988 found NTGs of 0.80 for Commercial Audits, 0.60 for Commercial Incentives, 0.70 for Industrial Audits, and 0.50 for Industrial Incentives (Rufo/Bester, 1989). This suggests that NTG ratios (exclusive of spillover) for large program portfolios in California may have been relatively stable over time.

Which NTG (Net) Estimation Method Should I Use?

The answer to this question depends on the specific circumstances and goals of the evaluation. Answers to the following types of questions can help to guide the choice of method(s) for a particular program or portfolio:

- **What are the policy goals of the program?** If short-term program efficacy or resource acquisition is the primary policy objective, methods that quantify nonparticipant spillover and

broader market effects may not be of interest. Evaluations of such programs need only focus on free ridership. In cases where stimulating market effects are a strong objective of the program, the selected method(s) will need to be able to measure the program's effects on both participants and nonparticipants. Methods 3, 4, and 6 are usually used to measure the net program effect on market sales of efficient products.

- **How mature is the program?** If the program is in its infancy and is promoting measures that are less well-known, a sales-based approach is not useful, since sales levels are likely to be very low at the beginning of the program. Mature programs that have been successful and run for at least three years are likely to have had at least some market effects and warrant the use of a sales based approach. An exception to this general maturity rule must be made for products that have been promoted by utility programs for many years in neighboring jurisdictions even though the program may be new in any particular area. Programs that promote CFLs are an example of this exception where evaluation planners should seriously consider starting with a market sales or demand modeling approach.
- **What is the program design?** *Does the program work with upstream suppliers or downstream customers?* For upstream program designs, it may not be possible to identify or develop an unbiased sample of participants to survey, thereby ruling out the use of supplier self-reports (Method 2). *Does the program promote customized measures only*? If so, the econometric and sales-based approaches are not feasible. For programs that promote custom technologies to large industries, a billing analysis approach cannot be used, leaving only the self-report approach.
- **How much budget do I have?** A limited evaluation budget may preclude use of more than one method to estimate net program effects. Even if the budget is allocated strategically, toward programs which account for the greatest share of overall portfolio savings, the available evaluation budget may not support the use of more expensive methods. Although there can be considerable variation, even within methods, of the cost involved, certain generalizations can be made. For example, simple billing analyses (Method 7) and sales based approaches (Methods 3 and 4) are less costly than a discrete choice approach (Method 5).
- **What data is readily available?** Sales-based approaches (Method 4), in particular, rely heavily on publicly available data sources for information. Since the reporting of sales to these sources is sometimes voluntary, the data are often incomplete, as is the case for various CFL sales data sources, which exclude sales from one or more important sales channels. If a good, complete and reliable data source is available to support the NTG calculations, then a sales-based approach may be the best choice for assessing the full influence (free ridership + spillover) of a particular program.
- **Is a suitable comparison group available?** Method 4, the paired comparison approach, relies entirely on finding a representative retailer operating in jurisdictions with and without rebates. Similarly, approaches involving comparisons of sales and the estimation of a demand model (Methods 3 and 6) also require a good deal of diversity in the market conditions and the availability of "no program" areas. In both cases, the lack of a suitable control group or the budget to collect data from remote areas are often limitations on the use of these methods.
- **What level of precision is desired?** Most administrators have a difficult time deciding how much precision is needed in estimating net effects. We recommend the level of precision be determined based on the needs of policy makers and the program. If a high level of precision is required and the budget is limited, this may rule out the use of multiple methods and/or use of more costly methods such as discrete choice.
- **Are there performance-based metrics which must be met?** The net measurement strategy may also be affected by the presence of performance-based metrics in the program

administrator's goals or contract. For example, the administrator's compensation may be tied to a maximum "no more than X percent" level of free ridership, or a minimum "at least Y percent" level of nonparticipant spillover. The adopted net measurement strategy must thus be capable of reporting results at this level of detail.

Each of the above methods has its particular strengths, as well as weaknesses. As noted above, there are many considerations in selecting a method or methods to implement. If there is sufficient budget and available data to support use of more than one method, we recommend that multiple methods be used and the results then be "triangulated". This can be done by either constructing a range of values for the NTG ratio, or by developing a point estimate based on applying the highest weights to the methods with the most reliable results. In our previous evaluation studies, when more than one method has been used to estimate the NTG ratio, we have found that the resulting values for the NTG ratio are generally similar and therefore can be used to develop a range of NTG values and/or a weighted average point estimate (e.g. NTG ratio range of between 0.79 to 0.89, point estimate of 0.85).

Conclusion

In conclusion, it is important to periodically assess and update free ridership estimates by performing a literature review, as was recently done in California, to discern any underlying trends or changes in assumed NTG ratio levels. Our review has found that despite the widespread changes in equipment markets, and multitude of NTG methodologies, in general, portfolio-level NTG ratio values have been relatively constant since the late 1980s. The choice of a specific methodology for measuring program and measure-specific NTG ratios is a complex one, which must consider the policy context, the level of market transformation, the specific program delivery approach, the size of the evaluation budget and the availability of comprehensive and reliable data sources.

References

California Public Utilities Commission. 2007. *Energy Efficiency Policy Manual*, *Version 3*. December 2007.

CALMAC 2000. California Measurement Advisory Committee Public Workshops on PY 2001 Energy Efficiency Programs, prepared in response to Administrative Law Judge's August 15, 2000 Ruling Establishing Schedule and Process for PY 2001 Energy Efficiency Program Planning, September 25.

CEC 1994. Staff Review of Utility DSM Evaluation Studies in Response to April 14, 1994 Committee Order as part of the 1994 Electricity Report. July 1994.

Itron, 2008. *Database for Energy Efficient Resources Draft 2006-2007 Ex Ante Net-To-Gross Ratio Update*, August 20, 2008. Available at www.deeresources.com

Peters, Jane and McRae, Marjorie, "Free-Ridership Measurement Is Out of Sync with Program Logic…or, We've Got the Structure Built, but What's Its Foundation?," *Proceedings from the 2008 ACEEE Summer Study*, August 2008

Rufo, Michael W., and Nancy L. Bester, "An Investigation of Commercial and Industrial Utility Demand-Side Management Program Impacts," *Energy Program Evaluation: Conservation and Resource Management, Proceedings from the 1989 International Energy Program Evaluation Conference*, Chicago, IL, August, 1989

Rick Ridge, Willems, Phillippus, and Jennifer Fagan, "*Self-Report Method for Estimating Net-to-Gross Ratios in California: Honest!"* Upcoming in Proceedings from the AESP's 19th Annual National Energy Services Conference and Expo, January 2009