Any Way You Slice It: Issues of Behavior and Influence in Net Impact Analysis

Christie Torok, Quantum Consulting Inc., Berkeley, CA
John Cavalli, Quantum Consulting Inc., Berkeley, CA
Mary O'Drain, Pacific Gas & Electric Company, San Francisco, CA

Abstract

The Commercial Energy Efficiency Incentives (CEEI) Program, run by Pacific Gas & Electric, is one of the largest programs of its kind. The CEEI program offers fixed rebates to customers who install specific electric energy efficient equipment, and customized financial incentives to customers who undertake large, complex projects that save gas or electricity.

For 4 years, Quantum Consulting has conducted an evaluation to determine the first-year gross and net energy, demand and therm impacts of the program. This paper addresses the relative merits, in both theory and practice, of the methodologies applied in the 1996 and 1997 program year evaluations to derive the net program effects and net-to-gross ratio. In this paper the term 'net-to-gross' is defined as the process of adjusting estimates of gross program impact for the effects of free ridership and spillover, resulting in an estimate of 'net impact.'

Three approaches were used to estimate net program impacts: billing data analysis, self-reported data analysis, and discrete choice analysis. By using three different approaches we created the opportunity for cross-validation and were able to derive information from the advantages of each individual approach.

Results for the 1996 and 1997 evaluations show each of these methods to be remarkably consistent. Specific issues in applying the three models are addressed in detail, as are the theoretical unpinnings of each model. We conclude that the most preferred of the three approaches is the two-stage discrete choice model, but recommend the continued use of multiple approaches to derive net impact results.

Introduction

The Commercial Energy Efficiency Incentives (CEEI) Program, run by Pacific Gas & Electric, is one of the largest programs of its kind. The CEEI program offers fixed rebates to customers who install specific electric energy efficient equipment, and customized financial incentives to customers who undertake large, complex projects that save gas or electricity.

For 4 years, Quantum Consulting has conducted an evaluation to determine the first-year gross and net energy, demand and therm impacts of the program. This paper addresses the relative merits, in both theory and practice, of the methodologies applied in the 1996 and 1997 program year evaluations to derive the net program effects and net-to-gross ratio. In this paper the term 'net-to-gross' is defined as the process of adjusting estimates of gross program impact for the effects of free ridership and spillover, resulting in an estimate of 'net impact.' For the sake of brevity, and avoiding redundancy, all specific data and model specifications will reflect the lighting end-use program.

Three approaches were used to estimate net program impacts: billing data analysis, self-reported data analysis, and discrete choice analysis. By using three different approaches we created the opportunity for cross-validation and were able to derive information from the advantages of each individual approach.

Net Billing Analysis Method

A billing analysis was used to estimate *statistically adjusted engineering* (SAE) coefficients to modify gross engineering estimates and calculate net energy impact. The net billing model specification incorporates both participants and nonparticipants into one model. A disadvantage of combining both groups into one model of net energy savings, is that the resulting sample is not randomly determined. In particular, participants self-select into the program and therefore are unlikely to be randomly distributed. There are certain unobserved characteristics that influence the decision to participate. If these characteristics are not accounted for in the model, the net savings model could produce biased coefficient estimates.

One solution to this problem is to include an Inverse Mills Ratio in the model to correct for self-selection bias. This method was developed by Heckman (1976, 1979)[1] and is used by others (Goldberg and Train, 1996[2]) to address the problem of self-selection into energy retrofit programs. This assumes that the unobserved factors that are influencing participation are distributed normally. Including an Inverse Mills Ratio in the model as an explanatory variable controls for the influence of the characteristics that cause participants to self-select into the retrofit program. This corrects for the self-selection bias in the net savings regression as the unobserved factors affecting participation are now controlled for in the model. As a result, standard regression techniques should produce unbiased coefficient estimates.

Goldberg and Train (1996) developed the technique of including a second Inverse Mills Ratio in the savings regression to account for the possibility that participation is correlated with the size of energy savings. The second Mills Ratio is interacted with a measure of energy savings, which allows the amount of net savings to vary with participation. The rationale for the second term is that those customers who have potentially large savings are more likely to participate in the program. Consequently, the unobserved factors that are influencing participation are also affecting the amount of savings.

To calculate the Inverse Mills Ratios, a probit model of program participation is estimated, and the parameters of this model are used to calculate an individual Inverse Mills Ratio for all participants and nonparticipants. This Mills Ratio is included in a net savings regression that combines both participants and nonparticipants into one model.

If the Mills Ratio controls for those unobserved factors that determine participation (i.e. the self-selection bias), and the other model assumptions are met, then the net savings model will produce unbiased estimates of net savings. The resulting SAE coefficients on the energy impacts (that have been interacted with the Mills ratios) are then used to adjust the engineering estimates of expected annual energy impacts (the original SAE coefficients) for the entire participant population. This is one estimate of net ex post energy impacts.

---

[1] Heckman, J. 'The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models.", Annals of Economic and Social Measurement, Vol. 5, pp. 475-492, 1976.

Heckman, J. "Sample Selection Bias as a Specification Error." Econometrica, Vol. 47, pp. 153-161, 1979.

[2] Goldberg, Miriam and Kenneth Train. 'Net Savings Estimation: An analysis of Regression and Discrete Choice Approaches', prepared for the CADMAC Subcommittee on Base Efficiency by Xenergy, Inc. Madison, WI, March 1996.

For the 1996 Evaluation, this net billing model was run with 428 nonparticipants and 591 participants. For the 1997 evaluation, this model was run on a sample of 487 nonparticipants and 679 participants. For the 1996 model, it was found that the net billing model results were significant at the 95 percent level in all cases but one, lighting warehouses. For the 1997 model, the net billing model results were significant at the 90 percent level in all cases. The parameter coefficients from the net billing model represent net participation within that technology. From these estimates, we can now "back out" an estimate of free ridership, by taking the product of these coefficients with their Mills ratio and dividing by the regression coefficients from the gross model. This equation has the following functional form:

$$(1 - FR)_m = \frac{Mills_m * \delta_m}{\beta_m}$$

Where,

$Mills_m$ is the mean Mills coefficient for all customers with technology m;

β_m is the SAE coefficient from the Gross Billing model for technology m; and,

δ_m is the regression coefficient from the Mills Model 1 regression for technology m.

Table 1 below illustrates the resulting estimates of the net-to-gross ratio, or one minus free-ridership for the 1996 and 1997 evaluations.

Table 1. 1996 and 1997 Net Billing Regression Analysis Estimates of (1-FR)

	Resulting (1-FR)	
	1996	**1997**
Lighting Offices	0.82	0.91
Lighting Schools	0.94	0.76
Lighting Hotel/Motel	1.40	0.61
Lighting Warehouse	0.84	-
Lighting Miscellaneous	1.02	1.21
HIDs	1.78	-

A drawback to this method is that large customers can exert such a significant influence on the results that they can overly bias the results. Therefore, very large customers are excluded from the sample. This is not an ideal solution, because potentially useful information from these customers is not used. In addition, larger customers are generally more likely to be free riders, so removing these customers may cause the model to underestimate true free ridership[3]. In addition, the usable sample is further reduced by the requirement that each customer have good historical billing data. Finally, the Inverse Mills Ratio method does not produce an estimate of spillover, rendering it an incomplete model of net impact. On the other hand, the results of the net billing analysis include estimates of the *magnitude* of net impacts, while the other two methods described below result in an estimate of the fraction of total impacts attributable to the program.

---

[3] It is not a feasible solution to use a standardized measure such as kWh per square foot because large and small customers are likely to have substantively different behavior. That is, behavior should not be assumed to be related only to concentration of energy use, but also may be effected by the absolute level of energy use.

Self-Report Method

The second approach is a "self report" method that is based on customers' responses to direct questions about their equipment choices and purchase decision process. Net participation is calculated by removing the free ridership component from the gross impact, and then adding the effects of participant spillover and nonparticipant spillover. Each of these components is estimated separately using self-report techniques. Specifically, survey questions regarding program awareness, program influence, and what the respondent's behavior would have been in the absence of the program, are analyzed. Scoring algorithms are applied to determine rates of free ridership and spillover among participants, as well as nonparticipant spillover. For the 1996 Evaluation, 808 participants and 4,258 nonparticipant were surveyed. For the 1997 Evaluation, a total of 860 participants and 4,168 nonparticipants were surveyed.

Free Ridership

Participants can be classified into four basic categories depending on the actions they would have taken in the absence of the CEEI program:

1. In the absence of the CEEI program, the participant would not have installed any new equipment
2. In the absence of the CEEI program, the participant would have installed standard efficiency equipment
3. In the absence of the CEEI program, the participant would have installed high efficiency equipment, but not as soon (more than one year later)
4. In the absence of the CEEI program, the participant would have installed high efficiency equipment at the same time (within the year)

Customers who fall into the first three categories are considered net program participants in the calculation of first year net impacts. Customers who fall into the fourth category are considered free riders. The self-report estimates of free ridership are based on these four categories. Responses consistent with category 4 were counted towards free ridership. Responses consistent with categories one through three were counted towards net participation.

Self-reported estimates of free ridership for the 1996 and 1997 evaluations are presented by technology group in Table 2 below. For both years, the technology group with the lowest rate of free ridership was the Delamp Fluorescent Fixtures category, comprised of fluorescent delamping actions implemented by the respondents. Technologies with relatively small sample sizes show larger fluctuations in estimated free ridership. The overall rate of free ridership remains consistent between the 1996 and 1997 results, at 0.25 and.0.26, respectively.

Table 2. Weighted Self-report Estimates of Free Ridership for Lighting Technology Groups 1996 CEEI Program

Technology Group	*Sample*	*Free Ridership*
Halogen	30	45.0%
Compact Flourescent Lamps	164	37.2%
Incandescent to Flourescent Fixtures	26	68.5%
Exit Signs	81	18.1%
Efficient Ballast Changeouts	13	24.7%
T-8 Lamps and Electronic Ballasts	391	24.1%
Delamp Flourescent Fixtures	118	13.4%
High Intensity Discharge	53	19.8%
Controls	57	57.2%

1997 CEEI Program

Technology Group	*Sample*	*Free Ridership*
Halogen	26	30.4%
Compact Fluorescent Lamps	165	36.7%
Incandescent to Fluorescent Fixtures	14	16.7%
Exit Signs	79	36.5%
Efficient Ballast Changeouts	12	54.5%
T-8 Lamps and Electronic Ballasts	323	23.1%
Delamp Fluorescent Fixtures	83	12.9%
High Intensity Discharge	52	49.2%
Controls	56	23.9%

Spillover

A spillover action is defined as a lighting action taken outside of the program which increases energy efficiency, and occurred as a direct result of the program' s influence. In counting the total number of surveyed participants and nonparticipants contributing towards spillover, the following four conditions, which reflect this definition of spillover, were used:

1. the action involved the installation of **high efficiency lighting equipment**, as recognized by the CEEI program
2. the respondent was **aware** of the program **before** making the decision to purchase new lighting equipment
3. the action was **not rebated** as part of the program
4. the respondent stated that this action was taken as a result of the **CEEI program's influence**

Questions that unequivocally show each of the above conditions to be either true or false were used to identify spillover adoptions in the participant and nonparticipant populations. The next step was to apply the analytical method, which is composed of three steps:

- Identification of the spillover rate
- Calculation of the impact per instance of spillover
- Estimation of the spillover contribution to the net-to-gross ratio

The spillover rate is the percentage of the participant or nonparticipant population that is identified as being influenced by the CEEI program to install non-rebated high-efficiency equipment. The spillover rate is estimated using self-reported information from the surveys. The estimate of impact per spillover adoption is based on the equipment installed as reported in the surveys. The contribution of spillover to the net-to-gross ratio can then be estimated as:

Participant Spillover:

NTGpart_spill = SP_RATEpart * POPpart*IMPACTpart_spill/IMPACTpop
Where,

NTGpart_spill = the participant contribution of spillover to the net-to-gross ratio
SP_RATEpart = the participant spillover rate
POPpart = the participant population, in number of sites
IMPACTpart_spill = the per participant site impact associated with spillover
IMPACTpop = the total CEEI Program impact

Nonparticipant Spillover:

NTGnp_spill = SP_RATEnp * POPnp*IMPACTnp_spill/IMPACTpop
Where,

NTGnp_spill = the nonparticipant contribution of spillover to the net-to-gross ratio
SP_RATEnp = the nonparticipant spillover rate
POPnp = the nonparticipant population, in number of sites
IMPACTnp_spill = the per nonparticipant site impact associated with spillover
IMPACTpop = the total CEEI Program impact

Spillover results for the 1996 and 1997 program year evaluations are presented in Table 3 below. In general, the estimates are reasonably consistent. However, the number of instances of spillover occurring within the sampled population is likely to be relatively small, particularly for the nonparticipant sample. Out of about 4,200 completed telephone surveys in 1996 and 1997, the number of spillover adoptions identified was 10 in 1996 and 13 in 1997. Therefore, the sample must be very large to ensure robust estimates, or will remain sensitive to small variations in spillover adoptions. One important lesson from these results is that a very large number of surveys are required to gather data sufficient to yield reliable results. We feel that, for nonparticipants, a sample of 4,200 is sufficient, but not by a large margin. In general, the more resources devoted to the data collection effort, the more reliable the self-report estimates will be.

Table 3. 1996 Participant Spillover Estimate

Avoided Cost Per Participant	$15,586
Spillover Rate	3%
Number of Participants	5,230
Number Contributing to Spillover	168
Spillover Avoided Cost	$2,622,950
Lighting Avoided Cost	$51,077,333
NTG Contribution from Participant Spillover	**5.14%**

Table 3 Continued. 1997 Participant Spillover Estimate

Avoided Cost Per Participant	$18,694
Spillover Rate	1.74%
Number of Participants	5,308
Number Contributing to Spillover	93
Spillover Avoided Cost	$1,730,714
Lighting Avoided Cost	$59,140,572
NTG Contribution from Participant Spillover	**2.93%**

1996 Nonparticipant Spillover Estimate

Avoided Cost Per Nonparticipant	$8,473
Spillover Rate	0.08%
Number of Nonparticipants	408,668
Number Contributing to Spillover	320
Spillover Avoided Cost	$2,710,747
Lighting Avoided Cost	$51,077,333
NTG Contribution from Nonparticipant Spillover	**5.31%**

1997 Nonparticipant Spillover Estimate

Avoided Cost Per Nonparticipant	$10,932
Spillover Rate	0.092%
Number of Nonparticipants	411,188
Number Contributing to Spillover	378
Spillover Avoided Cost	$4,137,013
Lighting Avoided Cost	$59,140,572
NTG Contribution from Nonparticipant Spillover	**7.00%**

There are some notable disadvantages in using self-reported data. One is that it relies on information respondents can remember. It is often the case that respondents recall recent changes with greater clarity, and changes from further back often have incomplete or inaccurate data. Further, in commercial settings, staff turnover can also limit the information available to more recent timeframes. Another trouble spot in this methodology is respondents' ability to state honestly and accurately what their behavior would have been in the absence of the program. Hypothetical questions such as these are difficult to answer with certainty.

One of the primary advantages to this method is the ability to use all of the data points. Unlike the net billing model, no matter how small or large a single customer is, each data point can be used in the analysis. In addition, the straightforward methodology allows it to be used in a wide variety of situations, where other more complex statistical methods may not easily apply.

Overall, the self-report method will tend to underestimate the incidence of spillover. First, adoptions will be missed due to incomplete information from respondents. In addition, those who the supply side affect are excluded from the analysis. For example, a respondent may have been influenced by vendor stocking practices, or a recommendation from a contractor who was influenced by the program, and is probably not aware the program contributed to his/her choice of energy efficient equipment. Therefore, requiring customers to state that the program influenced them to choose energy efficient equipment is likely to provide a lower bound estimate of the true spillover rate.

Discrete Choice Method

The third method of estimating the net-to-gross ratio is a two-stage discrete choice model. This model is used to simulate the decision to purchase various types of commercial equipment. Once estimated, the model is used to determine the probability of purchasing high-efficiency equipment in the absence of the program.

The probability of purchasing any given equipment option A can be expressed as the product of two separate probabilities: the probability that a purchase is made multiplied by the probability that equipment option A is chosen given that a purchase has been made. This can be written as:

*Prob (Purchase & Equipment A) = Prob (Purchase) * Prob (Equipment A | Purchase)*

The two-stage model adopted for this analysis estimates both of the right hand side probabilities separately. The first stage of the model estimates the probability that a customer makes a lighting equipment purchase and is referred to as the **purchase probability**. The second stage of the model estimates the type of lighting equipment chosen given that the decision to purchase has already been made and is referred to as the **equipment choice probability**. The product of the purchase probability and the equipment choice probability is the **joint probability** and reflects the probability that any one lighting equipment option is purchased. Once estimated, the model is used to determine the probability of purchasing high-efficiency equipment in the absence of the lighting program. This is simulated by setting the rebate and program awareness variables to zero in both stages of model.

The net-to-gross ratio is calculated using the probability of purchasing high-efficiency lighting equipment both with and without the existence of the retrofit program. The expected impact with the program is the joint probability of choosing high-efficiency equipment multiplied by the energy impact of the equipment. Similarly, the expected energy impact in the absence of the Lighting Program is the joint probability of purchasing high-efficiency equipment without the program multiplied by the energy impact of the equipment. The net-to-gross ratio is the net savings due to the program divided by the expected energy impact that results from having the program.

The purchase decision is specified as a logit model with a dependent variable having a value of either zero or one. In this application, customers are given a value of one if they made a lighting equipment purchase either in or outside the program and a zero if they did not purchase any lighting equipment. The purchase decision model specification is defined as:

$$\text{PURCHASE} = \alpha + \beta' X + \gamma' Y + \vartheta' Z + \varepsilon$$

The explanatory variables X contain information on rebate and program awareness that capture the effect of the lighting program. Building characteristics such as square footage and changes to the facility are contained in Y. Variable group Z contains variables indicating building type and type of lighting. The error term ε is assumed to follow a logistic distribution, consistent with the logit model specification.

The probability of making a lighting equipment purchase in the absence of the program is calculated by setting the rebate and program awareness variables equal to zero. The probability of making a lighting purchase is then recalculated using the logistic density function.

The estimated probabilities for different customer groups are shown below for the 1996 and 1997 program year evaluations. The two years' results are consistent with each other, and generally conform to expectations. Lighting program participants have a high probability of making an equipment purchase, with an estimated purchase probability of 0.70 and 0.74, for 1996 and 1997 respectively. Conversely, those that did not make any purchases have a low estimated purchase probability, at 0.20 for both years.

Table 4. 1996 Estimated Purchase Probabilities

Customer Group	With Program	Without Program
No Purchase	0.20	0.14
Participants	0.70	0.41
Purchase HE Outside Program	0.45	0.23
Purchase Std Efficiency	0.28	0.20

1997 Program Estimated Purchase Probabilities

Customer Group	With Program	Without Program
No Purchase	0.20	0.15
Participants	0.74	0.28
Purchase HE Outside Program	0.45	0.24
Purchase Std Efficiency	0.30	0.20

The second stage of the model is the estimation of the equipment choice, given the decision to purchase new equipment has already been made. The equipment choice decision is modeled using a conditional logit specification as follows.

$$\text{EQUIPMENT CHOICE} = \beta'\text{AWARE} + \beta'\text{PREDISP} + \beta'\text{SQFEET} + \beta'\text{CINDEX} + \beta'\text{SAVINGS} + \Sigma\, \beta'\text{BLDTYPE} + \varepsilon$$

Where,
AWARE = Awareness variables

PREDISP = Predisposition towards high efficiency equipment[4].
CINDEX = (cost – rebate) / cost
SAVINGS = Annual dollar amount of electricity savings expected from equipment
BLDTYPE = Building size and business type
ε = Random error term assumed logistically distributed.

Once both the purchase probability and the equipment choice probability are estimated, the two probabilities are multiplied together to determine the joint probability that a purchase is made and that an individual equipment option is selected. This joint probability is calculated twice. First, the joint probability is calculated using the original values for the program awareness variables and rebate variables. This gives the joint probability with the existence of the program. Next, the joint probability is calculated in absence of the program. This is done by setting the awareness and rebate variables equal to zero.

To calculate expected impacts, the joint probability of making a purchase with the program is multiplied by the gross impact associated with the technology. The expected impact is then summed across the high efficiency equipment options to get a total expected impact for each customer. The calculation is given by:

$\text{EXPECTED IMPACT}^{W} = \Sigma\, P^{W}_{j} * \text{IMPACT}_{j}$

Where P^{W}_{j} = Joint probability of choosing equipment option j with the program

IMPACT_{j} = One year impact associated with equipment option j.

The expected impact without the program is calculated in the same manner using the joint probability in absence of the program:

$\text{EXPECTED IMPACT}^{WO} = \Sigma\, P^{WO}_{j} * \text{IMPACT}_{j}$

Where P^{WO}_{j} = Joint probability of choosing equipment option j without the program.

The net impact associated with program is simply the difference in expected impacts with and without the program:

$\text{NET IMPACT} = \text{EXPECTED IMPACT}^{W} - \text{EXPECTED IMPACT}^{WO}_{j}$

The net-to-gross ratio is then the net impact divided by the expected impact with the program:

NTG = NET IMPACT / EXPECTED IMPACT

The estimated net-to-gross ratios by building type are shown for the 1996 and 1997 program year evaluations in Table 5 below. The Table shows notable variation in some of the building type results between the two years. However, the overall rates estimated for 1996 and 1997 were identical to two decimal places, 0.82.

[4] This variable is a function of respondent's answers to questions regarding their attitudes towards high efficiency, and is included to control for free-ridership effects.

Table 5. 1996 Estimated NTG Ratios by Building Type

Building Type	NTG
Office	1.01
Retail	0.74
College	0.69
School	0.77
Grocery	0.76
Restaurant	0.81
Health	0.72
Hotel	0.70
Warehouse	0.78
Personal	1.04
Comm. Services	0.66
Misc. Comm.	0.98

1997 Estimated NTG Ratios by Building Type

Building Type	NTG
Office	0.81
Retail	0.82
College/univ	0.83
School	0.59
Grocery	0.91
Restaurant	0.85
Healthcare	0.81
Hotel	0.84
Warehouse	0.80
Personal Service	0.84
Community Service	1.07
Misc. Com.	0.92

This method does not rely on historical billing data, or on respondents' correctly reporting the degree to which the program influenced their behavior, although it does require respondents to correctly state their awareness of the program. In addition, the model controls for the effects of business type, building characteristics, size, and other influencing factors to better isolate the effects of the program. The discrete choice methodology also excludes supply side effects of the program. If a respondent is unaware of the program, no net benefits will be attributed to their actions. A respondent may be influenced by supply side program effects, but remain unaware of the program. This methodology will overlook any impacts associated with such a respondent. In addition, a drawback of this approach is that it requires that costs and savings be assigned to all non-chosen equipment alternatives. This requirement introduces another potential source of measurement error.

Conclusions

As discussed above, three separate models were implemented to estimate the components of the net-to-gross ratio (free ridership and spillover). Table 6 below present the 1996 and 1997 results of each method, by business type and for the total program. Results, both within business type and overall, are weighted by the ex-post gross energy impacts. Results are presented for the total net-to-gross ratio, as well as the two primary components, free ridership and spillover. For the Mills ratio methodology, only free ridership is estimated, as discussed previously.

Table 6. 1996 Comparison of Net-to-Gross Ratios

	Discrete Choice Model			*Self Report*			*Mills*
Business Type	***NTG***	***1-FR***	***Spill***	***NTG***	***1-FR***	***Spill***	***1-FR***
Office	1.01	0.65	0.36	0.88	0.78	-	0.84
Retail	0.74	0.73	0.01	0.86	0.75	-	1.03
College/Univ	0.69	0.68	0.01	0.84	0.74	-	0.96
School	0.77	0.75	0.02	0.84	0.74	-	0.97
Grocery	0.76	0.76	0.00	0.85	0.74	-	1.06
Restaurant	0.81	0.81	0.00	0.84	0.73	-	1.05
Health Care	0.72	0.72	0.00	0.84	0.73	-	0.82
Hotel/Motel	0.70	0.68	0.01	0.68	0.58	-	1.41
Warehouse	0.78	0.74	0.04	0.92	0.81	-	1.14
Personal Svcs.	1.04	0.79	0.26	0.85	0.75	-	1.05
Comm. Svcs.	0.66	0.64	0.02	0.85	0.74	-	0.92
Misc.	0.98	0.73	0.25	0.88	0.78	-	1.08
Total	**0.82**	**0.71**	**0.10**	**0.85**	**0.75**	**0.10**	**0.99**

1997 Comparison of Net-to-Gross Ratios

	Discrete Choice Model			*Self Report*			*Mills*
Business Type	***NTG***	***1-FR***	***Spill***	***NTG***	***1-FR***	***Spill***	***1-FR***
Office	0.81	0.76	0.05	0.88	0.78	0.10	0.91
Retail	0.82	0.80	0.02	0.81	0.71	0.10	0.76
College/Univ	0.83	0.78	0.05	0.82	0.72	0.10	0.61
School	0.59	0.58	0.01	0.84	0.74	0.10	0.61
Grocery	0.91	0.67	0.24	0.84	0.74	0.10	1.21
Restaurant	0.85	0.83	0.01	0.82	0.72	0.10	0.76
Health Care	0.81	0.80	0.01	0.86	0.76	0.10	1.21
Hotel/Motel	0.84	0.82	0.01	0.75	0.65	0.10	0.76
Warehouse	0.80	0.79	0.01	0.79	0.69	0.10	1.21
Personal Svcs.	0.84	0.78	0.06	0.86	0.76	0.10	1.06
Comm. Svcs.	1.07	0.79	0.28	0.78	0.69	0.10	0.91
Misc.	0.92	0.82	0.10	0.78	0.68	0.10	1.21
Total	**0.82**	**0.76**	**0.05**	**0.84**	**0.74**	**0.10**	**0.87**

The results of the 1996 and 1997 evaluations are markedly consistent. In particular the overall net-to-gross ratio for the discrete choice and self-report methods differ between the two studies by less than one percent. The self-report results for free ridership and spillover individually are also very consistent across the two studies; also coming within one percent of each other. There is somewhat more variation between the two studies when one examines individual business type results, as well as for the Mills ratio approach. These variations are likely due to smaller and varying sample sizes.

For each year individually, it is clear that the discrete choice model was well validated by the self-report results. The total net-to-gross ratio was within five percent of the self-reported results in both years. Even at the business type level, the self-report results were within 20% of the discrete choice model results for all but two business types, in both years. Much of this variation can be attributed to the fact that the spillover estimates for the self-report approach were not estimated at the business type level. Rather, a single estimate of spillover was estimated.

An examination of the free ridership estimates among the discrete choice and self-report models at the business type level also provides strong validation for the two sets of results. The 1996 evaluation produced self-reported results that were within 10% of the discrete choice model results for all but three business types. In 1997, the self-reported results were within 20% of the discrete choice model results for every business type but one.

The free ridership estimates generated using the Mills approach yielded somewhat higher estimates of net participation in both years, relative to the other methods. This result supports the contention that censoring large customers from the analysis may bias free ridership estimates downward, as larger customers are more likely to be free riders.

The consistency of the 1996 and 1997 results across the three different approaches bodes well for each individually. The Mills ratios lack the estimate of spillover, and are also run on a reduced set of the data due to the censoring of large customers and customers without good billing data. The self-report values rely on customers to give accurate and unbiased responses to their hypothetical actions in the absence of the program. The most sophisticated, and preferred, of the three approaches is the two-stage discrete choice model, but significant information is derived from the use of multiple approaches. Each approach has advantages and disadvantages, which provide valuable checks and balances.

As a final note, many utilities are becoming more and more interested in issues surrounding market transformation. Fundamentally, the challenge of measuring spillover and free ridership effects is estimating what total energy use would have been in the absence of an existing program. This is the same question facing those who are interested in measuring market transformation effects. All three of the methodologies presented here could be utilized to measure market transformation effects of end-user behavior. Ideally, to capture the market transformation effects on customer purchases, an 'outside service territory' control group would be used in addition to an "in-service territory" nonparticipant sample.