Advancing the "Science" of Free Ridership Estimation: An Evolution of the Self-Report Method for New York Energy $mart[SM] Programs

Stuart Schare, *Summit Blue Consulting*
Jennifer Ellefsen, *NYSERDA*[1]

INTRODUCTION

Evaluations of energy efficiency programs often utilize self-reporting by program participants to estimate free ridership. While this method can be simpler and less expensive than other approaches, such as conducting a billing analysis with a control group, it has inherent drawbacks such as non-response bias, limited respondent recall of program influence on decision-making, and potential investigator bias in translating responses into partial free ridership values. However, as evaluators have refined the interview techniques and analytic methods, estimation of free ridership from self-reporting has become more of a reliable, systematic approach to quantifying program impacts.

This paper presents the approach used in 2006 to estimate free ridership in several of the business/institutional and residential energy efficiency programs administered by the New York State Energy Research and Development Authority (NYSERDA). This approach has evolved from prior evaluations in New York, as well as methods employed by evaluators in New England and California. The free ridership estimation method relies on experienced interviewers who are knowledgeable enough to probe respondents for details of program influence and who can characterize the responses in quantitative terms that can be used in the analysis. The core of the approach includes the following elements:

- Direct estimates from respondents of the likelihood of installations and/or the share of efficient measures that would have been installed without the assistance of the program;
- Quantitative and open-ended responses to questions regarding the influence of the program on equipment investment decisions;
- Scoring of open-ended responses by experienced interviewers using established guidelines to quantify the level of program influence; and
- Adjustments to direct free ridership estimates based on program-influence scores from both the respondents and the interviewers.

The underlying concept is that only impacts "caused" by the program should be included in the final program savings estimate, net of free ridership. However, absolute proof of causality is unattainable since one can never observe what would have happened (*i.e.,* what the efficiency of installed equipment would have been) in the absence of the program. Consequently, causality "must be justified or rationalized on the basis of *a priori* argument, outside evidence, intuition, theory, or some other informal means."[2] The necessity of this approach to attribution analysis, relying in part on intuition and outside assumptions, is supported by Heckman in his argument that 'there is no mechanical algorithm for producing a set of 'assumption free' facts or causal

[1] The views expressed in this paper are those of the authors and do not necessarily reflect the views of the New York State Energy Research and Development Authority.

[2] See Moffitt, R., "Causal Analysis in Population Research: An Economist's Perspective," Johns Hopkins University, 2003.

estimates based on those facts."[3] In the context of energy efficiency program evaluation, Wirtshafter and Sorrentino add that "any proof [of program influence] available will be, at best, circumstantial."[4] This is not only true for assessing attribution in the context of an energy efficiency program, but it is true for any policy assessment including those pertaining to education programs, health programs, fiscal policies and virtually any assessment of a business decision that purports to have a cause and effect.

By presenting details of the free ridership estimation method and a summary of results as the method was applied to NYSERDA's Loan Fund reduced-interest loan program, the authors hope to advance the understanding of an innovative approach to free ridership estimation for energy efficiency programs that builds on past practices and offers benefits that offset many of the typical drawbacks to the self-report method.

METHOD FOR FREE RIDERSHIP ESTIMATION

The methods discussed below are presented with respect to NYSERDA's Loan Fund reduced-interest loan program, but can be applied more generally to virtually any efficiency program—and, in fact, these methods have been used for Wisconsin Power & Light's Shared Savings program and for other NYSERDA programs. Estimates of savings attributable to the program are based on the responses of program participants to carefully derived questions (building on established work within the field)[5] regarding prior intentions, the importance of program factors such as financial incentives, and the likelihood that the same actions would have been taken even without the program. Based on methods employed in program evaluations in other states and on the experience of the evaluation team, each actor's responses were assessed in a systematic manner to produce estimates of free ridership.

Data were gathered through a series of survey questions asked of borrowers who participated in the program. Views on free ridership were asked in both direct questions—aimed at obtaining respondent estimates of the appropriate (partial) free ridership rate that should be applied to them—and in supporting, or influencing, questions that could be used to verify whether the direct responses were consistent with participants' views of the program's influence. Figure 1 presents the flow of the free ridership analysis, illustrating how the direct and influencing questions are used to generate a single free ridership estimate for each respondent.

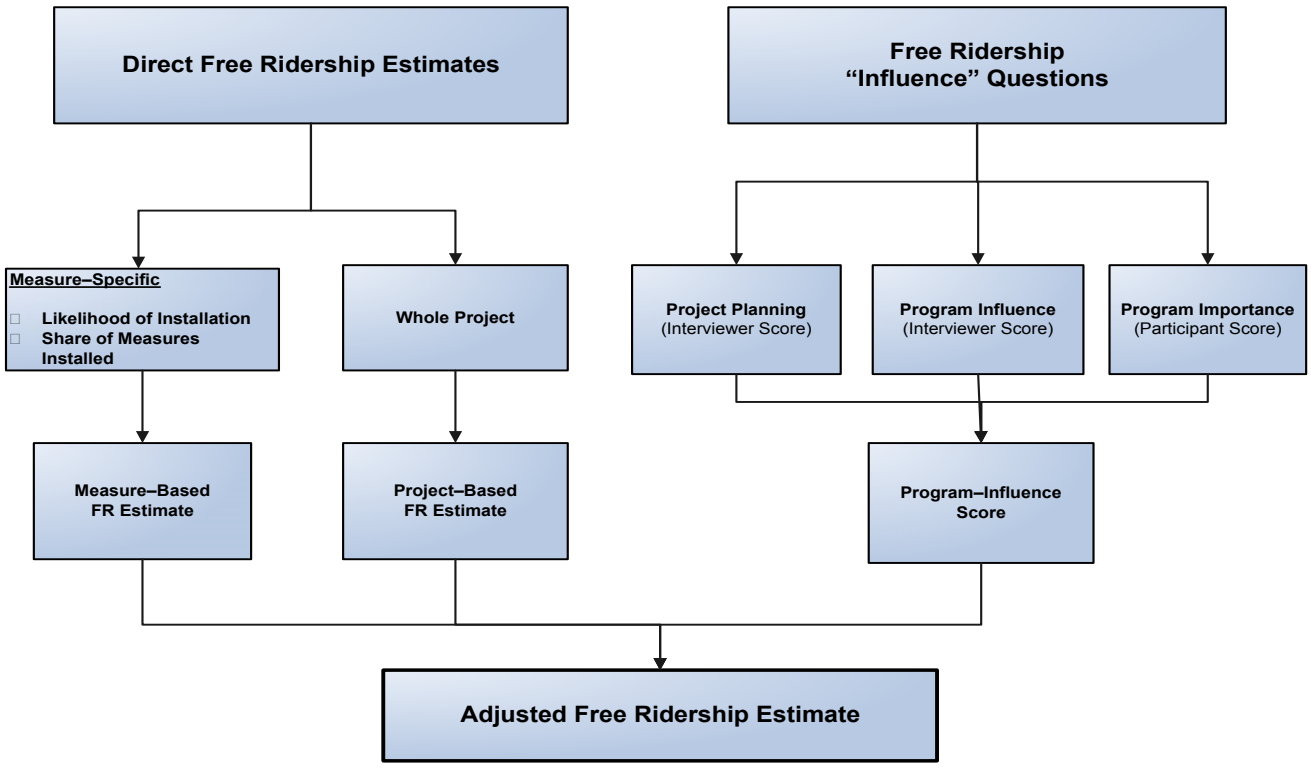
---

[3] Heckman, J., "Causal Parameters and Policy Analysis in Economics: A Twentieth Century Retrospective," *The Quarterly Journal of Economics*, Volume 115, No. 2, 2000, pp. 45-97.

[4] Wirtshafter, Robert M. and A. Sorrentino, "Proving Causality in Market Transformation Programs: Issues and Alternatives," *Proceedings of the 1994 ACEEE Summer Study on Energy Efficiency in Buildings,* American Council for an Energy-Efficient Economy, Volume 10, 1994, pp. 259-265.

[5] In particular, the evaluation team reviewed recently adopted evaluation frameworks from California and New England. See *Protocols for the Evaluation of Post 2005 California Energy Efficiency Programs,* August 2005 and *The California Evaluation Framework*, June 2004, both prepared for the California Public Utilities Commission by TecMarket Works. Also see *Standardized Methods for Free-Ridership and Spillover Evaluation—Task 5 Final Report (Revised),* prepared by PA Government Services, Inc. for National Grid, NSTAR Electric, Northeast Utilities, Unitil, and Cape Light Compact, June 2003.

Figure 1. Flow Chart of Free Ridership Analysis



Direct Free Ridership Questions

The direct free ridership questions separately addressed the following aspects of the projects:

1. Each major category of measures that were reported to the program (*e.g.,* HVAC, lighting, building controls), and

2. The project as a whole across all measure categories.

For the measure-specific questions, respondents were first asked when, if at all in the foreseeable future, they would have replaced existing equipment or installed new equipment if they had not participated in Loan Fund. Respondents were then asked to estimate the *likelihood* that they would have incorporated measures "of the same high level of efficiency" if not for the financial and technical assistance of the program. In cases where respondents indicated that they may have incorporated some, but not all, of the measures, they were asked to estimate the *share of measures* that would have been incorporated anyway at high-efficiency. This flexibility in how respondents could conceptualize and convey their views on free ridership allowed respondents to give their most informed answer, thus improving the accuracy of the free ridership estimates.

Additional direct free ridership questions were also asked to obtain a lower bound, an upper bound, and a best estimate of the overall energy savings attributable to the program across all measure categories. These questions focused on the incremental savings from installing high-efficiency equipment instead of standard-efficiency equipment. The questions were asked after the measure-specific questions so that respondents would have fresh in their minds the decisions they made on individual measures.

Free Ridership "Influence" Questions

In addition to the "direct" free ridership questions, the surveys included a number of "program influence" or supporting questions in order to clarify the role that program interventions (*e.g.*, financial incentives and technical assistance) played in decision-making, and to provide supporting information on free ridership. Responses to these questions were analyzed for each respondent, not just in aggregate, and were used to identify whether the direct responses on free ridership were consistent with how each respondent rated the "influence" of the program. Where significant program influence was identified, free ridership estimates were reviewed and potentially adjusted via a "rules-based" approach that placed bounds on the final free ridership estimate.

Questions addressing the following topics were included to reflect the role of the program in decisions to install higher efficiency measures:

- The borrower's plans (or lack thereof) prior to participating in the program to install the energy efficiency equipment included in the project
- The program's influence on the type or efficiency level of the equipment, or the amount of high efficiency equipment, included in the project
- The importance of the program, including technical and financial assistance, in the decision to incorporate high-efficiency equipment in the project.

Plans to Incorporate High Efficiency

Each respondent indicating any degree of planning for high efficiency prior to participating in the program was asked to describe these plans in detail and was probed for the equipment type, timing, quantity, and efficiency, as well as for any prior budgeting for the high efficiency equipment. Based on the response to this inquiry, the interviewer then assigned a "project planning" score using a 5-point scale and established guidelines to ensure consistency across respondents and interviewers (Table 1).

Table 1. Guidelines for Assigning High-Efficiency Project Planning Score

Score	Extent of Planning
1	No plans for high-efficiency equipment; respondent may have considered alternative technology options, but did not explicitly consider high efficiency.
2	Initial steps toward consideration of high efficiency such as requesting information on or discussing, in general, high efficiency options with vendors or contractors.
3	In-depth discussion or consideration of specific types of high efficiency equipment (*e.g.*, lighting controls, high-efficiency chillers), including their positive and negative attributes and costs.
4	Identification of specific equipment manufacturers and models, including assessment of their relative costs and performance characteristics.
5	High efficiency equipment and designs fully specified and explicitly approved or incorporated into project budget.

These guidelines were intended to assist interviewers in assigning project planning scores, but were not prescriptive since the descriptors do not necessarily apply to all projects and may not build upon one another sequentially. Interviewers were given the latitude to make reasoned judgments based on the specific claims of the respondents.

Program Influence on Project

Respondents were also asked whether their participation in the Loan Fund in any way influenced the type or efficiency level of the equipment, or the amount of high efficiency equipment, installed in the project. Each respondent indicating some degree of program influence was asked to describe how the program influenced the decision to install high-efficiency equipment in the project. Similar to the "project planning" score described above, the interviewer then used this response to assign a "program influence" score using a 5-point scale according to established guidelines (Table 2).

Table 2. Guidelines for Assigning Program Influence Score

Score	Characterization of Program Influence
1	No influence on the decision to install high-efficiency equipment. All equipment would have been installed at the same efficiencies even without the program.
2	Program helped in making final decision on equipment that had already been thoroughly considered.
3	Program lent credibility to the decision to invest in high efficiency and/or it provided information that helped expande the quantity, scope, or efficiency of the equipment.
4	Program identified a significant number of specific high efficiency options that were installed but that had not previously been considered and/or program were a major driver behind a significant increase in the quantity, scope, or efficiency of high-efficiency equipment.
5	Program was the primary reason that high efficiency equipment was installed in the project.

Similar to the guidelines for program planning, these guidelines were intended to assist interviewers in assigning program influence scores, but interviewers were given the latitude to make reasoned judgments based on the specific claims of the respondents.

Importance of Program in Project Decision-Making

In the final influence question, respondents were asked to provide their own assessment of the program's importance in the decision to install equipment at the high level of efficiency that was used at the project site. Similar to the previous program influence questions, a 5-point scale was used, with 1 indicating the program was "not at all important" and 5 indicating the program was "very important."

The general approach holds that if a borrower were not definitively planning to install all of the efficiency measures prior to participation, and if the program influenced decisions in the course of the project, then the program can reasonably be credited with at least a portion of the energy savings resulting from the high-efficiency equipment. Moreover, responses to all of the program influence questions provide an indication of the degree to which the energy savings should be attributed to the program.

Using the Survey Responses to Estimate Free Ridership

Direct Free Ridership Estimate

The direct free ridership estimate is based on both the measure-specific questions and the "whole project" questions. For each measure category for which the respondent had installed equipment through the program, the survey collected information on 1) when, if ever, the equipment would likely have been installed and 2) the *likelihood* that the same high efficiency equipment would have been used, or the *share of high-efficiency measures* that would have been installed. The response to the likelihood/share-of-measures questions were used as the initial free ridership

value for the measure category. This value was then discounted if the respondent indicated that the program influenced them to install the equipment more than one year earlier than they otherwise would have. The specific discount values (*i.e.,* adjustment multipliers) are presented in Table 3. For example, if a respondent indicated a 50% likelihood of installing high efficiency lighting, but not for another two years, then the 75% adjustment multiplier would reduce the direct free ridership estimate for lighting to 50% x 75% = 37.5%.

Table 3. Early Replacement Adjustment Multipliers

Early Replacement *Within ____ years of program participation*	**Adjustment Multiplier**
Within 1 year	100%
1+ to 2 years	75%
2+ to 3 years	50%
3+ to 5 years	25%
More than 5 years	0%

Each measure category was also assigned an energy savings value (in Btu as a common unit across electric and gas measures) from the savings recorded for that respondent in the program database. The direct free ridership estimate for each measure category (after any adjustment for early replacement) was weighted according to the relative savings from the category to determine a weighted average free ridership estimate across all measures.[6] This estimate formed one-half of the direct free ridership estimate for commercial borrowers.

A second "direct" free ridership estimate was determined based on answers to the direct free ridership questions regarding the lower bound, upper bound, and best estimate of the overall energy savings attributable to the program across all measure categories. If a "best estimate" was provided, this value was used as a second direct free ridership estimate in addition to the measure-based estimate discussed above. The final direct free ridership estimate, which would potentially be adjusted according to responses to the program influence questions, was the simple average of the measure-based estimate and the "best estimate." If sufficient information were available for only one of these values, then this value was used as the final direct free ridership estimate.

Adjusted Free Ridership Estimate

As previously discussed, additional questions were included in the surveys to support an analysis of the consistency of responses. Responses to these "program influence" questions were used to adjust the "direct" free ridership estimates using objective criteria described below. If made at all, adjustments are made to individual respondents' free ridership estimates—not to the aggregate free ridership value across respondents. As a general rule, adjustments were only made if one or both of the following conditions are true: 1) the respondent's average score for the three 1-to-5-scale questions was a 4.0 or greater, and/or 2) the respondent's initial free ridership score was beyond the bounds that could reasonably be expected based on responses to the influence

[6] For example, assume the lighting free ridership estimate is 50% and the HVAC free ridership estimate 0%. Further, assume lighting accounts for 80% of the project savings, and HVAC for 20%. The direct free ridership estimate for the project would be [50% x 80%] + [0% *20%] = 40%. If savings data were not available, then no measure-specific direct free ridership estimate was calculated.

questions. Specifically, the process for whether and by how much to adjust a respondent's direct free ridership estimate was as follows:

Step 1. Calculate an average program influence score (on a 5-point scale) from the scores assigned to the three sets of program influence questions regarding project planning, influence of the program, and the program's importance. A higher score for program influence and importance suggests greater program impact, but a higher score for planning indicates lower impact. Therefore, prior to calculating an average score across the three sets of questions, the planning score was inverted so that 1=5, 2=4, etc. In this way, a higher average score across these questions unequivocally represents greater program impact.

Step 2. Adjust the direct free ridership estimate within respondent's lower and upper bounds. Respondents were asked to provide a range of free ridership estimates for the project as a whole. This range can be used in conjunction with the program influence questions to adjust the direct free ridership estimate to better reflect respondents' views of the impact of the program on their decisions to install high efficiency equipment. If a respondent's average influence score is toward the extremes (4.0 or higher, or 2.0 or lower), then the direct free ridership estimate is adjusted to reflect the lower or upper bound values provided by the respondent. For example, if a respondent's score is the maximum possible value of 5.0 (implying that the program was very influential), then the lower bound free ridership estimate would be used. For successively lower scores, a free ridership value between the initial value and the lower bound would be assigned. For example, for a score of 4.5, the initial estimate would be averaged with the lower bound; in other words the adjusted value would be 50% of the way between the initial estimate and the lower bound value. Table 4 presents values, for each possible program influence score, representing the degree to which the initial free ridership estimate is adjusted toward the respondent's lower or upper bound values.

Table 4. Free Ridership Adjustment Within Respondent's Lower and Upper Bounds

Average Program Influence Score	**1.00**	**1.33**	**1.67**	**2.00**	**2.33 to 3.67**	**4.00**	**4.33**	**4.67**	**5.00**
Adjust FR value…	100%	75%	50%	25%	No adjustment made	25%	50%	75%	100%
…of the way toward the _____ bound.	Upper				No adjustment made	Lower			

Step 3. Further adjust the free ridership estimate within reasonable bounds based on the average program influence score. These bounds are intended to reflect the range of free ridership values that the evaluation team believed could reasonably characterize a Loan Fund project based on a respondent's answers to the program influence questions. If a respondent's score is the maximum possible value of 5.0 (implying that the program was very influential), then a reasonable free ridership value would be as low as 0%, and it is assumed that a direct free ridership estimate higher than 25% would be inconsistent with responses to the program influence questions. For successively lower scores, the range of reasonable free ridership values increases to successively higher values. Table 5 presents the full set of lower and upper bounds that are deemed reasonable for each possible average program influence score.

Table 5. Range of Reasonable Free Ridership Values Based on Program Influence Responses

Average Program Influence Score	1.00	1.33	1.67	2.00	2.33	2.67	3.00	3.33	3.67	4.00	4.33	4.67	5.00
Lower Bound Free ridership Value	75%	70%	60%	50%	40%	30%	25%	20%	10%	0%	0%	0%	0%
Upper Bound Free ridership Value	100%	100%	100%	100%	90%	80%	75%	70%	60%	50%	40%	30%	25%

After any necessary adjustment for the respondent's lower and upper bound estimates in Step 2 above, the free ridership estimate was compared to the appropriate range of values, according to the average program influence score, in Table 5. If the free ridership value fell outside of the bounds, then it was adjusted to a final free ridership estimate equal to the closest lower or upper bound value.

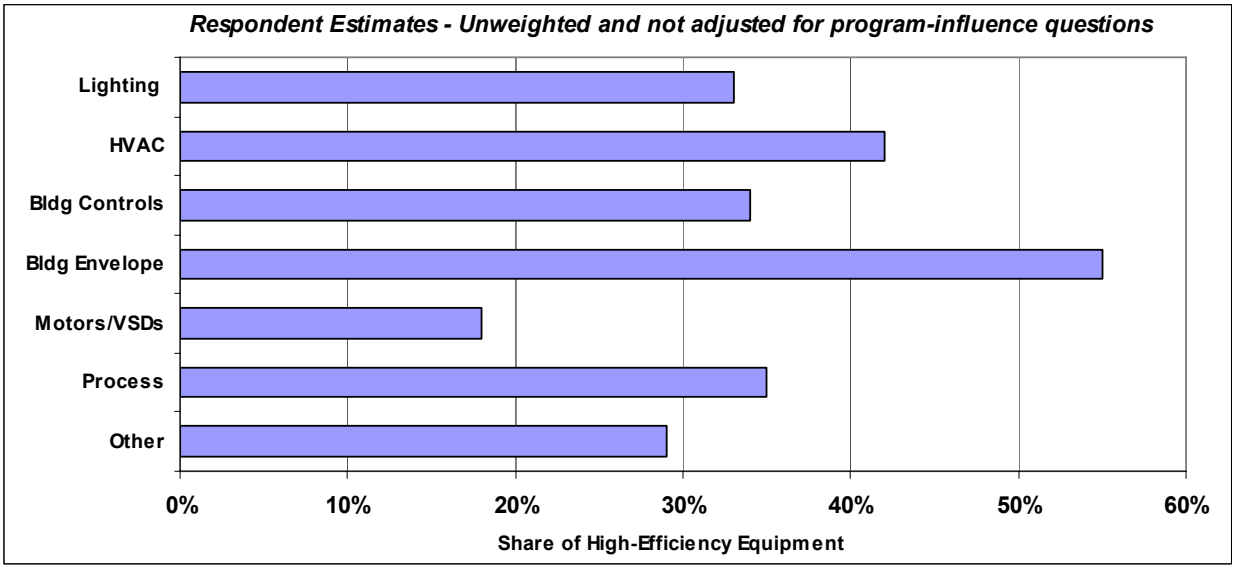
FREE RIDERSHIP RESULTS FOR THE LOAN FUND PROGRAM

The free ridership estimation method described above was applied in 2006 to NYSERDA's Loan Fund program, which offers reduced-interest loans to residential and commercial customers for purchase of high-efficiency equipment. In order to maintain simplicity and to better illustrate how the method is used, results of the analysis are discussed below for commercial borrowers only. Findings are based on telephone interviews with 61 commercial customers who participated in the program.[7]

According to the analysis, the measures most likely to be installed even without the Loan Fund were those related to building envelope (including doors, windows, air/duct sealing, and insulation) and HVAC equipment. Respondents indicated that more than 40% of the high-efficiency equipment in these categories would have been installed anyway (Figure 2). As discussed above, these estimates are based on questions specific to each measure category regarding the likelihood, scope, and timing of the investments in energy efficiency if the respondent had not participated in the program.

[7] See *Loan Fund Program - Market Characterization, Market Assessment and Causality (MCAC) Evaluation Final Report,* Summit Blue Consulting for NYSERDA, 2006.

Figure 2. Direct Estimates of High-Efficiency Installations without the Loan Fund, by Measure Category



Free ridership for the whole project (*i.e.,* across all measures) was estimated at 33% based on responses for the individual measure categories discussed above. For each respondent, this approach weighted the measure-specific free ridership values according to their relative energy savings as documented in program records. A similar free ridership value, 32%, was estimated in response to the series of questions culminating in respondents' "best estimates" of the share of energy savings that would have been achieved in the absence of the program.

As discussed above, "program influence" questions were used to adjust respondents' direct free ridership estimates. Respondents indicated that some energy efficiency measures were being planned, at least in part, for approximately two-thirds of all projects prior to participation in the Loan Fund (Table 6). Despite what may seem to be a significant amount of pre-program planning by participants, however, more than one-third of Loan Fund projects were not being planned at all. Furthermore, nearly three-quarters of all projects had been planned to only a moderate degree at most (3 or less on the 5-point scale), which generally indicates that the borrowers took some initial steps toward acquiring high-efficiency equipment—such as discussing energy efficiency alternatives with a contractor—but had not reviewed specific options in detail.

Table 6. Planning for High Efficiency Prior to Program Participation

	Share of Respondents
Share of borrowers planning to install at least some of the high-efficiency equipment prior to participation	67%
Level of planning for high efficiency *Average Score* ➔	2.51
No plans = 1	36%
2	16%
3	20%
4	16%
Plans fully developed/documented = 5	11%

Note: Total of "1 to 5" scores do not add to 100% due to rounding of individual percentage values.

Beyond planning issues, the other two sets of influence questions indicate that the Loan Fund significantly influenced borrowers in selecting high-efficiency equipment. Two-thirds of commercial borrowers responding to the survey report that the program in some way influenced "either the type or efficiency level of the equipment…or the amount of high efficiency equipment" installed (Table 7). For example, one borrower reported purchasing air conditioning equipment at a higher SEER level in order to participate in the Loan Fund, and another indicated that the Loan Fund allowed them to select high-efficiency equipment by improving what would have been only marginal paybacks.

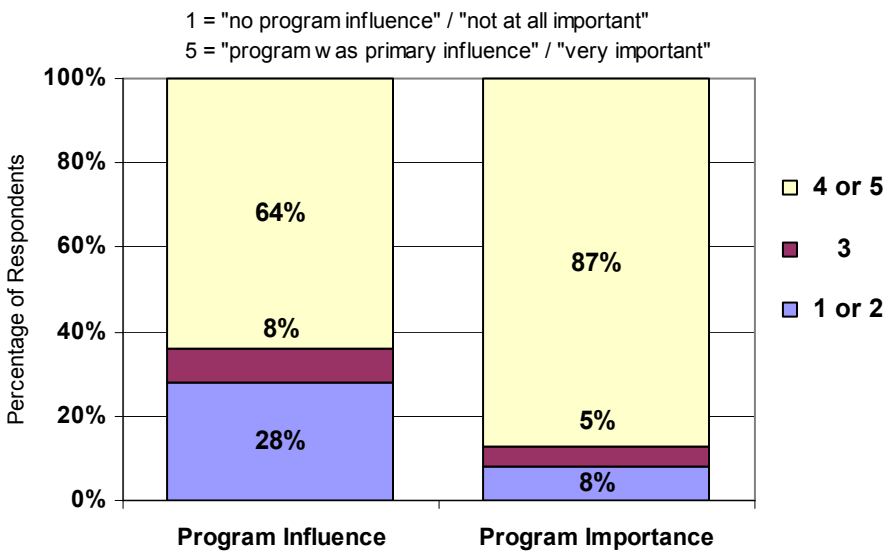
Overall, nearly two-thirds of commercial borrowers surveyed received influence scores at the upper end of the scale, implying that the program was a primary reason for the selection of high-efficiency equipment. As shown in Table 7, the respondents themselves rated the program influence even higher, with 87% assigning a 4 or 5 rating to the program's importance "in your decision to install energy efficiency measures" at the project site.

Table 7. Program Influence on Decisions to Incorporate High Efficiency Measures

	Share of Respondents
Program influenced type/efficiency of equipment or amount of high-efficiency measures *Share of respondents*	67%
Program influence on decision to install high-efficiency measures (Interviewer assessment) *4 or 5 on 5-point scale*	64%
Importance of program in decision to install high-efficiency measures (Respondent assessment) *4 or 5 on 5-point scale*	87%

To illustrate the significant influence of the Loan Fund, Figure 3 graphically presents the relatively high percentage of respondents with program influence scores at the upper end of the scale (4 or 5 on a 5-point scale). As discussed previously, the "Program Influence" rating is from the interviewers' assessments of respondents' answers to an open-ended question about program influence. The score for "importance" of the program on decision-making, as rated directly by respondents, shows an even more dramatic program impact, with 87% rating the Loan Fund a 4 or 5.

Figure 3. Influence of Loan Fund on Decision to Install High Efficiency Equipment



The figure suggests an overwhelming influence of the program on participating projects—a finding that would appear to be inconsistent with the initial free ridership estimates of more than 30%. A partial explanation from the interviews is that the program is, in fact, highly influential in the selection of equipment. However, in some cases the program was the critical influence for only some, but not all, of the measure categories for which equipment was installed; or the program may not be the sole driver but may be instrumental in the decision being made in a timely manner. Thus, a high level of program influence does not necessarily correspond to zero free ridership; and, in any case, the various financial, corporate, and personal considerations involved in making specific investments in energy efficiency makes it difficult for respondents to precisely estimate what would have happened in the absence of the program. The high degree of program influence indicated through multiple sets of questions—juxtaposed against respondents' estimates that roughly one-third of high-efficiency measures would have been installed even without the program—suggests that the most appropriate free ridership estimate would reflect an adjustment of the initial estimates to account for apparent inconsistencies.

Per the methods described above, potential adjustments to the direct free ridership estimates affected the final estimates for some respondents. To illustrate, consider a respondent whose average influence score is 4.5. Given the strong program influence indicated, the "best estimate" free ridership value would be averaged with the lower bound estimate to reflect that free ridership may be at the lower end of the range provided by the respondent (this adjustment only applies to commercial borrowers). Furthermore, the influence score suggests that free ridership should be fairly low, and can reasonably be expected to be no higher than 40%. Thus, if the respondent's free ridership value (after the initial adjustment above) is above this level, then the adjusted estimate would be reduced to 40%.

The impact of influencing factors on the free ridership estimate for Loan Fund borrowers is reflected in adjusted free ridership estimates of 27%, which is 5% to 6% below the direct free ridership estimates of 33% (based on the measure-specific method) and 32% (based on the project-wide estimate across all measures). The reduction from the direct free ridership estimate is due to the strong program influence that respondents indicated via the "influence" questions. It should be noted that the final free ridership rate presented here represents the energy savings that would have been achieved without the program. Furthermore, this rate reflects partial free

ridership across many of the respondents, and may not necessarily correspond directly to the percentage of program participants that are pure free riders.

CONCLUSIONS

The approach to free ridership estimation discussed above is based on the self-reporting method often used in program evaluation. However, it offers the unique benefits of a clearly defined and repeatable method for quantifying free ridership, while also incorporating qualitative information from program participants that is often used only as supporting illustration. The core principles embodied in the approach include the following:

1. **Direct estimation of free ridership from the perspective that is most appropriate for the project** and to which the respondent can best relate his program experience. This takes the form of either a "likelihood" that the high-efficiency measures would have been installed without the program or the "share" of high-efficiency measures that would have been installed without the program.
2. **Separate estimation of free ridership addressing the complete project across all measure types and, alternatively, addressing decisions to install specific measures.** The dual line of questioning allows respondents to provide a "big-picture" view of the program's influence on the project as well as to focus on specific measures, which may have been influenced by the program to varying degrees.
3. **Quantitative incorporation of qualitative responses based on interviewers' probing for details and causality.** This aspect of the approach relies on experienced interviewers who are able to apply appropriate judgment to assign "influence" scores reflecting the degree to which the program affected equipment-purchasing decisions.

In 2006, this free ridership approach was successfully applied in New York to the commercial New Construction Program and the Loan Fund, as well as Wisconsin Power & Light's Shared Savings program. Previous iterations of the approach have also been used in a variety of other **New York Energy $mart**[SM] program evaluations. This method overcomes one of the principal limitations of the self-report method—the difficulty of systematically converting the opinions of participating customers into quantifiable free ridership values. Importantly, this method also provides a highly defensible approach to estimating net program impacts, which have been a key input to benefit-cost analyses and policy decisions on the direction not only of NYSERDA's programs, but of energy efficiency programs across the country.