What Have We Learned about Success and Its Drivers in Comprehensive Residential Upgrade Programs?

Marjorie McRae, Ph.D., Research Into Action, Inc., Portland, OR
Jordan Folks, M.S., Research Into Action, Inc., Portland, OR

ABSTRACT

Evaluators are called on to assess program success and to identify factors contributing to successful and unsuccessful outcomes. Yet few studies have analyzed sufficient numbers of programs to identify the correlates of success and even fewer studies have attempted to demonstrate what can make or break *any* residential efficiency upgrade program – regardless of program design intricacies or the varying environments in which the program takes place. This paper attempts to answer these questions:

1. What defines a successful program?
2. What programmatic elements help avoid poor program performance, regardless of program design specifics or regional characteristics?
3. What programmatic elements lead to highly successful program outcomes, regardless of program design specifics or regional characteristics?

This paper explores 12 quantitative indicators of successful program outcomes and identifies the drivers and detractors of success based on data from a diverse set of 54 comprehensive residential upgrade programs from across the country. The programs were conducted by 41 state and local grantees in 32 states and territories that received federal funding for development and deployment of energy efficiency upgrade programs through the U.S. Department of Energy's Better Buildings Neighborhood Program (BBNP). Analysis of BBNP data reveals a robust metric for measuring relative program success and demonstrates that: a) contractor training helps prevent sub-par program outcomes and b) offering multiple pathways to participation and achievement of energy savings is critical to programmatic success. This paper fully presents our findings on the factors that define and drive success in widely varying comprehensive residential upgrade programs.

Introduction

From 2010 to 2013, the U.S. Department of Energy (DOE) administered the Better Buildings Neighborhood Program (BBNP) to support programs promoting whole building energy upgrades. BBNP distributed a total of $508 million to support hundreds of communities served by 41 grantees, with awards ranging from $1.4 million to $40 million per grantee. State and local governments received the grants and worked with nonprofits, building energy efficiency experts, contractor trade associations, financial institutions, utilities, and other organizations to develop community-based programs, incentives, and financing options for comprehensive energy saving upgrades. Each of the 41 grant-funded organizations targeted a unique combination of residential, multifamily, commercial, industrial, and agriculture sector buildings, depending on their objectives. Due to the inclusion of subgrantee-ran programs, the number of programs included in our analyses (n = 54) exceeds the number of primary BBNP grantees (N = 41). Thus, we analyzed data from 54 widely varying residential programs conducted by grantees and their subgrantees. For simplicity, throughout this paper we refer to both grantees and subgrantees as "grantees."

Using both data that grantees reported to DOE in partial fulfillment of their grant requirements and data collected by us, we conducted a series of statistical analyses to develop a quantitative definition of grantee success that corresponds to BBNP’s multiple program objectives and to identify program features and characteristics that predict success. We conducted analyses of program success for the single-family residential sector due to greater availability of data than for the nonresidential and multifamily sectors. Further, focusing on the single-family residential sector has significant merit; running a successful residential whole-house retrofit program constitutes a challenge for many program administrators.

Methods

Overview

Using 12 diverse indicators of success, we identified successful program outcomes – and their drivers – across a diverse set of comprehensive residential upgrade programs from across the country. The research began by defining numerical performance metrics corresponding to BBNP’s multi-faceted objectives. We then conducted latent profile analysis (LPA) to cluster programs into groups with similar performance on the 12 indicators of success. LPA is a statistical approach that aims to identify categories, or clusters, of entities (grantees and sub-grantees), based on continuous indicators (performance metrics). We used LPA to identify clusters of grantees that represent different domains or levels of success. We sought to identify clusters that were both theoretically sound and provided a valid representation of success among the BBNP grantees.

LPA revealed that programs clustered into three groups; their average group values on the 12 metrics were consistent with an interpretation of a most successful group, an average group, and a least successful group. After clustering programs into the groups described above, we used binary logistic regression to identify the respective factors that distinguished programs that fell into the least successful group and those that fell into the most successful group. The following sections further describe the methodology employed in our analysis of factors that drive or inhibit program success.

Defining Success Via Latent Profile Analysis

First, we identified quantifiable metrics of success for residential energy efficiency programs based on BBNP’s objectives and data availability. We then compiled performance metric data for each grantee and subgrantee and conducted LPA on the resulting dataset.[1] We used LPA as an exploratory approach to measuring relative grantee success – a comparative assessment based on the performance metrics we examined. While we had hypotheses as to how clusters of grantees may have demonstrated similar performance on the performance metrics, prior to executing the LPA we did not actually know how grantees would cluster together. Thus, LPA allowed us to assess if grantees fell into tiered levels of success or if they fell into clusters representative of different domains of success (e.g., a high cost effectiveness cluster, a large energy savings cluster, etc.). In sum, we used LPA to explore how grantees cluster along the performance metrics and subsequently defined the respective clusters based on their members’ average performance on the performance metrics in the LPA model.

[1] We conducted analyses of 2-, 3-, and 4-cluster models on the final set of twelve performance metrics. We found that the 3-cluster model yielded the most parsimonious and theoretically valid results.

Predictor Variables and Data Reduction Techniques

Next, we identified grantee and program characteristics that may predict program success and compiled the corresponding data. This dataset also included exogenous variables we deemed critical control variables, such as weather metrics, average energy price, median income, and other variables that may affect energy use, savings, and participation rates. Due to the large number of predictor variables of interest, we conducted a factor analysis on all continuous variables as a means of data reduction and to identify latent variables present in the dataset; the results of which informed the construction of index [predictor] variables that represented the factors identified in the factor analysis.

Regression Modeling

Next, we explored which programmatic elements were significant predictors to grantee success, when controlling for exogenous variables. Our analysis aimed to identify both the drivers and detractors of success among residential programs. Thus, we used two mutually exclusive sets of binary regression models to explore these relationships, where each set of regression models employed a different dependent variable (but tested for relationships with the same set of predictor variables): one set of models sought to identify which elements predicted membership in the least successful cluster and the other aimed to identify the elements predicting membership in the most successful cluster. Since standard maximum likelihood estimate-based logistic regression models perform poorly on small samples (Firth 1993), and the number of records in our dataset (n = 54) is considered a "small sample" (Long 1997), we used penalized maximum likelihood logistic regression, which corrects for small sample bias (Heinze & Ploner 2004).

First, we used bivariate logistic regression to explore whether any of the proposed predictor variables predicted membership in either the least successful cluster or the most successful cluster, respectively. Next, we ran multivariate regression models for each dependent variable using the independent variables identified as meaningful predictors ($p < 0.10$) in the aforementioned bivariate models.[2] We used a stepwise approach to add these variables into the respective models in order to derive optimal models (that is, until adding additional variables no longer improved the model). Finally, we used Tjur's R-Squared to measure the predictive power of each model (Tjur 2009). Tjur's R-Squared is more appropriate for penalized likelihood logistic regression, as it relies on the mean differences of the predicted probabilities of the model (as compared to other pseudo R-Squares, which rely on the maximum likelihoods that characterize standard logistic regression models).

Results

The Measure of Success

We used LPA as a means of measuring the multi-faceted concept of "success" among residential energy efficiency upgrade programs. First, we compiled grantee data on 12 diverse metrics of programmatic success to be used for the LPA. The performance metrics covered program/market saturation, cost effectiveness, program effectiveness, and wider economic impacts. The performance metric data are based on grantee level data aggregated from each grantee's project level data, and thus reflects how program-wide outcomes varied with respect to the 12 performance metrics. LPA is an exploratory technique, and our

[2] While we employed $p < 0.05$ as the threshold for statistical significance in this study, we retained predictors demonstrating $p < 0.10$ at the bivariate level for subsequent multivariate modeling in order to see if their relationship with the dependent variable strengthened when controlling for other factors.

analyses sought to identify groups, or clusters, of grantees that differed meaningfully in their performance on 12 metrics of program success. Results of the LPA revealed grantees clustered into three groups, and our analysis of each group's performance on the 12 performance metrics demonstrated that one group generally performed best on each of the metrics (the "most successful" cluster), another group generally performed worst on the metrics (the "least successful" cluster), and a third group demonstrated mid-range values on the performance metrics (the "average" cluster). Thus, the LPA revealed clusters of grantees that were more or less successful relative to one another. Figure 1 demonstrates these tiered levels of grantee success by exhibiting the cluster means for each performance metric included in the LPA model.[3]

While most indicator variables yielded mean cluster values that were consistent with a most, average, and least successful groupings interpretation, three indicator variables exhibited cluster means that deviated from this interpretation. Specifically, the most successful cluster had a somewhat higher program costs per job hour than the average group, the least successful cluster had a negligibly higher proportion of comprehensive projects that the average group (a difference of 0.22%), and the average cluster had a negligibly higher average MMBtu savings per project compared with the most successful group (a difference of 0.45).

Drivers of Success

We present two sets of results in this section: 1) significant predictors of membership in the least successful grantee cluster, and 2) significant predictors of membership in the most successful grantee cluster.

Predicting Membership in the Least Successful Cluster. First, we ran bivariate regression models to explore whether any of the proposed predictor variables predicted membership in the least successful cluster. Tested predictor variables included both programmatic elements (covering such areas as program design and financing) as well as exogenous controls (such as demographics and weather patterns of the grantee service area). Only two conceptual areas – contractor training and audit types – yielded any meaningful bivariate relationships. While several variables related to contractor training predicted membership in the least successful group at the bivariate level, we determined that "any contractor training offered" was the optimal variable to include in subsequent multivariate modeling attempts, since multicollinearity and assumptions of independence concerns prevented us from including multiple contractor training variables in the same multivariate model. Further, the lack of contractor training offerings was the strongest predictor of belonging to the least successful cluster (Tjur's R2 = 0.32). Thus, multivariate models predicting membership in the least successful cluster just included *any contractor training offered* and *number of audit types offered* as predictors.[4]

Multivariate results suggest offering any form of contractor training is the best way to mitigate lackluster program performance, regardless of other program elements or exogenous factors. As seen Table 1, grantees that offered contractor training were significantly less likely to be in the least successful cluster (Model 1). Further, while the number of audit types offered initially predicted membership in the least successful cluster (Model 2), this relationship is no longer statistically significant when any contractor

[3] While two of the metrics - *program's total contractor job hours invoiced* and *total program-wide present value of lifetime cost savings* – were not normalized to reflect the grantee's award amount (or more specifically, their residential outlays), correlations between either of these two metrics and residential outlays were both below 0.75, revealing that performance on these metrics were not direct functions of residential outlays. Further, when regressed on success cluster variables, these two metrics either did not predict (p > .5) or only marginally predicted (odds ratios = 1.00000) membership in any of the successful clusters.

[4] *Number of audit types offered* indicates the number of unique audit approaches available to participants: such as online, mail-in, phone-based, walk-through, or audits using diagnostic equipment.

training offered is also included in the model (Model 3).

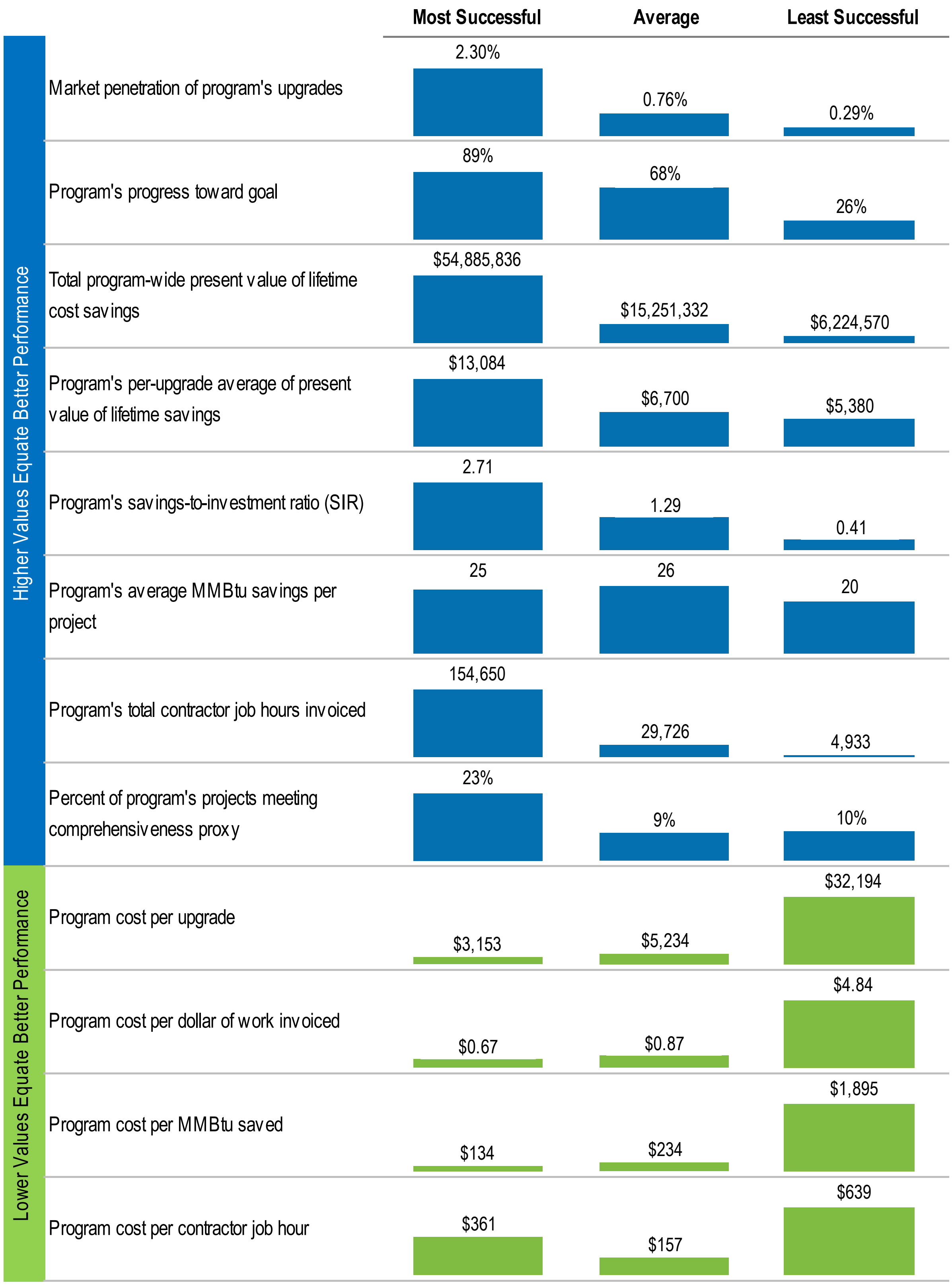

Figure 1. Performance metric Cluster Means (n=54)

These results demonstrate lack of contractor training is the strongest predictor of membership in the least successful cluster. Further, none of the exogenous control variables (such as energy prices or regional economic indicators) were associated with belonging to the least successful cluster or confounded contractor training's relationship with membership in the least successful cluster.

Table 1. Multivariate Logistic Regression Modeling of Least Successful Cluster Membership (n=54)

Variable	**Model 1**	**Model 2**	**Model 3**
Number of audit types offered	-	0.16*	0.56
Any contractor training offered	0.04***	-	0.07*
Wald test	9.56**	3.52*	9.04*
Tjur's R2	0.32	0.18	0.34

Note: Rows above the grey bar present odds ratios.

* $p < .05$; ** $p < .01$; *** $p < .001$

Predicting Membership in the Most Successful Cluster. Bivariate models predicting membership in the most successful cluster yielded several different programmatic elements and exogenous control variables as significant predictors. Specifically, bivariate regression models indicated grantees were more likely to be in the most successful group if they offered direct install options or did not require savings thresholds at the project level. Additionally, increased numbers of eligible upgrade contracting firms and audit type offerings were also associated with increased likelihood of being in the most successful group. The *timing of peak performance index* (and index representing how quickly a program was able to begin functioning at its best and then how long it was able to sustain its peak performance) as well as the program's *ramp-up time* (the length of time between the grant award date and the start of the aforementioned 'peak performance' period) also exhibited bivariate relationships with the most successful dependent variable. Staff experience also predicted success: grantees with at least one staff member with 15 years or more of relevant experience were significantly more likely to be in the most successful group. Further, three exogenous control variables – population of grantee's target area, the average cost of electricity in the grantee's state, and the constraints on energy use and savings opportunities index (which accounts for the weather patterns and age of housing stock in a grantee's target area) – also predicted membership in the most successful cluster. No other programmatic elements or exogenous control variables significantly predicted membership in the most successful group.

After identifying meaningful ($p < .1$) bivariate predictors of membership in the most successful group, we conducted multivariate logistic regression analysis using those predictors. We employed a stepwise approach to multivariate regression modeling. It quickly became apparent that three particular variables collectively predicted most successful cluster membership, at which point our iterations simply entailed adding and subsequently removing the rest of the independent variables one at a time (as no other variables ultimately retained significance nor explained away any of the three significant predictors when added to the model). As a result, our multivariate regression tables reported in this volume start with the "final model" (Model 1), and then subsequently demonstrate how the other previously meaningful ($p < .1$) predictors are no longer meaningful or significant once they are included in a multivariate model with the three primary predictors of most successful cluster membership (Table 2 and Table 3). Since all of these models do not fit on one page, we divided the models into two tables: one exhibiting the effect of adding exogenous controls (Table 2) and one demonstrating the effect of adding additional programmatic elements to the model (Table 3). The interpretation of the multivariate regression results are as follows.

Multivariate modeling reveals *offering multiple pathways to participation and achievement of energy savings* is critical to achieving the most successful program outcomes, regardless of other program elements or exogenous factors. Specifically, programs that include direct install options, offer multiple audit types, and allow larger numbers of contracting firms to perform upgrades are more likely to be in the most successful cluster. Further, these elements are predictors of being in the most successful group net of exogenous control variables, suggesting that offering multiple pathways to participation and achievement of energy savings ensure program success, regardless of the population size, energy costs in a program's service region, or housing stock- and weather-oriented constraints on energy use and savings opportunities (Table 2).[5]

Table 2. Multivariate Logistic Regression Modeling of Most Successful Cluster Membership, Testing Additions of Exogenous Controls (n=54)

Variable	**Model**						
	1	**2**	**3**	**4**	**5**	**6**	**7**
Constraints on energy use and savings opportunities index	---	1.43*	1.06	---	---	---	---
State-level average electricity cost (cents per kWh)	---	---	---	1.26*	1.3	---	---
Population of grantee's service area	---	---	---	---	---	1.00*	1.00
Direct install options offered	24.82***	---	21.12***	---	25.43***	---	24.72***
Number of audit types offered	3.89*	---	3.68*	---	4.75*	---	3.92*
Number of eligible upgrade contractor firms	1.02**	---	1.02**	---	1.02†	---	1.02*
Wald test	11.81**	3.74†	12.04*	4.157*	11.54*	3.58†	11.94*
Tjur's R2	0.55	0.11	0.56	0.10	0.61	0.10	0.58

Note: Rows above the grey bar present odds ratios. † $p < .1$; * $p < .05$; ** $p < .01$; *** $p < .001$

Further, multivariate modeling demonstrates that program elements associated with providing multiple pathways to participation and *achievement of energy savings* are the key programmatic predictors of belonging to the most successful cluster, as other programmatic elements (specifically: savings threshold requirements, ramp-up time, the peak performance index, and staff experience) were no longer significant when included in multivariate models alongside the *multiple pathways* indicators (Table 3). Additionally, we explored interaction effects with the three independent variables in Model 1, to assess if the effect of *direct install options offered* was modified by either *number of audit types offered* or *number of eligible upgrade contractor firms*. The resulting analysis demonstrated that the independent variables in Model 1 did not interact with each other and thus did not modify the effect of *direct install options offered*.

[5] Multivariate modeling demonstrates the number of eligible upgrade contractors as a predictor of most success is not a function of population size; thus, population size is not confounding this relationship.

Table 3. Multivariate Logistic Regression Modeling of Most Successful Cluster Membership, Testing Additions of Programmatic Elements (n=54)

Variable	**Model**								
	1	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
Savings threshold required for qualified projects	---	.13**	0.26	---	---	---	---	---	---
Ramp up time	---	---	---	0.22**	0.39	---	---	---	---
Timeliness index	---	---	---	---	---	1.58*	1.47	---	---
At least one team member had 15 years or more of relevant previous experience	---	---	---	---	---	---	---	4.61*	1.82
Direct install options offered	24.82***	---	17.80**	---	22.32***	---	27.67***	---	18.14***
Number of audit types offered	3.89*	---	4.37*	---	3.86*	---	4.12*	---	3.77*
Number of eligible upgrade contractor firms	1.02**	---	1.02**	---	1.02*	---	1.02*	---	1.02**
Wald test	11.81**	8.17**	11.45*	5.95*	12.09*	4.91*	11.65*	3.82†	12.40*
Tjur's R2	0.55	0.19	0.58	0.18	0.59	0.14	0.60	0.09	0.56

Note: Rows above the grey bar present odds ratios. † $p < .1$; * $p < .05$; ** $p < .01$; *** $p < .001$

Discussion

Our analyses suggest residential energy efficiency programs can mitigate poor performance outcomes by providing contractor training opportunities and, further, can achieve successful program outcomes by offering homeowners multiple pathways through which they can engage with the program and achieve energy savings. This section explores how these programmatic elements may contribute to program success.

Avoiding Poor Performance via Contractor Training

In order to deliver residential energy efficiency services effectively, programs rely on contractors with the skills needed to sell and perform the audit and upgrade work (State and Local Energy Efficiency Action Network Residential Retrofit Working Group 2011). However, some regions may lack a sufficiently large base of qualified contractors with experience and expertise in energy efficiency building science. Further, grantee experience suggest that even when a strong contractor base exists in a region, participating contractors can benefit from sales training, technical training, and training on program processes and requirements. Several studies support these findings. For example, one study found that contractors believed BPI certification was often a strong selling point when attempting to attract customers (GDS Associates, Inc. 2009), and a 2011 report from SEE Action found it is imperative to offer sales training to contractors because of the important role they play in outreach. Given these benefits, programs have frequently found contractors to be extremely interested in program-related training (Energy Market Innovations, Inc. 2012; NMR Group, Inc. 2012).

Thus, contractor training is a critical step to successfully delivering program services. Without contractor training, residential energy efficiency programs may suffer from lackluster results; the least successful grantees – who were significantly less likely to offer training than average and most successful grantees – achieved comparably lower market penetration, energy savings, and progress toward upgrade count goals than more successful grantees. Further, the least successful grantees had higher program costs per upgrade than average and most successful grantees; combining these findings on cost and training, we speculate these higher costs may owe in part to a lower quality contractor base that ineffectively or inefficiently delivered audit and upgrade services.

Crafting Highly Successful Programs by Offering Multiple Pathways to Participation and Achievement of Energy Savings

The results of the regression analysis demonstrate the importance of offering multiple pathways to participation and achievement of energy savings in order to achieve the most successful program outcomes. Allowing participants to enter the program and achieve energy savings in a variety of ways makes participation easier for customers and takes advantage of the strengths of various program design structures, while mitigating their limitations. Specifically, our regression analyses suggest providing multiple audit types, direct install options, and larger numbers of contracting firms that can perform upgrades are key components of the most successful residential upgrade programs. These three elements constitute multiple pathways to participation and achievement of energy savings; the following sections further explore the benefits associated with these specific predictors of successful program outcomes.

Offering Multiple Audit Types. Offering multiple audit types, such as online, mail-in, phone-based, walk-through, or in-depth audits (which we subsequently refer to as "diagnostic audits" due to their typical use of diagnostic equipment), provides potential participants with a variety of ways to begin engaging with a program and identifying ways that they can save energy.[6] Since certain audit types may be more appropriate for or appealing to different homeowners, offering multiple audit types successfully accommodates potential participants' varying wants and needs.

Further, prior research has shown that "there is no…'correct' model of retrofit decision-making… Nor is such a model likely to emerge in the future. Like most other types of behavior, energy related decision-making is multi-faceted" (Sanstad et. al 2010). Thus, offering prospective participants multiple audit types increases the types of customers appealed to, thus potentially increasing program-wide participation and conversion rates.[7]

While diagnostic audits are commonly viewed as the gold standard, studies have found similar conversion rates for diagnostic and other types of audits (Scott et. al 2014; ECONorthwest 2010). Further, there are many benefits associated with offering less intensive audits in tandem with diagnostic audits. Walk-

[6] There is no industry standard terminology for what we term in this paper "diagnostic audits." By diagnostic audit, we mean the most comprehensive of audit types, which typically use diagnostic equipment (such as blower door equipment and infrared cameras) to improve the identification and quantification of energy savings opportunities. Consistent with the lack of industry standard terminology, this audit approach itself is unstandardized. Software and diagnostic tools used in these audits vary from program to program and even project to project within a given program.

[7] We interpret the finding that offering multiple audit types is associated with grantee success as suggesting its value lies in increasing the types and therefore number of customers that pursue an audit. Our data provide no insight into possible variations among audit types in rates of conversion to upgrades or in resulting upgrade savings.

through audits conducted by experienced contractors can identify considerable savings recommendations and identify "hot leads" in a more cost-effective manner than diagnostic audits, for example. Additionally, multiple grantees reported diagnostic audits constituted barriers to participation due to lack of familiarity with the concept of diagnostic audits, homeowner inability to stay home during the duration of the audit, cost to the participant (despite audit incentives), and skepticism regarding the value of audits relative to the cost. These barriers are less pertinent to less comprehensive audits, such as online or walk-through audits. Skeptical or frugal homeowners may be more likely to pursue a lower cost audit option, which may be perceived as less of a financial risk. Further, costly diagnostic audits may constitute an equity issue; offering lower cost audit options expands the pool of homeowners that can afford to participate, increasing the amount of program-wide savings as it facilitates the participation of the traditionally hard-to-reach low- and middle-income populations.[8]

Providing Direct Installation of Low-Cost Measures. Regression analyses revealed conducting direct install of low-cost measures was the strongest predictor of membership in the most successful cluster. Grantees reported direct install options, which were often included in the audit prior to a more comprehensive upgrade project, could serve as both sources of significant energy savings as well as "sweeteners" to encourage participation in the audit or a subsequent upgrade project.[9] Directly installing low-cost measures (such as LEDs, showerheads, and faucet aerators) during an audit allows programs to claim direct energy savings prior to a comprehensive upgrade project (as well as garnering savings from audit participants that do not pursue an upgrade project), which can increase program cost effectiveness. Direct installations also serve a quality control function, as trained building science experts, rather than homeowners, install measures and ensure that they are installed correctly. Additionally, research has found direct install activities have high customer satisfaction, may motivate customers to participate in a program who may not have participated otherwise, were associated with efficient lighting remaining in sockets longer, and were more likely than other delivery methods to result in the installation of lighting measures (Peters et. al 2010).

Having a Large Number of Contractors Eligible to Conduct Upgrades. Having a large number of firms that are eligible to complete program upgrade projects makes it easier for participants to find a qualified contractor; in addition, some participants may appreciate the ability to shop for contractors in order to find the best quote. A recent baseline study of a whole house retrofit program in California found most homeowners who had recently completed renovation projects costing at least $3,000 (including participants and nonparticipants from various energy efficiency programs) chose contractors that they had previously worked with, had prior relationships with, or found via word of mouth (DNV GL 2014). Additionally, the study found that only about half of those completing renovations contacted more than one contractor. Since homeowners primarily rely on existing relationships and referrals when selecting upgrade contractors, having a large number of contractors eligible to conduct upgrades increases the probability that a homeowner's preferred contractor is performing upgrades through the program, which, in turn, increases the likelihood that a homeowner will complete an upgrade project through their local energy efficiency program. Moreover, having a large number of eligible contracting firms maximizes the number of projects that can be conducted at a given time. Increased numbers of eligible contractors can also magnify program and energy efficiency

[8] Offering free diagnostic audits remedies any cost-to-participant-related barriers. However, the provision of free audits result in higher program costs. Some grantees concluded free audits lowered conversion rates; willingness to pay an audit was associated with greater likelihood of pursuing the upgrade.

[9] As stated, the success strategy of direct measure installation was typically coupled with onsite audits. However, grantee experience suggests the strategy is not limited to onsites; a few grantees directly installed measures at times other than onsite audits.

upgrade awareness via contractor-led advertising and outreach efforts.[10] These findings suggest increases in the number of eligible upgrade contracting firms can result in more program-wide energy savings.

Programs seeking to maximize the number of eligible upgrade contractors have a variety of avenues for doing so. Since contractors may be deterred from participating in programs that are overly complex and burdensome, easing the contractor experience (such as simplifying and minimizing participation steps and paperwork) may help increase the number of contractors seeking program eligibility. Further, program-to-contractor outreach can raise awareness of the program among nonparticipating contractors, which could in turn increase contractor participation. However, we caution program administrators against relying on overly lax contractor eligibility criteria, which may maximize the number of eligible contractors yet reduce the project quality on average. While proper quality assurance and quality control (QA/QC) techniques could minimize quality of work issues, our regression analysis demonstrates programs should provide training opportunities for their contractors in order to avoid sub-par program outcomes. Thus, having a large pool of trained eligible upgrade contractors is key to the most successful programs.

Conclusions

This study defined and quantified a multi-faceted measurement of relative program success among BBNP grantees and statistically identified factors associated with achieving success. Using LPA as an exploratory technique, our analyses sought to identify groups, or clusters, of grantees that differed meaningfully in their performance on 12 metrics of program success. Results of the LPA revealed grantees clustered into three groups, and our analysis of each group's performance on the 12 performance metrics demonstrated that one group generally performed best on each of the metrics, another group generally performed worst on the metrics, and a third group demonstrated middling values on the performance metrics. Regression analyses demonstrated four programmatic elements predict cluster membership.

Specifically, our regression analyses revealed that not providing contractor training was the strongest predictor of membership in the least successful cluster, and program designs that allowed for multiple pathways to participation and achievement of energy savings predicted membership in the most successful cluster. Regression results identified the following as critical components of multiple pathways to participation and achievement of energy savings: offering direct install options and multiple audit types, and having a large number of eligible contractors than can perform upgrades. Since this study analyzed program data from 54 diverse grantees and subgrantees spanning widely varied regions and demonstrated that exogenous elements neither explained nor confounded variation in success, the statistical findings are particularly insightful for the energy efficiency industry as they elucidate what can make or break a residential program regardless of broader contextual factors.

Further demonstrating the value of this volume's findings, the regression results are intuitive, reinforced by other qualitative and quantitative grantee findings, and supported by the literature. Our rigorous analysis of a wealth of information confirms many program design and implementation approaches identified as effective by the industry literature and program administrator experiences. The findings provide the energy efficiency community with greater confidence in its understanding of how to make residential upgrade programs successful.

[10] Residential nonparticipant awareness of their local BBNP program was significantly higher among the most successful grantees (37%) than among average (32%) and least successful grantees (21%). About one-quarter of surveyed residential participants reported learning about the program from their contractors (compared to 66% for publicity sources such as advertising and 37% for program sources).

References

DNV GL. 2014, July. *Baseline Characterization Market Effects Study of Investor-Owned Utility Whole House Retrofit Programs in California (Final Report).* CPU0073.01. California Public Utilities Commission.

ECONorthwest. 2010, April. *Process Evaluation of the PG&E 2006-2008 Home Energy Efficiency Survey (HEES) Program.* http://www.calmac.org/publications/PG%26E_HEES_Report.pdf

Energy Market Innovations, Inc. 2012, September. *California HVAC Contractor & Technician Behavior Study*.

Firth, D. 1993. "Bias reduction of maximum likelihood estimates." *Biometrika.* 80:27-38.

GDS Associates, Inc. 2009, February. *New York Home Performance with ENERGY STAR® Program Market Characterization and Market Assessment Evaluation.*

Heinze, G., and Ploner, M. 2004. *A SAS-macro, S-PLUS library and R package to perform logistic regression without convergence problems (Technical Report)*. Section of Clinical Biometrics, Department of Medical Computer Sciences, Medical University of Vienna, Vienna, Austria.

Long, J. S. 1997. *Regression Models for categorical and limited dependent variables*. Thousand Oaks, CA: Sage Publications.

NMR Group, Inc. 2012, September. *Process Evaluation and Market Characterization and Assessment: Green Jobs - Green New York Residential Program.*

Peters, J., Moran, D., and Frank, M. 2010. *Lighting Program Assessment: Residential Direct Distribution.* Portland, OR: Bonneville Power Administration

Sanstad, A., Diamond, R., Sanquist, T., and Lutzenhiser, L. 2010. *U.S. Single-Family Homeowners' Decision-Making Regarding Energy Retrofits: A Literature Review*. U.S. Department of Energy.

Scott, K., Kociolek, E., and Castor, S. 2014. Customer Engagement Experiment: Which Follow-Up Strategies Turn Home Energy Audits Into Energy Savings? *ACEEE paper;* Energy Trust of Oregon.

State and Local Energy Efficiency Action Network Residential Retrofit Working Group. 2011, June. *Roadmap for the Home Energy Upgrade Market*.

Tjur, T. 2009. "Coefficients of determination in logistic regression models—A new proposal: The coefficient of discrimination." *The American Statistician* (63), pp. 366-372.