

Learning from Successive Evaluations of a Low-Income Program

*Lori Megdal, Ph.D., Megdal & Associates, Acton, Massachusetts
Margaret Cush Grasso, and Elaine Akley, KeySpan Energy Corporation, Islandia, New York
Bruce Bennett, GDS Associates, Inc., Manchester, New Hampshire¹*

ABSTRACT

A multi-year process evaluation was conducted on the Long Island Power Authority's (LIPA's) Residential Energy Affordability Partnership Program (REAP). This is part of LIPA's Clean Energy Initiatives and is designed to assist low-income customers with greater affordability of their energy bills. The multi-year process evaluation allowed the evaluation team to examine whether there were trends to indicate whether changes in the program had been effective or not. The analysis looked at if and how the demographics had changed in the program. The various analyses were then examined to ensure that any trends being seen were, in fact, actually true for the program rather than an artifact of the changing demographics. Satisfaction ratings were obtained in the survey by asking for scores of one to five for very dissatisfied to very satisfied. This is a common inquiry in process evaluations. This data is ordinal ranking data rather than interval data. Often, however, means are analyzed and t-statistics provided for comparisons. These statistics are not valid for ordinal ranking data. Valid comparisons on satisfaction ratings were made in this evaluation by using a Wilcoxon Rank Sum and Kruskal-Wallis Chi-square (χ^2) test to analyze for statistical significance. This paper presents this logic and the procedure that can be simply used to obtain valid statistical significance testing for comparing satisfaction rating data. The paper presents the methodology, results, and provides an example of the usefulness of multi-year process evaluations and the use of the Kruskal-Wallis Chi-square (χ^2) test.

Introduction and Program Background

The Long Island Power Authority's (LIPA's) Residential Energy Affordability Partnership Program (REAP) is part of LIPA's Clean Energy Initiatives and is designed to assist low-income customers with affordability of their energy bills.

The REAP Program provides:

- Installation of comprehensive electric energy efficiency measures and
- Extensive energy education and counseling.

KeySpan Energy, under contract with LIPA, is managing the program. Presently, it is being implemented for KeySpan by Honeywell Utility Solutions. A partnership is also in place with the Community Development Corporation of Long Island (CDC) for their participation in program implementation. Other players in the process include Vermont Energy Investment Corporation (VEIC, consultant), Optimal Energy (consultant) and Appliance MD (refrigerator replacement).

The program targets two low-income customer segments. The first segment includes low-income customers who qualify for the US Department of Energy's Low-Income Weatherization Assistance Program (WAP). The second segment includes households who do not qualify for WAP and

¹ Sharyn Barata of B&B Resources, Inc. was a significant contributor for other elements of the Process Evaluation not reported upon in this paper. She conducted and analyzed the many in-depth interviews. Opinion Dynamics Corporation also was part of the overall project team and conducted the telephone surveys. CMC Energy Services conducted the site visits led by C. Joseph Iandolo. The project also received input and assistance from Ken Tohinaka of VEIC and Ralph Prah of Prah Associates as consultants to KeySpan for the LIPA programs.

whose income level is no more than 60% of the median household income level for the LIPA service area (a population with higher incomes than those in the first segment). The focus in the second segment is to attract customers with primary electric space heating and central air conditioning equipment.

The REAP Consumer Education and Counselling component works with participants to identify actions they can take to lower their electric bills. Participants are required to accept certain responsibilities in order to participate in this program. Through these commitments, the benefits of any installed energy efficiency measure will be maximized. The education component includes information related to:

- Set back thermostat operation and management,
- Inspection, maintenance and replacement of central air conditioning equipment filters,
- Use and value of installed compact fluorescent lighting retrofits,
- Electrically heated hot water conservation measures,
- Electric hot water heater temperature settings, and
- Refrigerator control settings.

All measures are installed at no cost to the participant. Only measures that meet cost effectiveness tests (produce positive utility net benefits) are installed. The number of measures installed for any participant is dependent upon the evaluation of their usage and environment. So the measures installed vary significantly from one participant to another. Potential measures include: High-efficiency compact fluorescent screw-in (CFL) lighting products; Replacement of very inefficient refrigerators with new ENERGY STAR rated units; Water conservation and electric energy efficiency measures for those with electric water heaters (e.g. aerators, showerheads and electric hot water heater insulating jackets); Combustion systems safety checks for those without electric heat; Wall and/or attic insulation; Air leakage sealing; Electric heating system maintenance, modification or repair; Air conditioning system maintenance and/or modification; Ducted distribution system modifications, and Other custom measures that are determined to be cost effective.

The LIPA REAP program has been fully implemented since February 2000. The REAP program served 4,250 low-income customers in 2001 and another 4,206 in 2002. Program installations through 2002 included almost 43,000 CFLs, over 2,600 new ENERGY STAR replacement refrigerators, and many other measures.

The Megdal & Associates team was contracted to conduct impact and process evaluations from 2001 through 2003, with the final multi-year process evaluation completed in March 2004. The evaluation reports produced included: *2001 Preliminary Impact and Process Evaluation* (April 2002), and the *2002/2003 Impact Evaluation Report* (June 2003) and the *2003 Process Evaluation of LIPA's Clean Energy Initiative REAP Program* (March 2004). The latter is a process evaluation that compared findings across multiple periods of time and is the basis for the material provided in this paper.

Methodology

The 2003 process evaluation included staff and stakeholder interviews and a participant telephone survey. The evaluation project was led by Dr. Lori Megdal of Megdal & Associates. The project manager, Margaret Cush Grasso, from KeySpan Energy was directly involved and played a key role in the evaluation. The stakeholder interviews were conducted by Sharyn Barata of B&B Resource, Inc. The telephone survey was conducted by Opinion Dynamics Corporation. The analysis of the telephone survey was conducted by Bruce Bennett of GDS Associates, Inc. under the direction of Dr. Megdal.

The design, analysis, and results of the telephone survey are the focus of this paper. The

telephone survey analysis was designed to be a multi-year analysis to allow comparisons between the findings for participants at different times during the program refinement process. The surveys made inquiries on process and impact related issues. The evaluation focused on several subgroups of interest. Given this, these subgroups were selected as different strata for a random sample. In this way, we could ensure that questions referring only to a stratum still could provide meaningful results.

The evaluation team stratified the samples along many dimensions. The primary dimension was timing. There was a preliminary process evaluation that included program participants through September 2001. These have the designation of Time Period 1 (as the earlier process evaluation results are compared to those found in the later survey). The second group (Time Period 2) are those that participated in the program after the preliminary process survey was conducted but before program changes occurred as part of the annual program improvement process and made in response to the Preliminary Process and Impact Evaluation. These improvements were in place in June 2002. The Time Period 2 participants are those who participated from October 2001 through May 2002. The third group is those that participated after program improvements, June 2002 through December 2002 (Time Period 3). Time Period 1 participants were interviewed in 2001 and Time Period 2 and 3 participants were interviewed in 2003.

A significant program improvement since the Preliminary evaluation was providing more thorough duct repair services and in targeting participants for these. Given this, there was an interest to examine participant responses from those receiving duct repairs. Two of the strata examined were duct repair participants and participants with central air-conditioning (those that could possibly have been offered the duct repair service).

To ensure that overall results reflect the total program population, weighting was derived based on population participation and applied to the survey results.

Demographics and Measure Penetration by Time Period

All comparisons over time need to be concerned with whether the populations being examined are the same. This could be an issue given that these evaluations were initiated early on in the program, when recruitment efforts may have undergone some experimentation. A more solid foundation for interpreting these results is provided first by examining similarities and differences between the groups' demographics and using this information for further testing when differences are seen.

The telephone surveys were conducted randomly over the various demographic characteristics. As such, the differences seen in the sample demographics over time reflect the differences in the population demographics for participants over time. The staff, program, and stakeholder interviews indicated that program recruiting and targeting changed as the program matured and program changes were made. This is consistent with the changing demographics seen in the participant surveys.

The greatest differences seen between the time periods are in the percentage living in owner-occupied dwellings. These differences are due to some early recruitment efforts involving a few very large apartment complexes. As shown in Table 1, more than twice as many of this survey's respondents are in owner-occupied dwellings than those responding to the first survey in 2001 (Time Period 1), 39% to approximately 75% (average of Time Period 2 and 3). Correlated with that is the fact that almost twice as many of the survey respondents have someone over 65 in the household and a much smaller percentage have children in the home.

Table 1. Demographics for the Participant Samples

	Time Period 1: Through Sept. '01 (n=313)	Time Period 2: Oct. '01-May '02 (n=179)	Time Period 3: June '02-Dec. '02 (n=221)
Percent of Respondents			
Owner occupied	39*	77	73*
Single family	61	77	68
With someone 65 or older in HH	36*	62	57*
With children	51*	28	31*
With annual household income \$20,000 or less	42*	44	33*

* Z-score for a difference of proportions shows greater than a 95% statistical significance.

The most recent participants also have a lower percentage in the lowest income group (household income of \$20,000 or less), though these still represent almost one-third of the participants.

Though the participant demographics have changed over these time periods, the mix and number of measures provided has generally not changed. The evaluators' experiences have found that there are often large differences between measures installed in owner-occupied residences versus renter-occupied ones across residential programs (low-income and non-low-income). This low-income program is free to the participants without greater financial obligations from landlords and, therefore, this program attribute could contribute to the lower differentiation seen here than with other programs we have reviewed. We have seen programs where a difference between owned and rented property occurs due to difficulties in obtaining acceptance of the measures by both tenants and landlords. This did not occur in the LIPA program.

The measure penetration by time period was examined for 11 measure categories. There were two measures that showed penetration changes: CFLs and duct repair measures. The data shows that 96% of the participants in Time Period 1 and Time Period 2 received CFLs and 87% of the Time Period 3 participants received CFLs. The average number of CFLs has declined somewhat, from 5.4 in Time Period 1 to 4.6 in Time Period 3. At the same time, it is likely this decline coincides with fewer people in the home (given a lower percentage with children) and, therefore, fewer cost-effective opportunities. It is also possible that there were fewer opportunities depending upon whether more participation came from areas that already had greater prior program penetrations, such as the Rockaways (one specific geographic area in the LIPA territory that had significant efforts from earlier programs).

The percent receiving air sealing and duct repair has been increasing. This could be due to the increased HDMC staff training and focus upon duct repair, and their ability to obtain savings from homes that are eligible for these measures. An increasing proportion of participants have central air-conditioning (CAC), from 11% in Time Period 1 to 17% in Time Period 2 and 20% in Time Period 3. This could have offered greater opportunities for duct repair measures. At the same time, however, training and emphasis on duct repair might have affected recruiting, which in turn attracted more CAC owners. (The HDMC staff were trained to test and conduct duct repairs and the HDMC staff led the participant recruitment.)

Examining Refrigerator Replacement and Education Over Time

In the spring of 2002 one brand of refrigerator used by the Program had a greater occurrence of manufacturer's defects than others. At the same time, having a new local refrigerator contractor helped

the program more quickly shift to another brand while the problems were being resolved, replacement refrigerators to be provided in a timely manner and allow other refrigerator repair issues to be handled more quickly. The effect of these counter-balancing issues was investigated. Customers were asked to express their satisfaction by responding whether or not they were completely satisfied with their new refrigerator. Although a majority of the respondents were completely satisfied with the new refrigerators across all Time Periods, 10% of those that received refrigerators during Time period 2 (Oct. 2001-May 2002) were not completely satisfied with the new refrigerator. Yet, when asked to explain the reasons for their dissatisfaction the most common responses did not relate to the problems the program had with refrigerators during Time Period 2.

All survey participants were asked if they perform any maintenance to their refrigerator in order to maximize savings. Only 14% said they conducted refrigerator maintenance in the 2001 survey. As a result, refrigerator maintenance became one of the areas the program sought to improve. The 2003 survey, as seen in Table 2, shows continued improvement in the area of refrigerator maintenance. Twenty four percent of the respondents from Time Period 2 report performing refrigerator maintenance. Over one-third, almost 36%, of the Time Period 3 respondent said they conduct refrigerator maintenance. Consecutively fewer respondents are responding with neither, “no” (they do not do any refrigerator maintenance), nor “don’t know”. More people are certain they are conducting more refrigerator maintenance to save energy.

Table 2. Refrigerator Maintenance

	Do you conduct any refrigerator maintenance to maximize energy savings? (Response percentage)		
	Time Period 1	Time Period 2	Time Period 3
Responses (n)	313	179	221
Yes	14*	24*	36*
No	76**	71	62**
Don't Know	9*	5*	2*

* Z-score for a difference of proportions shows greater than a 95% statistical significance.

** 90% statistical significance between Time Period 1 and Time Period 3.

Increasingly, respondents claim LIPA field representatives explained the importance of performing specific refrigerator maintenance tasks on a regular basis. In addition, when asked what maintenance efforts were suggested by the field representatives, respondents recalled specific maintenance efforts that coincide with the maintenance tasks currently being performed by participants. The only deviation was that slightly fewer people recalled being told to clean their refrigerator coils than are actually conducting this task. There are fewer percentages of respondents “maintaining a tight seal” to save energy than those suggested to do so. In general, although there are fewer people conducting refrigerator maintenance than being suggested to do so, the trend from Time Period 1 through Time Period 3 is steadily increasing. This is supportive of the program improvements made: greater auditor training for greater customer education in this area.

Perception of Bill Reduction Due to Program Participation

In the telephone survey participants were asked if they experienced a reduction in their electric energy bills as a result of program participation. Half of the Time Period 2 participants felt the Program reduced their electric bills. Less than half (40%) responded positively in Time Period 3, though the

shorter post-installation period could mean their opinion might change when they have seen more electricity bills post-retrofit.

As discussed earlier, program demographics have varied over time. The perceptions of the program’s effect on electric bills was examined across the two time periods for the elderly versus non-elderly and then for owners versus renters. Over both time periods the elderly were less likely to report reduction in their electric bills. No consistent difference between owners and renters was seen across both time periods.

Recall and Effectiveness of the “Action Plan” Over Time

The LIPA REAP program assists participants with the development of an “Action Plan” where specific energy efficiency actions were identified to be taken by the customer. These actions are included with the LIPA REAP Partnership Agreement customers sign as part of their participation in the program. The lack of recollection of the “Action Plan” was one area of concern found in the Preliminary Process Evaluation after 2001.

Changes were made to the “Action Plan” and its presentation. Program changes were made to reduce the amount of program paperwork received by the participant and the “Action Plan” was redesigned to help increase its recollection. This was done through reducing paperwork to only one program document that was smaller and included the “Action Plan” with signature.

This issue was investigated again in the second process evaluation. Most respondents could not recall being assisted by the LIPA field representative with the development of an “Action Plan”. No increasing trend in recall was seen. There was some concern whether improvements in this area could be masked by the fact that the percent of elderly participants has grown over these time periods. So the responses were then examined for households with elderly versus those with no elderly residents. Table 3 shows that the percentages of participants that recall completing an “Action Plan” have remained between 35 and 50% for both households with elderly and without elderly residents. There is no trend within elderly households or within non-elderly households on recall rate, confirming no trend in recall rate and that such a trend is not masked by the increasing proportion of participants that are elderly.

The survey results suggest the improvements that were made have not shown to be effective at this time.

Table 3. Action Plan Development Assistance from LIPA

	Did the LIPA field representative work with you to develop an electric energy saving "Action Plan"? (Response percentage)					
	Time Period 1		Time Period 2		Time Period 3	
Responses (n)	114	199	111	67	126	91
	Eld	Non Eld	Eld	Non Eld	Eld	Non Eld
Yes	37.2	38.9	30.1	49.7	34.6	41.2
No	43.5	41.5	43.2	35	46	40.6
Don’t Recall/Don’t Know	19.3	19.5	26.7	15.4	19.5	18.2

Used to examine whether there was a trend in recall within the elderly or within the non-elderly over time.

Of those that called preparing an “Action Plan”, the most common action items taken by respondents were to “turn off the lights when not in the room”, “clean refrigerator coils periodically”,

“check the temperature setting on the refrigerator”, and “to turn down the heat at night or when not at home during the winter months”.

A majority of the respondents said the aforementioned actions helped to reduce electric energy bills, as seen in Table 4. It is likely that Time Period 3 participants did not have enough time to tell whether the actions saved them energy. Given this, Time Period 3’s lower percentage positive response and greater “Don’t know” responses may be a function of this lack of time post-implementation rather than a final measurement. Future evaluations could test this hypothesis.

Table 4. Action Plan’s Assistance in Reducing Electricity Costs

	Did actions help to reduce electric energy bills? (Response percentages)		
	Time Period 1	Time Period 2	Time Period 3
Responses (n)	120	59	69
Yes	77	79*	47*
No	13*	9*	23*
Don’t Recall/Don’t Know	10	11*	31*

* Z-score for a difference of proportions shows greater than a 95% statistical significance.

Participants were also asked to rate the helpfulness of the LIPA REAP program with the implementation of their “Action Plan” after the initial visit. As shown in Table 5, most of the respondents felt that the Program was at least of some help. However, a greater percentage of respondents in the 2001 survey (63.8%) felt that the Program was “Extremely” helpful. Whereas, a smaller percentage of respondents to the 2003 survey (44.8% and 39.5% from Time periods 2 & 3, respectively) felt that it was “Extremely” helpful. A larger percent of respondents from Period 3 (11.5%) felt that the Program was of no help. At the same time, a shorter time between retrofit and the survey could be influencing this number. So there needs to be some qualification and re-testing in future evaluations to the trend implied in Table 5. At the same time, the program should examine whether there are program changes that may have caused these responses and whether program improvements are warranted. The future analysis should also be examined for statistical significance through a Wilcoxon Rank Sum and Kruskal-Wallis Chi-square (χ^2) test as discussed below (as was used with the satisfaction rating comparisons done in this study).

Table 5. Helpfulness of the Program with Implementation of Action Plan

	How helpful was the Program with Implementation of your Action Plan after the initial field visit? (Response percentage)		
	Time Period 1	Time Period 2	Time Period 3
Responses (n)	120	66	83
Extremely Helpful	64	45	40
Of some help	27	50	40
No help	4	5	12
Don’t Recall/Don’t Know	5	1	10

Satisfaction Over Time and Use of the Wilcoxon Rank Sum and Kruskal-Wallis Chi-square (x2) Test

An important reason for the telephone survey in the process evaluation was to obtain from participants how satisfied they were with various components of the program including their satisfaction with the work performed. As was the case with the previous survey, participants were asked to rate their satisfaction on a scale of 1 to 5 with 5 being “extremely satisfied” and 1 being “extremely dissatisfied”.

Overall, 85% and 81% of the customers from Time periods 2 and 3 respectively, were satisfied with the program (responding with a score of 4 or 5). Compared to the 2001 survey when 75% of the respondents were satisfied, overall program satisfaction has improved. This is a positive finding which supports the views of the staff members interviewed who said the LIPA REAP program has been improving.

Table 6 presents the percent of satisfied (those “extremely satisfied” and those “satisfied”) by program component.

Table 6. Satisfaction by Program Component

Program Component Percent Satisfied (rating a ‘4’ or ‘5’) with ...	Time Period 1	Time Period 2	Time Period 3
Number of Responses (n)	313	179	221
Process for Scheduling Audit	68%	79%	78%
Quality of Work	73%**	81%**	84%**
Attitude of Initial Contractor	n/a	62%	55%
Attitude of Follow-up Installation Contractor	61%	84%	87%
Amount of Work Performed	70%	76%	73%
Electric Energy Savings	64%*	60%*	49%*
Explanation of How Appliances Use Energy	63%	66%	67%
Actions Recommended to Saving Energy	71%	67%	72%
Computer Printout Showing Electric Energy Use	51%	41%	45%
% Happy with New Refrigerator	n/a	90%	91%
Program Overall	75%	85%	81%

* Statistically significant at probability greater than Chi-square at 0.002 (greater than 99%) via a Kruskal-Wallis test.

** Statistically significant at probability greater than Chi-square at 0.21 (close to 80%) via a Kruskal-Wallis test.

There are two trends shown as statistically significant in Table 6, increasing satisfaction with the quality of work provided, and a decreasing satisfaction with the electric energy savings. Further exams were made to ascertain whether these trends were because of changes in the demographics of the participants over time or whether they are true program changes.

Often process evaluations report satisfaction inquiries in terms of frequency satisfied (as presented in the text above), mean score or both. Little attention has been given to whether one way or the other is more accurate. The detailed exam here has found that the trends discussed show up in the frequency satisfied but not in the mean scores. This issue was then further examined.

The satisfaction inquiry asked participants to provide a score from one to five on their level of satisfaction. This is actually an ordinal scale ranking more than a specific measurement of a degree of satisfaction (i.e., you can’t say that four is twice as much satisfaction as two or that five is twenty-five percent greater than four).

Creating a mean, and using t-statistics with it, requires an interval scale. Satisfaction is not actually measured on an interval scale. This means that the use of means and a t-statistic are not a valid statistical test for changes or comparisons of satisfaction.

There is a test that uses the ranking provided in this type of satisfaction question. It also has fairly simple statistical tests. The Wilcoxon Rank Sum and Kruskal-Wallis Chi-square (χ^2) test is specifically designed to test statistical significance for ordinal ranking data.

This test was used to analyze whether the satisfaction ranking was statistically significantly different over time. The tests provide a chi-square approximation which can be used for statistical significance. The tests were performed using the nonparametric procedure in SAS® (NPAR1WAY) with the time periods as the class variable.

The Kruskal-Wallis chi-square test did find that the trends seen in the increasing satisfaction with quality of work and a decreasing satisfaction with electric energy savings were statistically significant (Table 6).

Program participants were generally satisfied with all program components except for the computer printout showing electricity usage. We found these results to be consistent across all three time periods.

Components showing a decline in satisfaction were electric energy savings, and the computer printout that showed electricity usage. Why these elements have lower satisfaction and what can be done to improve them was noted as probably being worthy of further exploration and program improvement trials.

The elderly/non-elderly analysis shows that both groups exhibit a trend of declining satisfaction for electric energy savings, as can be seen in Table 7. There is a difference, however, between the trends exhibited for the satisfaction with the computer printout that shows electric energy use. There is an increasing satisfaction trend among the non-elderly for the computer printout while the elderly show more mixed results. Elderly participants generally show a decrease in satisfaction with the computer print-outs but the ratings do not provide a consistent trend over time. The program may want to monitor the satisfaction of the computer printout separately for the elderly and non-elderly as they move forward or making any changes to this offering.

Table 7. Satisfaction by Program Component – Elderly and Non-Elderly (unweighted)

Elderly

Program Component Percent Satisfied (rating a '4' or '5') with ...	Time Period 1	Time Period 2	Time Period 3
Number of Responses (n)	109	122	140
Process for Scheduling Audit	82%	83%	87%
Quality of Work	86%	88%	86%
Attitude of Initial Contractor	n/a	92%	92%
Attitude of Follow-up Installation Contractor	93%	87%	85%
Amount of Work Performed	82%	79%	84%
Electric Energy Savings	77%*	71%*	57%*
Explanation of How Appliances Use Energy	78%	78%	76%
Actions Recommended to Saving Energy	79%	82%	74%
Computer Printout Showing Electric Energy Use	75%**	80%**	65%**
Program Overall	83%	89%	85%

* Statistically significant at probability greater than Chi-square at 0.001 (greater than 99%) via a Kruskal-Wallis test.

** Statistically significant at probability greater than Chi-square at 0.18 (greater than 80%) via a Kruskal-Wallis test.

Non-Elderly

Program Component Percent Satisfied (rating a '4' or '5') with ...	Time Period 1	Time Period 2	Time Period 3
Number of Responses (n)	198	53	72
Process for Scheduling Audit	76%	87%	80%
Quality of Work	74%	81%	85%
Attitude of Initial Contractor	n/a	84%	86%
Attitude of Follow-up Installation Contractor	84%	88%	91%
Amount of Work Performed	74%	79%	74%
Electric Energy Savings	66%**	67%**	57%**
Explanation of How Appliances Use Energy	72%	69%	74%
Actions Recommended to Saving Energy	77%	69%	78%
Computer Printout Showing Electric Energy Use	69%**	74%**	79%**
Program Overall	77%	81%	78%

** Statistically significant at probability greater than Chi-square at 0.16 (greater than 84%) via a Kruskal-Wallis test.

This analysis also provides an example of the importance of examining differences in demographics between the time periods and how that could affect the appropriate interpretation of the time trend analysis.

Both owners and renters exhibit statistically significant trends at the 80% level for increasing satisfaction for the quality of work and decreasing satisfaction for electric energy savings. This is shown within Table 8 below. The earlier analysis of satisfaction with the program components, as shown in Table 6 above, led an evaluation reviewer to hypothesize that these trends could have been caused by the difference in the populations (more owners than renters) over time rather than actual program trends. The finding that the same trends were exhibited by owners and renters in both cases shows that the trends are not caused by the demographic differences but are likely actual program trends.

Table 8. Satisfaction by Program Component – Owner and Renter
(unweighted)

Owner

Program Component Percent Satisfied (rating a '4' or '5') with ...	Time Period 1	Time Period 2	Time Period 3
Number of Responses (n)	129	135	160
Process for Scheduling Audit	75%	85%	85%
Quality of Work	73%**	87%**	85%**
Attitude of Initial Contractor	N/a	91%	91%
Attitude of Follow-up Installation Contractor	71%	87%	89%
Amount of Work Performed	72%	79%	80%
Electric Energy Savings	65%*	69%*	57%*
Explanation of How Appliances Use Energy	73%	76%	74%
Actions Recommended to Saving Energy	73%	79%	76%
Computer Printout Showing Electric Energy Use	69%	76%	68%
Program Overall	73%	85%	83%

* Statistically significant at probability greater than Chi-square at 0.03 (greater than 97%) via a Kruskal-Wallis test.

** Statistically significant at probability greater than Chi-square at 0.18 (greater than 80%) via a Kruskal-Wallis test.

Renter

Program Component Percent Satisfied (rating a '4' or '5') with ...	Time Period 1	Time Period 2	Time Period 3
Number of Responses (n)	178	40	55
Process for Scheduling Audit	80%	80%	72%
Quality of Work	82%**	83%**	84%**
Attitude of Initial Contractor	n/a	87%	85%
Attitude of Follow-up Installation Contractor	93%	89%	72%
Amount of Work Performed	80%	82%	84%
Electric Energy Savings	73%**	74%**	58%**
Explanation of How Appliances Use Energy	74%	75%	77%
Actions Recommended to Saving Energy	81%	76%	76%
Computer Printout Showing Electric Energy Use	72%	86%	83%
Program Overall	84%	90%	82%

** Statistically significant at probability greater than Chi-square at 0.18 (greater than 80%) via a Kruskal-Wallis test.

Advantages of Conducting “Over Time” Process Evaluations

Examining issues over time can add a significant amount of information and depth to a process evaluation. This is the appropriate way to examine whether changes being made as part of program maturation to increase the quality of the program actually are doing so. In the LIPA-REAP exam, the program staff interviews had indicated that the program had been improving with program maturation, the quality of services provided had improved. This was verified by the over time analysis where participant satisfaction with the quality of work performed had consistently increased across the time periods and this increase was statistically significant. The demographics in the program had changed over time, with increases in the proportion of participants that are elderly and home ownership. Yet, analysis was conducted that examined the trends in quality of work performed and found that this trend was not strongly dominated by any one population category. The increase in the satisfaction with quality of work is, therefore, a firm finding and not an artifact of the changing demographics.

Similarly, the decreasing satisfaction with electric energy savings was found to be a consistent and statistical significant trend overall and among each of the demographic groups examined. So this is also a definitive finding and not an artifact of the changing demographics with program maturation. One hypothesis is that the program's first year found it recruiting, or being able to recruit, those with the greatest opportunities with electric energy savings. This would create greater satisfaction with the electric savings for those participating in the early years than those in later years. If greater opportunities can not be found for electric savings among the low-income households served by REAP, the program may need to ensure that program literature describes the opportunity in a way to ensure that people can be served while not creating unrealistic expectations about electric savings.

Program managers knew they had more problems with the new refrigerators used for replacement during a specific time period. (Once discovered, the problem was corrected.) The over time surveys were able to ascertain if this had created significant problems in overall participant satisfaction. The analysis found that participant satisfaction with their replacement refrigerators was high and there were no significant differences between the time periods.

During the initial process evaluation the program had discovered relatively low recollection of the “Action Plans”, an important part of the energy education system provided. So the program made changes to the “Action Plans” that they had hoped would help improve its effectiveness. However, the

over time analysis found that the recollection of the Action Plans had not improved. Therefore, the changes made were not effective and other changes would need to be tested.

The participants' satisfaction was rated on a scale from one to five (where five is very satisfied). This is a very common practice among process evaluation surveys. This is ordinal ranking data rather than interval data. As such, means and t-statistics are not valid methods for testing statistical significance of comparisons or changes. The most appropriate test is one that takes into account that this is an ordinal ranking rather than an interval scale. A Wilcoxon Rank Sum and Kruskal-Wallis Chi-square (χ^2) test was used to analyze whether the satisfaction ranking was statistically significantly different over time. This test was relatively straightforward to perform. It proved useful and allowed valid meaningful interpretations of the comparisons. It is recommended that this test be used for similar evaluations in the future.

References

- Blalock, Jr., Hubert M. (1979) *Social Statistics*. McGraw-Hill Book Company. New York: New York. Pp. 367 – 369.
- Megdal & Associates (2004) *2003 Process Evaluation of LIPA's Clean Energy Initiative REAP Program* submitted to KeySpan Energy Corporation on behalf of the Long Island Power Authority.
- SAS/STAT® User's Guide: Volume 2, GLM-VARCOMP*, Version 6 Fourth Edition, SAS Institute, Inc. Cary: North Carolina, pp. 1193 – 1210.