# Chapter 2 - Evaluation Overview and Methods

## *Background*

The New York Independent System Operator (NYISO) collaborated with wholesale electricity market stakeholders, including NYSERDA, in 2000-2001 to develop and implement emergency and economic demand response programs to access customers' abilities to shed load in response to high prices and/or situations where the reliability of the electricity grid might be jeopardized.  NYSERDA participated in the NYISO working group that created these price responsive load (PRL) programs and developed complementary Enabling Technologies for Price Responsive Load Management and Peak Load Reduction programs to promote expanded participation.

During the Fall of 2001, an evaluation of the these programs, commissioned jointly by NYISO and NYSERDA, was conducted by Neenan Associates with support from the Consortium for Electric Reliability Technology Solutions (CERTS), particularly by staff from Lawrence Berkeley National Laboratory (LBNL) and Pacific Northwest National Laboratory (PNNL).  The study included several components including surveys of customers, to improve the understanding of participant demographics, curtailment strategies and satisfaction with the programs, and an analysis of customer performance data to quantify benefits for participants and for the overall marketplace (e.g., price reductions, reliability enhancements, etc.).
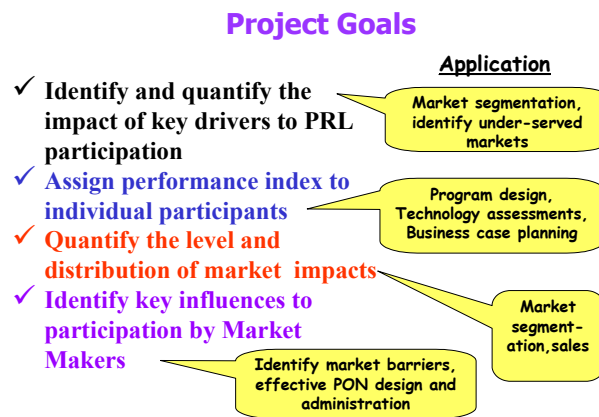


**Fig. 2-1: Evaluation Project Goals**

The feedback from this evaluation assisted NYISO and NYSERDA in:  1) quantifying the benefits of customer participation; 2) determining what aspects of the NYISO programs were attractive to customers and which ones needed to be modified; and 3) modifying NYSERDA program offerings to target lowering barriers to participation.

**2002 NYISO PRL Evaluation**


The NYISO's PRL programs were continued in 2002 and NYSERDA continued to provide funding for enabling technologies though Program Opportunity Notices (PONs). Consequently, these entities desired to extend the comprehensive evaluation of the previous year with two new areas of focus. First, the 2002 PRL program analysis focused on characterizing barriers to participation in DADRP. Although the number of subscribers to DADRP increased slightly, bid activity and the amount of scheduled curtailments was lower in 2002 than in 2001. Because day-ahead market participation is widely viewed as being a critical element of a robust electricity market, identifying barriers to participation in DADRP was deemed to be of the utmost importance for this year's evaluation. Accordingly, this year's survey and subsequent analyses were specifically designed to better characterize those barriers.

NYSERDA funding for 2001 and 2002 was focused on demonstrating the value of enabling technologies to customer PRL program participation, with the expectation that by doing so, firms that manufacture and sell such devices would be enticed into the market and assume the role of recruiting and servicing participants to PRL programs as a means of creating demand for their products. Moreover, commodity retailers and LSEs would use the available PON funds to create customer interest in switching from the default POLR service to their competitive offerings. Finally, PON funds were expected to increase market entry by specialized curtailment service providers (CSPs) seeking to develop a profitable portfolio of PRL resources.

The presence of diverse and committed market makers is an important element of developing the overall retail market structure. After two years of experience, NYSERDA desired to characterize the role its funding plays in how these businesses viewed PRL program participation as a business proposition. So, while last year's process analysis concentrated on how LSEs and CSPs viewed NYSERDA PON performance, with regard to meeting their immediate needs, this year's analyses focused on characterizing how PRL was viewed by existing and potential market makers as contributing to their long-run business goals and interests. Thus, a different survey and evaluation methodology were developed and implemented.


*Project Team*

**2002 NYISO PRL Evaluation**

The evaluation of 2002 PRL program performance was conducted by Neenan Associates and a team of researchers associated with the Consortium for Electricity Reliability Technology Solutions (CERTS).[1] NYISO and NYSERDA provided funding for Neenan Associates, which was responsible for project management and deliverables. Funding for the CERTS team was provided by the Department of Energy. The analysis involved almost a dozen researchers that contributed collectively over five man-years of effort.
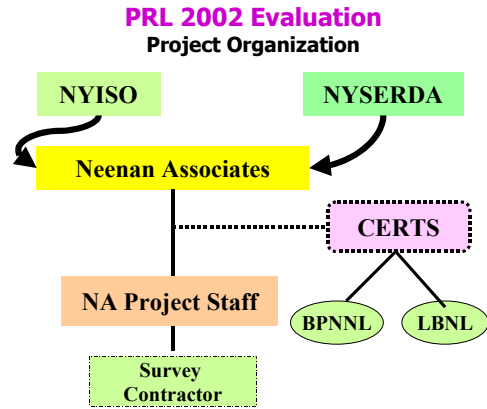


**Fig. 2-2: Evaluation Project Organization**

## *Approach*

As was the case last year, the project team analyzed NYISO market data to quantify the actual MW reductions, the improvements in system reliability and the impacts on electricity prices. The contribution of NYSERDA PON participants was derived from these overall PRL program benefits. In addition, a survey of program participants and non-participants was implemented to: 1) characterize customer preference for various PRL programs; 2) assess customer familiarity with NYSERDA programs and whether/why they chose to participate or not participate in them; 3) determine the important correlations among customer characteristics (e.g., sector, size, load curtailment strategy) and PON participation; 4) determine the level of satisfaction with PON and PRL program features and obtain recommendations for improvement; and 5) evaluate customer needs and payback expectations regarding enabling technologies.

The instrument developed last year served as the basis for this year's survey, but some important modifications were made to accommodate this year's special focus on DADRP. As a result, the survey administration process differed from that of 2001 whereby surveys were mailed to customers, 111 of which filled them out and returned them to Neenan. This year's survey was administered to 144 customers directly by means of a telephone interview, two-thirds of which

---

[1] The CERTS researchers are associated with Lawrence Berkeley National Laboratory and Pacific Northwest National Laboratory.

**2002 NYISO PRL Evaluation**

were conducted by a vendor, and one-third by the CERTS team research scientists. Like last year, several prizes were awarded by a lottery, as an inducement to participate in the survey.

To guide the survey design and evaluation effort, a set of hypotheses was constructed to reflect the issues that NYISO, NYSERDA, and other stakeholders identified as requiring more information before much needed resolution could be achieved. Based on discussions with the NYISO and NYSERDA and others, such as the NYSDPS, end-use customers and customer representatives, the project team drafted a set of issues and corresponding hypotheses that were then circulated for review.

These hypotheses then served as the foundation for the survey design and subsequent analyses. The hypotheses were constructed as testable propositions. Each posed a question, the answer to which could be construed as affirming the proposition, or lending doubt as to its validity, using accepted statistical methods. To ensure that the results of the analysis of these propositions contributed to issue resolution, the propositions were constructed to minimize Type I errors (accepting that the proposition was true, when in fact the survey results did not support such a conclusion).  An example is provided below.

$H_o$: Particularly "comfort-sensitive" customers are less likely to participate in PRL programs than other customers

$H_a$: Comfort sensitive customers participate at the same rate, which implies that the program design is not biased against such customers

Two survey versions were developed to test, in part, these hypotheses.  First; a base survey that would be administered to customers by a vendor via scheduled telephone interviews about 20 minutes in length was created. The time constraint limited the breadth of questions that could be asked and dictated that most responses had to be closed ended (respondents chose from an established list of alternatives). This base survey then became the foundation for developing a second instrument, called the PRL audit.

This enhanced survey was designed to be administered by experienced engineers, which allowed greater latitude in recording customers' responses to the questions asked. By probing issues with respondents, the interviewer would be able to record subtle but important nuances that distinguish customers and contribute to explaining behavior.  In addition, the PRL audit, which required forty-five minutes to complete, included additional inquiries. The genesis of the PRL audit was research conducted by the CERTS team last year, when they developed and field tested
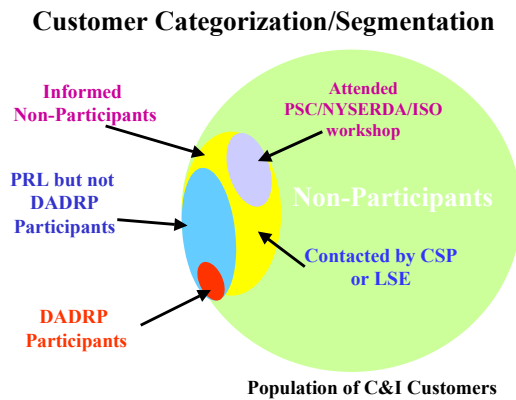
**2002 NYISO PRL Evaluation**

protocols for gathering extensive data on customers' equipment inventory, characteristics, and usage that would help resolve many issues related to why customers are reluctant to participate, or participate in only a limited fashion in PRL programs, despite an apparent larger capability.

The base survey instrument was designed in three stages. In the first, a base rate and rank instrument was developed using many of last year's questions to develop a longitudinal database on preferences and customer characteristics. New sections were added to address the focus on DADRP and to explore customer preferences for alternative bidding methods, using the hypotheses as the foundation for what questions to ask. Finally, structural changes were made to the instrument to accommodate the direct administration of the survey by a vendor.

Subsequently, an alternative instrument was developed in which the research team identified areas in the base survey where, due to ambiguity about customer circumstances or narrow interpretations of wording, the questions explored only the surface of a deeper issue. The CERTS staff then developed more, in depth probing protocols and a complete PRL audit was prepared, and reviewed.

The data for the surveys described above can be used to evaluate customers' revealed preferences. Each was confronted with a decision to participate or not, and the data collected can be used be used to characterize what factors were most important in the decisions. However, the results are only applicable to situations where the exact same programs are offered. They do not provide insight into the response to different program configurations.

A set of conjoint-type questions was added to both surveys. Respondents were asked to make 20 separate choice decisions. In each, they were offered alternative program designs each described by a specific but different level of five feature characteristics (event notice, event duration, curtailment benefit level, noncompliance penalty level, and start time). The responses to these questions provide the data needed to develop a stated preference choice model that associates customers' likeliness to participate with program features.

**Customer Categorization/Segmentation**



Fig. 2-3: Customer Segmentation

**2002 NYISO PRL Evaluation**

### *Survey Administration*

The research team tested the base survey and the PRL audit instruments with selected customers. Based on the results, refinements were made to adjust the wording to better fit customers' perspectives, and length of the instrument was adjusted to fit the target completion time.

The base survey was administered to participants and non-participants by a survey vendor during September and early October. The CERTS teams conducted the PRL audits in the same period. Four sample frames were constructed. Three were compiled from NYISO subscription records that contain the names of all PRL program participants, which were sorted into three categories: those that participated in DADRP, those that participated in EDRP only, and those that participated in EDRP and ICAP/SCR. (Customers that participated only in ICAP/SCR were not included in this analysis.) The three categories constructed are not exclusive since all DADRP participants also participated in EDRP and some in ICAP/SCR. However, because of the focus this year on barriers to DADRP participation, this partition was necessary to ensure that the questions on the survey were properly addressed.

The fourth sample frame was constructed to represent non-participants, customers that did not join the program this summer. It was comprised of customers that attended one of six briefings on PRL programs conducted around the state in April and May by NYSERDA, NYSDPS, and NYISO. The workshop introduced customers to the programs, demonstrated how program provisions worked and provided examples of the potential benefits of participation.

Over 300 customers attended a workshop, about one-quarter of which (75) joined one or more PRL programs in 2002. The remaining customers constitute a subpopulation of informed non-participants (INP), customers that were provided with extensive information about the programs, but elected not to participate in 2002. Last year, the INP sample frame was constructed from names and addresses provided by LSEs of customers that they had contacted about program participation. The means by which customers were contacted varied widely, from participation in a workshop to receiving a letter or bill stuffer announcing the program, which provided insight into the value of information to the decision to participate of not. But, the lists were not representative of the population in general, so extrapolation of the results was difficult.

Because this year's survey was conducted through a telephone interview, telephone contact information was required for all customers in the INP sample frame. This requirement

**2002 NYISO PRL Evaluation**

made compiling the sample frame from LSE and CSP contact lists impractical. Instead, the population of workshop attendees was used to represent non-participants, albeit they likely do not represent the population of all customers. LSEs and CSPs used these workshops as a means of informing their customers of the programs, and they likely were biased toward larger customers or customers with which they have established a relationship that goes beyond the usual communication of information. Second, customers that attended are likely those that either had previous experience with a similar program, have or are considering the installation of enabling technologies, or have usage patterns conducive to PRL program participation. The survey results are described in Chapter 3.

## *Data Sources and Uses*

Data used in the analysis consisted of secondary data acquired from NYISO, and primary data collected directly from customers via surveys administered by the project team. Secondary sources of collected data are illustrated in the table below and include the following:

| Project Database Elements, Sources, and Uses | | |
|---|---|---|
| **Input** Import or manual data entry (some range checking) | | |
| **Retrieval** Queries for counts and reports | | |
| **Data Elements** | **Source** | **Use** |
| Participant subscription information | NYISO registration forms | Sampling frame for participant survey administration |
| Non-participant information | PSC and CSP sponsored workshop lists | Sampling frame for participant survey administration |
| CSP and host utility information | NYISO CSP list | Participation analysis |
| Event and performance data (computed) | NYISO | Analysis of event performance |
| Survey administration data | Neenan | Track survey administration (unique ID, mail merges, sent & reply dates, etc.) |
| Survey response data | Survey instruments | Report and evaluate end-use response to participation, response and program features |
| Other end-use firm related data | Survey instruments and/or follow-up interviews | Additional data for elasticity analysis and participant segmentation |

**Table 2-1: Project Data Requirements**

- Program subscription and performance data bases

- NYISO hourly prices (LBMPs) and load

- Customer Survey - a survey developed and administered to PRL program participants and other customers for the purposes of characterizing their satisfaction with the programs and collecting data that can be used to quantify how program features contribute to their willingness to participate and respond to curtailment events.

- PRL Audits – a more detailed, complex, and adaptive survey instrument completed by a randomly selected group of participants. It includes a detailed equipment inventory representing the participant's load management capability, and information about the firm's operation and objectives.

## *Evaluation Plan*

A careful analysis of the responses to the 2002 customer acceptance survey will help answer a number of key questions about participation, performance, and customer acceptance of the NYISO Demand Response Programs. Answers to these questions are of particular interest to the NYISO, NYSERDA and DOE, the project funders, to the NYSDPS, in order to craft public policy, and CSPs seeking to operate successful retail PRL programs. Moreover, these findings also have implications for the design of and participation in similar programs that might be implemented elsewhere in the country as part of FERC's standard market design.

Much of the initial analysis of these survey results will focus on differences between informed non-participants and on participants in EDRP and/or ICAP/SCR. There is keen interest in knowing more about participants in DADRP, but there are still only a small number of them. Some analysis can be attempted with these customers as part of a general analysis plan, but much of what we learn about customers in DADRP will be gained through the extended analysis of the data collected through the PRL audits.

As with the 2001 evaluation, one of the primary objectives of this year's PRL evaluation will be to better understand customers' decisions regarding participation in the NYISO's several demand response programs. It is perhaps convenient to think of these decisions as falling into four groups. We would like to use these data to better understand customers':

- Current Participation Decisions

- Continued or Future Participation Decision

- Load Reduction Subscription Rates

- Actual Event Performance.

Current participation decisions include those by informed non-participants not to enroll in any program and program participants that have elected to enroll in one or more of the NYISO's three programs: ICAP/SCR, EDRP, and DADRP. Despite the substantial increase in enrollment

**2002 NYISO PRL Evaluation**

this year in ICAP/SCR and EDRP, it is still critical to gain a better understanding of what motivates the decision to enroll in a PRL program. Furthermore, these programs are new, and continue to evolve; we must know which customers would continue in the programs if critical program features were changed.

Subscription rates indicate the load customers initially plan to curtail during an emergency event, or, in the case of DADRP, in real-time, if their bids were accepted in the DAM. If a customer belongs to both EDRP and ICAP/SCR, participation levels may differ by program, reflecting the different performance requirements and measurements.

Clearly, these decisions about program participation and performance are jointly determined by the characteristics of customers (e.g. type of business, number of hours open, number of production shifts, peak time of electricity use, etc.), the particular features of the various programs, and perhaps, even by conditions in the market (e.g. expectations about the level of wholesale prices in the DAM or in the RTM, etc.). Factors affecting decisions by new participants in 2002 may differ from those firms also in the programs during 2001. Financial assistance from NYSERDA or others in purchasing or installing load management equipment is hypothesized to influence decisions, as could past experience with load management programs and the usefulness of information received about the current programs. We gain important insights into how these factors interact to influence the customers' decisions through two levels of analysis.

### *Top Level Analysis*

The first, top-level analysis involved a careful examination of the survey raw data and the construction of some basic frequency tables, summary statistics, and cross tabulations. No analysis should proceed without a solid understanding of these data. In this "top" level analysis, much can also be learned about these important decisions through some straightforward hypothesis tests about differences in the means of key measures of satisfaction or preference between important subgroups of the survey respondents.

|  | Participation Status | |
|---|---|---|
| | Participated | Did Not Participate |
| Attended | 40 | 10 |
| Did Not Attend | 25 | 25 |

*Attendance Status*

**Table 2-2: Participation in NYSERDA/PSC Workshop**

Neenan ASSOCIATES

CERTS
CONSORTIUM FOR ELECTRIC RELIABILITY TECHNOLOGY SOLUTIONS

**2002 NYISO PRL Evaluation**

The hypotheses constructed to guide the analyses were evaluated using chi-square tests for independence of table rows and columns.  For example, in the (hypothetical) cross-tab shown, 80% of those who attended (a briefing) participated (in a PRL program), while only 50% of non-attendees participated.  The chi-square procedure is used to determine the likelihood that the two dissimilar rows could in fact be samples from the same population (i.e. with the same underlying probability of participation).  If this probability is sufficiently small (5% is a common threshold), the (null) hypothesis that they are from the same population is rejected, in favor of the hypothesis that the rows represent different populations.  For the values in the example, the likelihood that two such different proportions would result from random samples of the same population is less than 0.2%.  The null hypothesis would be rejected in favor of the hypothesis that briefing attendance is significantly associated with participation.

## *Comprehensive Analysis*

Informative as these simple hypothesis tests can be individually, however, they do not account for other factors that might be related to the differences that led to rejecting or failing to reject some of the hypothesis tests. In more in-depth analyses, we attempt to control for these other factors by constructing theoretically consistent behavioral models and applying more extensive multiple regression analysis. The details of the evaluation methods associated with these extended analyses are discussed below.

## *Evaluation Methods*

**Choice modeling** – Two different choice modeling activities can be performed using the collected survey responses.  First, conjoint survey questions asked customers to choose between alternatives with different features. By imposing behavioral assumptions (consistent with economists' notion of demand) on conjoint data to characterize customers' decision-making behavior, this choice model utilizes econometric techniques to quantify the relative contribution of individual features to the value the customer realizes from participation, the results of which are interpreted as the impact of features on the likelihood of program participation. Once fully configured, the choice model supports the evaluation of alternative program designs, represented by alternative feature levels, on expected participation, which is useful for both program design and modeling expected participation and price response.

**2002 NYISO PRL Evaluation**

A second choice model can be developed to explain firm's current PRL program participation decision.  Self-reported firm characteristics and actions taken by New York State agencies, market participants, and other institutions are used as predictors in assessing the likelihood of a customer's choice to join an emergency program (EDRP and/or ICAP/SCR), the day-ahead program (DADRP), or no PRL program whatsoever.  Such a model provides important insight into the kind of customers who are likely to join a PRL program.

**Models for Evaluating Customer Satisfaction and Price Responsiveness**

| | |
|---|---|
| **Satisfaction** | **Customer solicited ratings and rankings** |
| **Arc Elasticity** | **Measures participant's average performance over all events. No other explanatory factors included.** |
| **Demand Elasticity** | **Full behavioral specification of demand that accounts for price and non-price factors that effect usage** |
| **Choice Model (1)** | **Uses stated preferences for alternative program designs to evaluate how customers value program features** |
| **Choice Model (2)** | **Uses self-reported firm characteristics to indicate customer participation in current PRL programs** |

**Fig. 2-4: Evaluation Models**

**Market price simulation** utilizes a statistical representation, developed from historical data, of how supply conditions influence market-clearing prices to estimate what the prices would have been if the PRL curtailments had not materialized. This method is easier to apply, but its accuracy depends upon the degree to which a statistical model can capture the peculiarities of market pricing that led to extreme prices, and the ready availability of market characteristic data such as constraints and generation availability.

**Price Elasticity**: Two different measures of elasticity are of interest:

The **own-price elasticity** measures how consumption of electricity varies with respect to the price paid for electricity. Generally, data over an extended period of time where the price of electricity varied are required to estimate this elasticity, although if electricity consumption is considered to be truly discretionary, e.g., foregoing air conditioning for a few hours, then PRL curtailments are consistent with this measure of price responsiveness.

The **substitution elasticity** measures how firms facing time-varying electricity prices alter their usage to shift electricity from the higher priced periods to other times, which is the case for PRL load curtailment situation where customers do not forego usage altogether, but instead re-adjust the timing of its consumption.

**2002 NYISO PRL Evaluation**

Elasticities can be measured simply, using the arc elasticity method, or derived from a complete representation of the customer's demand of electricity. The more simple arc elasticities are derived from event performance, calculated as the change in the customers' usage, relative to the CBL, during the event divided by the change in price, measured as the difference between the PRL price, either explicit or implicit, and the basic tariff or contract price the customers would normally pay. The data needed for such calculations are readily available.

Estimating fully specified demand equations and deriving the substitution elasticity can produce a more insightful representation of response. The substitution elasticity measures a customer's ability and willingness to produce outputs using different levels of inputs, which characterizes the underlying production of service process in a physical sense.   The substitution of interest here is between electricity at times of high prices (during events) and other times when prices are lower. The higher the substitution elasticity, the more price responsive the customers. Estimating substitution elasticities for individual customers requires interval data for the entire period during which the customer participates in the program (usually the summer months) along with weather data and firm characteristics (operating or output measures, labor schedules, etc.) necessary to account for factors other than price that influence changes in load from hour to hour.

**Other Performance Indicators** that provide insight into the character of customer participation and curtailments include:

- Curtailment performance relative to subscription measures how well customers estimated their ability to respond when they registered for the program. Higher performance by this metric (under equivalent price incentives and penalties) indicates that the participant understands its capabilities well, and therefore will performs more uniformly over all events. Low performance variance is useful to system dispatchers when they consider deploying the available resources, and want to predict the outcome as precisely as possible.

- Curtailment performance relative to CBL measures what proportion of the current level of usage the customer curtails when an event is called. Higher performance by this measure is valuable as it lowers resource acquisition and transactions cost per delivered kWh of curtailed load.

## *Methods Employed*

The analyses conducted are summarized in the table below and described in more detail in the chapters that follow.

**Table 2-3: Summary of Evaluation Methods and Data Requirements**

| Method | Description | Data Requirements |
|---|---|---|
| Top End Analysis | Use statistical tests to evaluate rate and rank survey responses and test hypotheses. | Survey responses for both Base survey and PRL audit |
| Curtailment responsiveness | Characterize individual and group response to events | |
| Arc demand elasticities | Price-weighted simple price elasticities | Event CBL and curtailments, and base service electricity rate |
| Performance Indices | Metrics based on relative measures of load curtailment capability | Event CBL and curtailments |
| Behavior Modeling | Characterize how observable customer characteristics survey responses contribute to the decision to participate | |
| Revealed Preferences | Define characteristics and factors that explain why customers chose to participate or not | Responses to rate and rank survey questions, and customer characteristics. |
| Stated Preferences | Use customer choices in hypothetical decision situations to deduce the value of product characteristics to likelihood of participation. | Responses to conjoin survey questions and customer characteristics. |
| Market Impacts | How curtailments effected market prices | Hourly LBMPS and corresponding loads for the DA and RT markets, by zone, and other market condition information such as available generation and transmission node constraints. |
| Reliability Benefits | The value of curtailments in preventing forced outages | |
| Collateral Benefits | How price changes are transformed into lower purchase costs to consumers. | |