IMPROVING THE METHODS USED TO EVALUATE VOLUNTARY ENERGY-EFFICIENCY PROGRAMS

Martin Schweitzer and Marilyn A. Brown
Oak Ridge National Laboratory

ABSTRACT

Despite progress that has been made in recent years, further improvements are needed in the methodologies commonly used to evaluate the energy savings arising from voluntary energy-efficiency programs. These voluntary programs are characterized by the fact that they do not involve mandatory codes or standards but instead use information and incentives to further the adoption of energy-efficient technologies and practices. Voluntary programs frequently are aimed at long-term transformation of markets that make lasting changes in consumer patterns of energy use. To date, many of the evaluations of such programs have focused on the direct effects to program participants and have not addressed the associated market transformation to the extent possible. Using information gathered through an extensive methodological review, the authors describe useful approaches taken in previous evaluations and draw conclusions concerning the best methods available for forecasting and measuring the impacts of voluntary programs.

Introduction

This paper presents key findings, to date, from an ongoing study of voluntary energy-efficiency programs being conducted by Oak Ridge National Laboratory (ORNL) for the U.S. Department of Energy (DOE) and the U.S. Environmental Protection Agency (EPA). Specifically, DOE and EPA commissioned a review of the literature to determine the best methods available for evaluating three of their “marquee” programs: Building America, Rebuild America, and Energy Star. These programs seek to promote lasting structural and behavioral changes in energy-efficiency markets through the voluntary actions of consumers. They do this by providing information, decision tools, technical assistance, performance monitoring, and labeling to certify energy efficiency. While these market interventions are often delivered to consumers one participant at a time, the goal is to create broad, sweeping, and sustained impacts on entire markets for energy efficiency.

Each of these three voluntary programs is described below.

- **Building America** is a private/public partnership designed to produce new homes using 30-50% less energy than typical homes. Building America forms teams of architects, engineers, builders, equipment manufacturers, material suppliers, lenders, construction trades, and community planners. These teams take a “systems engineering approach” to perform research and identify cost-effective energy-efficient building methods and technologies that will make up the “next generation” of building systems. DOE funds five teams—including more than 50 different companies—that design, build, and test prototype homes. The new techniques employed in those houses can then be used by the partners (and others) in their other construction efforts. DOE’s money is used for research, training, and monitoring, but not for construction (U.S. Department of Energy 2000).

- **Rebuild America** encourages the formation of partnerships of local governments, private businesses, and other organizations that help communities identify and solve problems related to energy use in commercial buildings, public facilities, and multifamily housing units. The program
provides the partners with the technical tools (e.g., handbooks, workshops, referrals to technical and business experts) necessary to plan and execute building retrofit projects that make use of innovative technologies and approaches. Each partnership chooses target buildings, sets goals for energy savings, seeks financing, and decides how to implement their project. The underlying goal is to accelerate the adoption of energy-efficient techniques and practices in commercial, institutional, and multi-family residential buildings (U.S. Department of Energy 2000).

- **Energy Star** is a voluntary program, sponsored by DOE and EPA, that is designed to increase consumer awareness of, and desire for, energy-efficient products and structures. It has three separate foci: equipment; homes; and non-residential buildings. The part of the program that focuses on equipment operates by awarding the Energy Star label to appliances and other equipment that meet stringent standards for energy efficiency. It is targeted at manufacturers, retailers, and consumers. The homes element bestows the Energy Star label on those residential structures that meet certain energy-efficiency specifications. The non-residential buildings component involves the formation of partnerships between the federal sponsors (DOE and EPA) and the owners and operators of commercial and institutional buildings. The non-federal partners are provided with information regarding energy-efficient structures and are given the opportunity to obtain the Energy Star Buildings label for their structures (U.S. Environmental Protection Agency 2000).

Programs such as these, that seek to transform the market for energy-efficient products and services, are significantly different from demand side management (DSM) and financial assistance programs that focus only on direct participants, and they must be evaluated in fundamentally different ways. While the direct impacts of programs like Building America, Rebuild America, and Energy Star Buildings (e.g., energy saved in buildings constructed or retrofit by participating partners) are certainly important, the additional spillover effects to non-participants throughout the country are potentially very substantial and must also be measured. The time horizon of the effects of market transformation interventions can be considerably longer than for traditional rebate-style resource acquisition interventions, and the impacts can be more diffuse and therefore more difficult to trace (Sebold et al. 2001).

Assessments of these three programs to date have primarily focused on the direct effects on program participants. It is the purpose of this paper to provide information on the best available methods for conducting future evaluations that go beyond this participant focus and address the market transformation effects of these programs. This paper summarizes the most important ideas gathered from our extensive literature review and concludes with a set of recommendations for evaluating voluntary energy-efficiency programs.

### Market Effects and Market Barriers

When attempting to quantify how markets have been transformed as a result of energy-efficiency programs, it is necessary to distinguish between market *changes*, which are influenced by energy and product prices as well as by the program effort, and market *effects* resulting just from the program itself. Market effects can include changes in a variety of areas, such as the energy-efficiency attributes of manufacturers’ products, the stocking and pricing behaviors of retail and wholesale businesses, and customer awareness and purchasing patterns (Eto, Prahl & Schlegel 1996). Market effects generally result from program efforts that reduce or eliminate pre-existing barriers to the adoption of energy-efficient practices and technologies (Fig. 1).
While commonly referred to as market barriers, the factors that impede the adoption of new technologies can be further divided into “market failures” and “market barriers.” Market failures occur when there is a flaw in the way markets operate. They are conditions of a market that violate one or more neoclassical economic assumptions that define an ideal market for products or services such as rational behavior, costless transactions, and perfect information (Brown 2001; Jaffe and Stavins 1994). Market failures can be caused by:

- misplaced incentives (as in master-metered apartment buildings when landlords pay the utility bill and tenants have no incentive to conserve);
- distortionary fiscal and regulatory policies (as in electricity pricing policies where prices do not reflect the real-time cost of electricity production);
- unpriced costs (such as air pollution and other environmental externalities associated with the extraction, production, distribution, and consumption of fossil fuels); and
- insufficient and incorrect information (e.g., about the performance of new technologies and the contribution of different appliances to a home’s electricity bill).

“Market barriers” refer to obstacles that are not based on market failures but which nonetheless contribute to the slow diffusion and adoption of energy-efficient innovations (Levine et al. 1995; DOE Office of Policy and International Affairs 1996). These barriers include:

- low priority of energy issues (as occurs when potential energy savings from a particular technology are small on an individual basis but large when summed across all consumers),
- capital market imperfections (such as the interest rate gap, where energy suppliers can obtain capital at lower interest rates than can energy consumers), and
- incomplete markets for energy-efficient features and products (as when energy efficiency is one of a large number of features that come in a package and cannot be purchased as an option).
The existence of market failures and barriers that inhibit socially optimal levels of investment in energy efficiency is the primary reason for considering public policy and program interventions. In many instances, feasible, low-cost policies and programs can be implemented that either eliminate or compensate for market imperfections and barriers, enabling markets to operate more efficiently to the benefit of society.

**Theory-Based Evaluation**

To accurately assess the effects of a market transformation program, the program evaluator must understand the underlying “logic” or “story” of the program in question. This means that the intended process by which the program (or “market intervention”) addresses existing barriers to the use of energy-efficient products and practices and makes lasting changes to the pre-program market structure must be understood and clearly described. The basic premise of “theory-based evaluation” is that it is important to articulate the way(s) in which a program will change its targeted market(s) so that the actual changes can be compared to a forecast of projected outcomes (Hastie et al. 2000). “Market influence diagrams,” “market effects tables,” and “dynamic models” are all good ways of explicitly showing what a particular program is expected to achieve and the mechanisms by which this will be accomplished (Fig. 2).

**Fig. 2.** Example of Market Influence Diagram (Source: Eto, Prahl, and Schlegel, 1996)

Characterizing the Market
An evaluation of a market transformation program should begin with a “market characterization study,” which examines the structure and dynamics of the market prior to a program intervention. In general, it is most cost effective to gather these baseline data at the beginning of the program. A market characterization study would include a description of a market’s size, its key participants, and the process by which products are manufactured, sold, and installed (Bronfman 1998). Such a study provides a pre-program baseline description and allows subsequent market changes to be identified (Peters et al. 1998).

Caulfield, Richardson, and Ridge (2000) provide an excellent example of a market characterization study, which they conducted for Pacific Gas and Electric Company’s (PG&E) Food Service Technology Center (FSTC). Their market structure diagram (Fig. 3) shows that by working with manufacturers, PG&E is able to influence the decisions made by end users, both directly and through intermediaries such as designers, dealers, and buying groups. The diagram also identifies all important market actors and illustrates the complexity of their interactions. For example, the diagram shows that the FSTC is directly involved with manufacturers, who interact with their own representatives as well as with large customers. Manufacturer representatives, in turn, are engaged with designers and dealers, who have direct relations with end users. At the same time, customers are also directly influenced by a variety of other market actors.

Market changes can be revealed through market effects studies undertaken as a follow-up to the initial market characterization work. The Northwest Energy Efficiency Alliance uses an exemplary approach whereby they develop a detailed baseline of information (including hypothesized market barriers) and then track changes and attribute effects through the use of frequent market progress evaluation reports (Kunkle & Lutzenhiser 1999).

The diffusion of innovation model (Fig. 4) posits distinct stages in the adoption process and emphasizes information and communication flows and change over time. It suggests that it is important to
Data Collection Techniques

Much of the data used to evaluate market transformation programs—particularly in the early stages before good data on sales and energy savings are available—come from surveys and interviews with a wide range of market actors (Bronfman 1999). Surveys can be administered in person, by mail, over the telephone, or via the Internet. A recent example of web-based data collection is provided by Martin et al. (1999). In this study, an on-line questionnaire was used to gather information from users of Industrial Assessment Center web sites regarding the energy and cost savings resulting from their participation in the program. While self-selection bias must be guarded against, this method of soliciting data can be very cost-effective.

The examination of reports and articles documenting the findings from past efforts can also be very useful. The use of focus groups, which involves intensive interaction with key market participants, is a good way to identify market barriers and failures (Peters et al. 1998). Focus groups and surveys also can be used to determine how a program might better reach the various market actors and lead to greater adoption of energy-efficient technologies and practices.

Tracking sales and price data and collecting information on product stocks maintained by wholesalers and retailers provide additional sources of information (Brown, Webber & Koomey 2000; Nadel 1999). Interviews with appliance dealers was a major source of data for the evaluation of the Super Efficient Refrigerator Program. Data bases kept by the Association of Home Appliance Manufacturers and the California Energy Commission were also very useful. In addition, the content of newspaper advertisements was examined (Lee & Conger 1996). New York is requiring program implementation contractors to collect pre- and post-program market information (e.g., stocking practices, sales and prices), and then subjecting these data to an independent audit prior to use in evaluations (Nadel 1999).
When conducting studies of market effects, it is important to get information from enough different types of market actors, and a sufficient number of each, so that the findings will be representative of the entire market. Data collected from multiple market segments of participants is required to adequately understand market activities. Studies that had a limited view of the market or too few data points have difficulty drawing conclusions (Peters et al. 1998).

**Analytical Methods**

Determining that a particular program has had a market effect involves ascertaining that a change in the market has occurred and that there is a plausible mechanism linking that change to the program intervention. To show such a causal link between an energy-efficiency program and a market change, one can identify relevant market barriers and failures that the program is designed to address and develop measurable indicators of their existence and magnitude. Subsequent measurements that show a diminishment in those factors, coupled with the occurrence of intended changes in the market, provide strong evidence of a program effect (Peters et al. 1998).

Because market transformation programs are intended to make lasting changes in market structures over time, it makes more sense to use time-series designs that track the same market for an extended period than to use cross-sectional designs that compare conditions in different areas (Conlon, Weisbrod & Samiullah 1999). A share capture model, which can be used to extract market effects from overall market changes, was successfully employed in an evaluation of the Green Lights program and is potentially useful for measuring the effects of the Energy Star equipment labeling effort (Horowitz, Lewis & Coyle 2000). Differences in responsiveness to product and electricity prices between a comparison period (with no intervention) and a treatment period (when the Green Lights Program operated) were examined. The share capture model produces coefficients that allow estimates to be made of the number of electronic ballasts that penetrated the market because of the price effect and as a result of the program itself.

Factor analysis can also be useful in analyzing data. An innovative approach to identifying whether or not market transformation has taken place (but not the extent of the change) is the binomial test used by Boston Gas Company to evaluate its programs (Spellman et al. 2000). This method uses a large number of indicators of change and runs a statistical test to see if a convincing majority of those indicators has moved in the appropriate direction.

When measuring the direct effects of information and technical assistance programs on the participants (as opposed to broader market effects), traditional evaluation techniques—such as tracking the adoption of targeted technologies and practices, measuring pre- and post-treatment energy use, and comparing performance by treatment and control groups—continue to be effective (Riggert et al. 1999; Green & Skumatz 2000; Hicks & Von Neida 2000). Figure 5 illustrates a typical pre- and post-treatment design with control group.
Leading and Ultimate Indicators of Success

Indicators of success are simply measurable factors that show whether a desired program effect has been achieved. These indicators can measure ultimate program outcomes, such as the actual adoption of new technologies and the associated energy savings. However, there is often a considerable lag time between the initiation of a market transformation program and the availability of these kinds of data, due to the time-consuming need for a variety of market actors to become aware of new products and modify their behavior. To address that situation, many evaluators use “leading indicators” as evidence of changes in the market that can be observed prior to the availability of sales or energy consumption data (Lee & Conger 1996). The indicators used can vary depending on how long the program under study has been in operation (Berry & Schweitzer 1999).

Effective leading indicators can include the following (DeCotis et al. 2000):
- changes in awareness of the targeted product or service by customers and mid-market actors;
- changes in market actors’ intentions;
- changes in products offered by manufacturers (e.g., number of different Energy Star products produced);
- increase in stocking, offering, or promotion of the product in question by current providers;
- growth in the number of providers;
- the depth and breadth of customers and mid-market actors;
- decrease in the unit cost of energy-efficient products; and
- number of program participants (e.g., number of builders and developers involved in Building America).

![Fig. 5. Schematic Diagram of Pre- and Post-Research Design With Control Group (Source: Soderstrom, Berry, Hirst, and Bronfman, 1981)](image)

The Super Efficient Refrigerator Program used a number of leading indicators to assess progress during the first several years of program operation (Lee & Conger 1996). These included:
· training provided by manufacturers;
· dealer promotional activities;
· number of SERP models on retail floor;
· SERP refrigerator prices; and
· customer awareness of SERP refrigerators.

Interviews with appliance dealers were a major source of data for these indicators. Other important sources included data bases kept by the Association of Home Appliance Manufacturers and the content of newspaper advertisements. Similarly, to assess a program to promote high-efficiency gas furnaces, telephone surveys of furnace contractors were conducted and written sales data were examined. By doing this in both the state where the program operated and in a neighboring state, the evaluators were able to estimate the program’s market effects (Kushler, Schlegel & Prahl 1996).

Ultimate indicators of success (Sebold et al. 2001), which tend to be measurable later in a program’s life span, include:
· the number of energy-efficient products or services adopted by consumers;
· number of square feet of floor space retrofitted;
· actual energy savings associated with the use of energy-efficient products or services;
· fuel bill savings resulting from utilization of the targeted products or services; and
· associated reductions in greenhouse gas emissions.

It is possible to identify "near-term indicators," "intermediate indicators," and "final indicators." For example, in an energy-efficient motors program, a near-term indicator could be the retail cost difference between standard and premium-efficiency motors early in the program's life, an intermediate indicator could be the number of customers familiar with premium efficiency motors, and a final indicator could be the resulting energy savings from use of energy-efficient motors (Berry & Schweitzer 1999). Boston Gas Co. developed 270 market indicators to measure progress for four market transformation programs, an average of 68 indicators for each one. Four general categories of indicators of market effects were examined: product awareness and promotional activity, including attitudes among end-users; product knowledge, including level of training and expertise among trade allies; product performance and reliability; and product availability and penetration. Different indicators were developed for each market actor or issue. This method allows market effects to be measured over short time periods (Spellman et al. 2000).

Conclusions

Based on the literature review conducted for this study and the experience gained from ORNL’s long history of evaluating energy-efficiency programs, the following recommendations are offered for evaluating programs such as Building America, Rebuild America, and Energy Star:
(1) understand the underlying “logic” of the program in question and the mechanisms by which it is intended to change its targeted market;
(2) begin with a “market characterization study” that examines the pre-program structure and dynamics of the market;
(3) conduct periodic “market effects studies” as follow-ups to the market characterization work;
(4) take a lesson from the diffusion of innovation model and examine the awareness and attitudes of market actors as well as the relevant social networks;
(5) select appropriate data collection methods, which are likely to include surveys and focus groups involving a wide range of market actors;
(6) to establish that a program has induced an observed market change, look for a reduction in the market barriers and failures that the program is designed to address;
because market transformations occur over time, time-series analysis tends to be more effective than cross-sectional designs; and additional innovative approaches should also be considered; and “leading indicators” of market effects should be used to evaluate programs relatively early in their life span, while “ultimate indicators” of program effects become more attractive later in time. Several new research methods exist that enable program evaluators to reveal specific effects of market transformation programs. By going beyond impacts from direct participants to impacts on the market at large, a more complete picture of a program’s transformational properties can be obtained.

References


