Title

Program Evaluation and Incentives for Administrators of Energy-Efficiency Programs: Can Evaluation Solve the Principal/Agent Problem?

Author
Blumstein, Carl

Publication Date
2009-03-30

CSEM WP 184

Program Evaluation and Incentives for Administrators of Energy-Efficiency Programs: Can Evaluation Solve the Principal/Agent Problem?

Carl Blumstein

March 2009

This paper is part of the Center for the Study of Energy Markets (CSEM) Working Paper Series. CSEM is a program of the University of California Energy Institute, a multi-campus research unit of the University of California located on the Berkeley campus.

2547 Channing Way
Berkeley, California 94720-5180
www.ucei.org

Program evaluation and incentives for administrators of energy-efficiency programs: Can evaluation solve the principal/agent problem?

Carl Blumstein
University of California Energy Institute
2547 Channing Way
Berkeley, California 94720-5180
Email: blumstei@berkeley.edu

Abstract

This paper addresses the nexus between the evaluation of energy-efficiency programs and incentive payments based on performance for program administrators in California. The paper describes problems that arise when evaluators are asked to measure program performance by answering the counterfactual question, what would have happened in the absence of the program? Then the paper examines some ways of addressing these problems. Key conclusions are 1) program evaluation cannot precisely and accurately determine the counterfactual, there will always be substantial uncertainty, 2) given the current state of knowledge, the decision to tie all of the incentive to program outcomes is misguided, and 3) incentive programs should be regularly reviewed and revised so that they can be adapted to new conditions.

Introduction

California policy puts energy efficiency "first in the loading order" (CPA, CEC and CPUC 2003). This policy has lead to substantial funding for energy-efficiency programs aimed at reducing the consumption of electricity and natural gas. In California these programs are administered by electricity and natural gas utilities. Utility administration has given rise to a principal/agent problem.

In the principal/agent problem one party, the principal, hires another party, the agent, to take actions on his behalf. The principal wants the agent to take actions that will make the size of some performance criterion, such as the net value of saved energy, as large as possible. The outcome depends on the agent's actions and decisions, his technological and economic opportunities, and on chance. The principal can observe none of these directly though he may know some information about the range of technological and economic opportunities and he may know the probabilities attached to the possible chance outcomes. The principal's problem is to design a mechanism for compensating the agent that will induce the agent to come as close as possible to maximizing the principal's performance criterion.[1]

In our case, the principal is the regulator (the California Public Utilities Commission (CPUC)) and the agents are the regulated energy utilities. The principal/agent problem is how to design an incentive mechanism that will cause the energy utilities to maximize some performance criterion for the energy-efficiency programs that they administer.

An obvious initial problem is that utilities are in the business of selling energy. Historically, at least to some degree, utility profits have depended on the volume of sales. This disincentive for the promotion of energy efficiency has long been addressed in California and a number of other US states by "decoupling" (See, for example, Eto *et al*. 1997). Decoupling breaks the link between utility sales and utility earnings by adjusting rates when actual sales differ from

[1] This description of the principal/agent problem is a shortened paraphrase of the description given in Joskow and Schmalensee (1986).

projected sales. If sales are above expectations, rates are lowered to hold earnings constant; if sales are below expectations, rates are increased to hold earnings constant. Under decoupling a successful energy-efficiency program will not directly cause a decrease in utility earnings.

In addition to direct effects on utility earnings, energy-efficiency programs have what Eto *et al*. (1998) call "hidden" costs. According to Eto *et al*., hidden costs consist of both the very real, but unobservable, management costs associated with the additional effort and organizational changes required to implement successful energy-efficiency programs, and the opportunity costs associated with net lost revenues from activities [like the construction of new plant] foregone because of the pursuit of energy-efficiency programs. There may also be hidden benefits to the utility for administering energy-efficiency programs. These might include a larger base over which to spread overhead costs and the avoidance of competition from other entities that might interfere with future business opportunities. However, these benefits are associated with the control of program resources, not necessarily with good program performance.

Efforts to find incentive mechanisms that will counteract the hidden costs of good program performance and align utility interests with the public interest have been underway for some time. Interest in the problem in the first half of the 1990's waned as a result of the trend toward industry restructuring and deregulation in the US (CPUC 2003). Failures of restructuring, particularly the failure of the California electricity market (Blumstein, *et al*. 2002), and increased concern about climate change have reinvigorated the search for improved incentive mechanisms (see, for example, Jensen 2007).

In California the search for better incentive mechanisms is far along with new rules promulgated in 2007(CPUC 2007a). The new rules reward good performance by a utility by giving the utility a share of the customer savings that result from the utility's energy-efficiency programs. But, while there is disagreement about what is wrong, there is a general consensus that things are not working well in California. Recently the CPUC suspended the 2007 rules and initiated a new rule making to re-examine the incentives (CPUC 2009).

Solution of the principal/agent problem can be straightforward when two conditions hold: 1) the principal has a single, accurately quantifiable objective for the agent to pursue and the agent's contribution to achieving the objective can be separated from the contributions of other factors, and 2) the proper alignment of financial incentives is necessary and sufficient to insure that the agent will act in the principal's interest. Unfortunately, neither of these conditions holds in California today.

At the heart of the difficulties in California are problems of evaluation. The regulator's objective in California is to "maximize achievement of cost-effective energy efficiency" (CPUC 2008). In California the intention is to tie utility incentives to utility performance in achieving this goal. Utility performance is supposed to be determined by the evaluation of its energy-efficiency programs. In the three-year period 2006-2008 about $2 billion was made available for utility energy-efficiency programs. Approximately 5 percent of these funds were dedicated to program impact evaluation, which was carefully separated from program implementation. Evaluators were selected by the regulator and were not allowed to have any engagement with the utilities. The evaluator's problem is to determine a counterfactual—that is, what would have happened if a utility program had not existed. Unfortunately, as this paper discusses, this problem is proving to be difficult to solve. Attempts to determine counterfactuals are consuming very substantial program evaluation resources in California without doing much, if anything, to improve program outcomes.[2]

[2] That there are difficulties in determining counterfactuals certainly does not mean that evaluation should be abandoned. Evaluation can and does play an important role in assisting program administrators in improving the design and conduct of the programs that they administer. (See, for example, Vine 2008 and Peters and McRae 2009)

An assumption implicit in California's approach to utility incentives is that the principal/agent problem will be solved solely by providing financial rewards for good outcomes. This is almost certainly too simple a view about how complex organizations like utility companies behave. As Eto *et al*. (1998) point out, organizational changes are needed when a corporation adopts new objectives. Failure to make these changes can seriously compromise energy-efficiency program performance. For example, when energy-efficiency programs are placed too far down in the corporate hierarchy, programs may be subject to inappropriate controls by low-level procurement officers and risk managers. Without immediate access to top management, decision making can be impaired and innovation can be stifled. Failure to make the necessary organizational changes can occur even when the changes are in the utility's interest. One reason this may happen is that such changes create both winners and losers among the corporation's employees. This can result in an internal principal/agent problem for the corporation. Because they stand to lose from change, some of the corporation's employees (its agents) resist change and obscure the need for change from the corporation's leadership.

The rest of this paper is organized as follows, first, a discussion that provides more detail about what's not working and then a discussion of possible ways that the situation can be improved. Key conclusions are that 1) program impact evaluation cannot precisely and accurately determine the counterfactual, there will always be substantial uncertainty, 2) given the current state of knowledge, the decision to tie all of the incentive to program outcomes is misguided, and 3) the incentive programs should be regularly reviewed and revised so that they can be adapted to new conditions.

What's not working?

Bias in favor of measures that produce quantifiable results

A well-recognized problem is that the current incentive mechanism biases the utilities' efforts in favor of results that can be quantified by program evaluators. Under the current California rules, incentives are tied to customer savings. These incentives create a bias in favor of utility programs that produce direct effects on customer energy use as opposed to programs that have indirect effects. In practice, this means that the incentive mechanism encourages the utilities to favor programs that directly result in the installation of energy-efficiency measures as opposed to programs, like public education or contractor training, that lead only indirectly to energy-efficiency actions because customer savings from programs with indirect effects are more difficult for program evaluators to quantify.

This is an example of the long-standing tension between "resource acquisition" and "market transformation." In resource acquisition utilities "acquire" energy efficiency as a substitute for new supply. In market transformation the objective is to change conditions in the market so that energy-efficient actions are taken without the need for subsidies or other interventions. As discussed below, market transformation and resource acquisition are complementary strategies since resource acquisition programs create conditions that lead to market transformation and market transformation programs create conditions that lead to participation in resource acquisition programs. If energy prices and energy-efficiency technology were static, market transformation would be an end point for energy-efficiency efforts. In practice, prices and technology are changing and the promotion of energy efficiency involves repeated cycles of subsidy, education, training, and the promulgation of performance standards (see, for example, Blumstein *et al*. 2000)

The bias in favor of quantifiable measures makes it unlikely that the balance between resource acquisition and market transformation will be near the optimum. What typically happens now is that at the beginning of a program cycle funds are set aside for market transformation programs and utilities earn a small fixed percentage of program costs.

Difficulties in measuring "free riders" and "spillover"

A key evaluation issue that must be addressed in dealing with the principal/agent problem is identification of the consequences of the agent's actions as opposed to the consequences of other factors. In California one of the ways this is playing out is debates about "free riders" and "spillover."

In the parlance of energy-efficiency programs, free riders are participants in a program who receive an incentive payment or other assistance but would have acted even without the program. To first order, payments to free riders do not accomplish anything, they are just transfers from one set of consumers (the non-participants) to another (the free riders)[3]. It is not desirable to reward utilities for the energy savings of free riders for two reasons: 1) the payments are unearned and 2) payments for free-rider savings would bias utility programs in favor of programs in which consumers already had a strong predilection to participate.

But, identifying free riders so that free-rider savings can be excluded from the calculation of incentive payments is easier said than done. Current practice is to determine who is a free rider by asking program participants a series of questions to determine if it was their intention to act even in the absence of the program. But this is not reliable. As Peters and McRae (2008) point out,

> "The self-report method for measuring free-ridership assumes intentions are [perfect predictors of] behavior. If someone reports, "I would have done it anyway," they are assigned a free-ridership value of 100%. Yet any student of behavior knows that, while better than attitudes and beliefs, intentions are only a weak predictor of behavior."

Evaluation of free-ridership can be more sophisticated than simply asking a direct question about intentions. For example, the New York State Energy Research and Development Authority (NYSERDA) employs a multi-question survey approach that has evolved from its own experience and insights from similar research in other states. NYSERDA relies on experienced interviewers who are knowledgeable enough to probe respondents for details of program influences and who can characterize the responses in quantitative terms (Saxonis 2007). However, in spite of their greater sophistication, these methods continue to suffer from the difficulties associated with determining counterfactual behavior from self-reported intentions.

"Spillover" is the other side of the free rider issue. Spillover occurs when the effects of an energy-efficiency program spill over to affect other behavior. Examples of spillover would be a consumer taking action as the result of an energy-efficiency program but not receiving any of the direct benefits offered by the program (non-participant spillover) or a program participant stimulated to pursue additional energy saving actions that are not subsidized by the program (participant spillover). Spillover might occur because a consumer was persuaded by advertising associated with the program, or as a result of contact with satisfied program participants, or because the existence of the program has caused suppliers to change the products and services that they offer. The idea here is that the existence of large-scale energy-efficiency programs has a transforming effect on the market. Consumers see efficient technologies at work, practitioners learn by doing, and suppliers change their stock of goods and services.

While spillover is obviously desirable, it is not taken into account in the California incentive mechanism.[4] This is, at least in part, because spillover is difficult for program evaluators to quantify. One way to determine spillover effects is to look at cross-sectional data. Horowitz

[3] Possible second order effects include that the free rider's experience with the program might cause him to encourage others to participate in the program.

[4] In a Decision in October 2007 the CPUC (2007b) directed its staff to explore during 2008-2009 the ability to credibly quantify and credit non-participant spillover. The CPUC intends to modify its evaluation protocols to include spillover if this proves to be feasible.

(2008) compares consumption data from US states with strong commitments to energy-efficiency programs to data from US states with weak commitments to energy-efficiency programs. He finds evidence of substantial spillover within the states with strong commitments to energy-efficiency programs.

Identifying the spillover from a specific energy-efficiency program or portfolio of programs is more challenging. Hoefgen *et al*. (2008) describe an effort to assess spillover from a program to promote the sale of compact fluorescent lamps (CFLs) in Massachusetts. In the evaluation two methods were used to construct a counterfactual. Here the counterfactual, called the "baseline" by Hoefgen *et al*., is the CFL sales that would have been made if there had been no program promoting the sale of CFLs. The first method relied on state-level sales from selected states with active programs, including Massachusetts, along with national CFL shipment data. The evaluators subtracted sales in areas with active programs from total national sales and treated the per-household CFL sales level in the remaining states as the counterfactual for per-household sales in Massachusetts. The second method used a single-state comparison area, Michigan, to construct the counterfactual. Hoefgen *et al*. find a very large spillover—they estimate that CFL sales due to spillover are *greater* than CFL sales subsidized by the program. However, the two methods produce estimates of the counterfactual that differ by about 20 percent—not especially large as these things go, but very consequential when millions of dollars in incentives are on the line.

Difficulties in determining cost effectiveness (the non-energy benefits problem)

"All cost-effective energy efficiency" is easier said than measured. The problem arises because it can be difficult to separate the costs of energy-efficiency actions from other costs. In California, cost effectiveness is determined by the Total Resource Cost (TRC) test. In this test the net present value of supply costs avoided by an energy-efficiency measure are compared to the total cost of the measure. The total cost of the measure includes both costs paid by the energy-efficiency program and costs paid by the program participant.

For example, if an energy-efficiency program provides a subsidy of 20 percent for the incremental cost of installing double glazed windows, then the total cost includes both the 20 percent provided by the program and the 80 percent provided by the participant. The problem in this example is that people install double glazing for a number of reasons in addition to energy saving. These other reasons include reduction in interior noise and greater comfort when sitting near windows because of reduced radiant heat loss. It is appropriate to allocate all of the cost paid by the energy-efficiency program (that is, the utility's costs) to saving energy. However, it is not appropriate to allocate all of the program participant's costs to saving energy—some of these costs should be allocated to the other benefits (often referred to as non-energy benefits) that motivated the participant's expenditures. Other examples of non-energy benefits include positive impacts on occupant health and productivity and reduced maintenance costs from installation of energy-efficient equipment and the adoption of energy-efficient practices (Birr and Singer 2008).

The consequence of the difficulties in separating out non-energy benefits is a bias in favor of simple measures where non-energy benefits tend to be small. Opportunities where relatively small subsidies could cause design changes that promote both energy efficiency and non-energy benefits are lost. An example is comprehensive whole-house retrofit programs in which investments are undertaken by homeowners to gain a range of non-energy benefits (Knight *et al*. 2006). This kind of lost opportunity is serious because a key part of the energy-efficiency agenda is to make energy efficiency an integral part of processes like home renovation.

Delays in obtaining results

Evaluation takes time. For instance, in California program impact evaluations for program activities conducted in the 2004-2005 period had yet to be fully completed as of January 2009. This kind of delay in completing evaluations causes delays in incentive payments based on performance and weakens the link between incentives and performance.

What might be done?

Just do it. Pick some indicator such as the difference between actual and forecast and forge ahead

One potential way forward is to simply commit to the best available performance evaluation procedures and be done with it. In deciding whether to pursue this approach one must weigh the trade off between the regulatory costs associated with an ongoing process against the cost of introducing inefficient biases into the energy-efficiency program.

As discussed above, current incentive schemes in California are biased in favor of resource acquisition strategies that are easy to quantify. This would not be serious if easily quantified resource acquisition were all (or most of) what needs to be done.

Do it better. (For example, use randomized controlled trials)

Evaluators, when confronted with the difficulties in evaluating the performance of utility energy-efficiency programs and, especially, the free-rider problem, often respond with suggestions for improvements in evaluation methods. One approach that is often recommended is randomized controlled trials (RCTs). In an RCT a population is randomly divided into two groups, a treatment group and a control group. The treatment group is treated (for example, with the offer of an incentive to take some energy-efficiency action) and the control group is not treated. Subject to some conditions, comparison of the behavior of the treatment group with the behavior of the control group will allow us to identify the effect of the treatment.

The key condition that must hold is that control group is not influenced in any way by the treatment. In practice this means that the treatment group and the control group are completely isolated from each other—the groups must be small enough so that the likelihood of interaction between them is very small. That is, the opportunity for spillover is very small.

If an RCT is designed so that the control group is not influenced by some intervention to encourage energy efficiency, then we will be well along in answering the question, what is the effect of this energy-efficiency intervention on the treatment group? Unfortunately, an RCT that satisfies this non-influence condition is not likely to be of much interest in the evaluation of the impact of utility programs. The problem is that, as noted above, the treatment group in an RCT that satisfied the non-influence condition would necessarily be quite small. But, what the regulator is trying to accomplish *requires* large programs. Programs need to be large because spillover and structural change in markets are essential parts of the desired program outcomes.

This is not to say that efforts to improve the evaluation of energy-efficiency programs should be abandoned. While counterfactuals cannot be known with certainty, it may be that the development of improved evaluation techniques will allow us to narrow the range of uncertainty. Even if the range of uncertainty cannot narrowed, it would be useful to have quantitative estimates of uncertainty to guide our use of evaluation tools and manage expectations for what can be accomplished through impact evaluation.

Evaluation also has an important role to play in improving the design and conduct of energy-efficiency programs (Vine 2008, Peters and McRae 2009). New approaches may be valuable in this regard. For example, although RCTs cannot tell us about spillovers, they can provide useful information for the design of energy-efficiency programs since it is useful for purposes of program design to know how an isolated individual will respond to a treatment such as an incentive

Outsource market transformation

Another approach to the bias toward resource acquisition in California's current incentive schemes is to outsource market transformation programs like advertising and education and training to another entity, typically a not-for-profit corporation. An example of this approach is the Northwest Energy Efficiency Alliance (NEEA, see http://www.nwalliance.org/). NEEA operates market transformation programs that are coordinated with a number of utilities and agencies that operate energy-efficiency programs in the US Pacific Northwest.

A difficulty with this approach is that resource acquisition and market transformation are not always easy to separate. Ideally, the two strategies are complements, not substitutes. For example, one might want to provide incentives to manufactures to develop and bring to market more efficient washing machines (a market transformation strategy) and also provide subsidies to consumers to purchase the more efficient washing machines when they come on the market (a resource acquisition strategy). Thus, there are obvious advantages to coordination between market transformation programs and resource acquisition programs. Although NEEA apparently coordinates market transformation efforts among many organizations with some success (Northwest Economic Research 2008), such coordination is typically easier inside a single organization as opposed to between two or more organizations. The case for a separate market transformation organization is best when the area encompassed by the market is much larger than the areas served by the resource acquisition programs.[5]

The creation of an additional organization also increases the difficulty of the identification problem. If we want to tie the incentive payment to the effects of a utility's actions, we will now need to separate the effects of the market transformation organization from the effects of the utility's energy-efficiency program.

Difficulties notwithstanding, the outsourcing option should not be taken off the table. It provides an opportunity to create a state-wide program as opposed to several utility service area programs. The experience of NEEA suggests that a not-for-profit market transformation organization can be effective.

The existence of a separate organization for market transformation makes more tangible the possibility that an alternative organization could take over the operation of some or all of the utility administered programs. The threat of entry by competitors may create a stronger incentive for the utilities to perform well. The importance of the threat of entry will depend on the extent to which the benefits to a utility from administration of energy-efficiency programs derive from the control of resources (as opposed to good program performance). Of course, the threat of entry may also cause the utilities to be reluctant to cooperate with potential competitors.

"Professionalize" the practice

Part of the incentive problem is that while we want to provide incentives for "good" conduct it is often difficult to tie good conduct to good results in an unambiguous way.

Drawing an example from an apparently unrelated area, consider the treatment of disease. The physician's role in the treatment of disease is to supplement the body's own defenses, but sometimes this is not necessary and sometimes it is not sufficient. That is, sometimes the patient will recover without treatment (a potential medical free rider) and sometimes the patient will not

[5] See Blumstein *et al*. 2005 for further discussion of the advantages and disadvantages of alternative administrative arrangements.

recover even with treatment. Given the uncertain relationship between treatment and outcome, how should we compensate the physician? The answer is often that the physician's compensation is not dependent on outcome. Rather, in the US, she receives a fee for her services regardless of outcome. Fee-for-service compensation is effectively, in the language of utility regulation, a cost-plus arrangement.

What then prevents the physician from giving treatments to every patient, even patients who the physician knows will recover without treatment? The answer is that we rely on professionalism. Professionalism for the physician includes, in addition to scientific and technical knowledge about the practice of medicine, a set of norms about ethical practice. These norms proscribe unnecessary treatment. To the extent that this works[6], it is primarily because the physician internalizes the profession's norms. There may also be external sanctions either from peer groups or from the legal system. But, external sanctions are usually associated with damage to the patient as a result either of treatment or failure to treat. This is not to say that self regulation based on professionalism is not facing challenges (see, for example, RCP 2005).

Is there any similarity between health professionals and the practitioners of energy efficiency? While the analogy can certainly be pushed too far, I think there is some similarity. Consider this definition of "profession" from the Royal College of Physicians (2005),

> "An occupation whose core element is work based upon the mastery of a complex body of knowledge and skills. It is a vocation in which knowledge of some department of science or learning or the practice of an art founded upon it is used in the service of others. Its members profess a commitment to competence, integrity and morality, altruism, and the promotion of the public good within their domain."

This definition could fit the work of many energy-efficiency practitioners. Not only does the practice rest on a complex body of knowledge, but also many of the practitioners come to the profession for very altruistic reasons. It would be easier to use professional norms to regulate the quality of energy-efficiency programs if we understood better how such regulation works. But, while it is hard to imagine the practice of medicine without the guidance of strong professional norms, we do not know how to create such normative structures for new professions. That said, it is now the case that regulators give no weight to professionalism of the staff in the construction of incentives to encourage good performance by utility companies. Given the obvious power of professional norms and the altruism of many energy-efficiency practitioners, this appears to be a serious mistake.

Include some non-quantitative measures in the evaluation of performance

Since some of the important characteristics of good energy-efficiency programs are difficult to quantify, it is probably desirable to include some non-quantitative measures in the evaluation of utility performance. Examples of such measures are discussed below. Although they are non-quantifiable, they can be graded—that is, judged to be poor, fair, good, excellent, etc. Grades might be determined by a panel of expert reviewers.

An example characteristic that is hard to quantify is corporate commitment. As noted above, organizational changes are needed when a corporation adopts new objectives. The organization that results from these changes is one measure of corporate commitment. Where in the corporation hierarchy is the energy-efficiency program placed? How many management layers are there between the energy-efficiency program's manager and chief executive officer? Does the energy-efficiency program have supportive arrangements for legal services, personnel, and

---

[6] There is some evidence that patients in programs that compensate the physician based on the number of patients in the physician's practice (referred to in the US as capitation) receive less treatment than patients whose physicians are compensated on a fee-for-service basis (Gosden *et al*. 2000). In the US there is increasing interest in the establishment of plans that pay physicians for performance (Rosenthal and Dudley 2007).

purchasing? Characteristics like these may be difficult to quantify, but they are certainly observable and important to take into account.

Another characteristic that is difficult to quantify but observable is support for professionalism. An evaluation of support for professionalism could address questions such as, Has the corporation succeeded in hiring and retaining a strong cadre of energy-efficiency professionals? Are there career paths for professionals? Are there training opportunities that support professional development?

Reduce performance incentives

The alternative to providing incentives is cost-plus compensation. Joskow and Schmalensee (1986), who are proponents of incentive regulation, are nonetheless wary of incentive schemes that get it wrong. In their view the greater the uncertainty in assessing an agent's performance, the less incentives should be part of compensation. As uncertainty grows they recommend moving toward (but never all the way to) cost-plus compensation.

Shifting toward cost-plus compensation would mean that most of the energy utilities' earnings from the energy-efficiency programs that they administer would be based on program expenditures, not performance. That is, utilities would be able to recover their costs plus an additional fee. The maximum allowed costs and the fee would be set in advance by the regulator. In California, where 100 percent of the energy utilities' above-cost compensation is now based on performance, this could be viewed as a step backwards.

The problem for the regulator is that the regulator wants to reward performance but measuring performance depends on a counterfactual that cannot be known accurately for reasons discussed above. The regulator must find a balance between the desire to keep program incentives tightly focused on energy saving goals and the concern that the mechanisms for measuring program performance will create distortions in the conduct of the program and will cause uncertainty that frustrates program planning and discourages commitment to sustained effort.

Provide for regular review and, if necessary, adjustment of incentive mechanisms

When there are difficulties in assessing utility performance Joskow and Schmalensee suggest, ". . . that incentive schemes must be regularly redesigned, just as tariffs are now." However, they also caution that, ". . . compensation rules must be kept fixed for reasonably long periods (and utilities must anticipate that this will happen) if they are to have noticeable effects on behavior." Unfortunately, finding the balance between "regularly redesigned" and "fixed for reasonably long periods" is easier said than done.

Conclusion

In California the focus of program evaluation has shifted from its original purpose, which was to learn what needed to be done to improve the design of energy-efficiency programs, to a new purpose, providing the basis for incentive payments to energy-efficiency program administrators. Since the stakes are large—up to $450 million in incentive payments for the 2006-2008 program cycle—discussions about evaluation are likely to become increasingly adversarial and more likely to become the purview of advocates whose job is not to seek the truth but rather to make the best case possible for their clients. This is the wrong direction to be heading.

To change this direction we need first to recognize that program evaluation cannot precisely and accurately determine the counterfactual question, what would have happened in the absence of a utility's energy-efficiency programs? There will always be substantial uncertainty.

Next, we need to reduce the stakes. Given the current state of knowledge, the decision to tie all of the incentive to program outcomes is misguided. The advice of Joskow and Schmalensee

(1986) about dealing with uncertainty in the measurement of utility performance, while given in a somewhat different context, is very relevant. This suggests that most of the utilities' compensation should be cost plus and only a relatively small share should be tied to the performance of the energy-efficiency programs that they administer.

This is not to say that performance incentives are unimportant. Rather it is to recognize the difficulties in quantifying good performance and to protect against the very perverse effects that can result from overreliance on poor performance measures. As long as the evaluation of performance continues to play some role in utility compensation there will be opportunities for the regulator to encourage improvements in performance. This might be done by making changes in the incentive mechanism but could also use other regulatory tools such as increased supervision of efficiency programs or the introduction of competing program administrators.

Finally, we need to be prepared to deal with changing circumstance and to exploit new knowledge. Again following Joskow and Schmalensee (1986), incentive programs should be regularly reviewed and revised so that they can be adapted to new conditions. These reviews will be less disruptive and less contentious if the amount compensation tied to performance evaluation is reduced.

References

Birr and Singer 2008. David Birr and Terry E. Singer, *NAESCO Analysis of Non-Energy Benefits of Efficiency Retrofits for ESCOs and ESCO Customers*. Washington, DC.: National Association of Energy Service Companies, http://www.naesco.org/resources/industry/documents/NAESCO%20NEB%20Report%2012-11-08.pdf (Accessed 2009-01-12)

Blumstein *et al*. 2000. Carl Blumstein, Seymour Goldstone, and Loren Lutzenhiser, "A theory-based approach to market transformation." *Energy Policy*, 28 137-144

Blumstein *et al*. 2002. C. Blumstein, L. S. Friedman, R. J. Green, "The History of Electricity Restructuring in California." *Journal of Industry, Competition and Trade,* 2:1 9-38

Blumstein *et al*. 2005. Carl Blumstein, Charles Goldman, Galen Barbose, "Who should administer energy-efficiency programs?" *Energy Policy,* 33 1053-1067

CPA, CEC and CPUC 2003. California Power Authority, California Energy Commission and California Public Utilities Commission, *Energy Action Plan.* http://www.energy.ca.gov/energy_action_plan/2003-05-08_ACTION_PLAN.PDF. See also http://www.energy.ca.gov/2008publications/CEC-100-2008-001/CEC-100-2008-001.PDF. (Accessed 2008-12-21)

CPUC 2003. California Public Utilities Commission, *History of Shared-Savings Incentive Mechanisms for Energy Efficiency Programs*. Attachment 2, Decision 03-10-057 October 16, 2003 http://docs.cpuc.ca.gov/word_pdf/FINAL_DECISION/30826Att2.pdf (Accessed 2009-01-15)

CPUC 2007a. California Public Utilities Commission, *Interim Opinion on Phase 1 Issues: Shareholder Risk/Reward Incentive Mechanism for Energy Efficiency Programs*. Decision 07-09-043 September 20, 2007 http://docs.cpuc.ca.gov/PUBLISHED/FINAL_DECISION/73172.htm (Accessed 2008-12-26)

CPUC 2007b. California Public Utilities Commission, *Interim Opinion On Issues Relating To Future Savings Goals And Program Planning For 2009-2011 Energy Efficiency And Beyond*, Decision 07-10-032 October 18, 2007 http://docs.cpuc.ca.gov/word_pdf/FINAL_DECISION/74107.pdf (Accessed 2009-03-18)

CPUC 2008. California Public Utilities Commission, *California Long-Term Energy Efficiency Strategic Plan: Executive Summary*. http://www.californiaenergyefficiency.com/docs/EESP_ExecutiveSummary.pdf (Accessed 2008-12-28)

CPUC 2009. California Public Utilities Commission, *Order Instituting Rulemaking to Examine the Commission's Energy Efficiency Risk/Reward Incentive Mechanism.* Rulemaking 09-01-019 January 29, 2009 http://docs.cpuc.ca.gov/word_pdf/FINAL_DECISION/97023.pdf (accessed 2009-02-18)

Eto *et al*. 1998 J. Eto, S. Stoft, and S. Kito, "DSM Shareholder Incentives: Current Designs and Economic Theory." *Utilities Policy***,** 7:1 47-62

Eto *et al*. 1997. Joseph Eto, Steven Stoft and Timothy Belden, "The theory and practice of decoupling utility revenues from sales." *Utilities Policy***,** 6:1 43-55

Gosden *et al*. 2000. Gosden T, Forland F, Kristiansen IS, Sutton M, Leese B, Giuffrida A, Sergison M, Pedersen L. "Capitation, salary, fee-for-service and mixed systems of payment: effects on the behavior of primary care physicians." *Cochrane Database of Systematic Reviews 2000*, (3): Article No. CD002215.

Hoefgen *et al*. 2008. Lynn Hoefgen, Susan Oman, Angela Li, Gail Azulay, and Ralph Prahl, "Market effects: claim them now or forever hold your peace." In *Proceedings of the 2008 ACEEE Summer Study on Energy Efficiency in Buildings*, 9:150-161. Washington, D.C.: American Council for an Energy-Efficient Economy.

Horowitz 2008. Marvin J Horowitz, "Energy efficiency and changes in energy demand behavior." In *Proceedings of the 2008 ACEEE Summer Study on Energy Efficiency in Buildings*, 7:148-155. Washington, D.C.: American Council for an Energy-Efficient Economy.

Jensen 2007. Val R. Jensen, *Aligning Utility Incentives with Investment in Energy Efficiency*. National Action Plan for Energy Efficiency. http://www.epa.gov/cleanenergy/documents/incentives.pdf (Accessed 2008-12-29)

Joskow and Schmalensee 1986. Joskow, Paul L. and Richard Schmalensee, "Incentive regulation for electric utilities." *Yale Journal on Regulation*, 4: 1-49

Knight *et al.* 2006, Robert L. Knight, Loren Lutzenhiser, and Susan Lutzenhiser, "Why comprehensive residential energy efficiency retrofits are undervalued." In *Proceedings of the 2006 ACEEE Summer Study on Energy Efficiency in Buildings*, 7:141-150. Washington, D.C.: American Council for an Energy-Efficient Economy.

Northwest Economic Research 2008. Northwest Economic Research, "*Independent Evaluation of NEEA's Accomplishments: Retrospective Assessment 2008*." http://www.nwalliance.org/research/reports/2008_NEEA_Retrospective_Assessment.pdf (Accessed 2009-01-07)

Peters and McRae 2008. Jane S. Peters and Marjorie McRae, "Free-Ridership Measurement Is Out of Sync with Program Logic…or, We've Got the Structure Built, but What's Its Foundation?" In *Proceedings of the 2008 ACEEE Summer Study on Energy Efficiency in Buildings*, 5:219-234. Washington, D.C.: American Council for an Energy-Efficient Economy.

Peters and McRae 2009. Jane S. Peters and Marjorie McRae, "*Process Evaluation Insights on Program Implementation*." Berkeley, California: California Institute for Energy and Environment. http://ciee.ucop.edu/energyeff/documents/proc_eval_whtppr.pdf (Accessed 2009-03-17)

RCP 2005. Royal College of Physicians. *Doctors in society: medical professionalism in a changing world*. Report of a Working Party of the Royal College of Physicians of London. http://www.rcplondon.ac.uk/pubs/books/docinsoc/ (Accessed 2009-01-04)

Rosenthal and Dudley 2007. Meredith B. Rosenthal and R. Adams Dudley, "Pay-for-performance: will the latest payment trend improve care?" *Journal of the American Medical Association* 297:740-744

Saxonis 2007. William P. Saxonis, "Free Ridership and Spillover: A Regulatory Dilemma" In *Reliable Program Results? You Betcha!*, 533-545. Chicago, Illinois: International Energy Program Evaluation Conference.

Vine 2008. Edward Vine, "Strategies and policies for improving energy efficiency programs: Closing the loop between evaluation and implementation." *Energy Policy*, 36 3872–3881

Acknowledgements

Helpful comments and suggestions from Jeanne Clinton, Karen Notsund, Jane Peters, Ed Vine, and two anonymous referees are gratefully acknowledged.