Estimating the Benefits from Buildings Technologies: Issues, Challenges, and Lessons Learned from Blue Ribbon Panel Reviews

Andrew Nicholls and Sean McDonald, Pacific Northwest National Laboratory
John D. Ryan, U.S. Department of Energy

ABSTRACT

The Federal government often uses "Blue Ribbon" panels to evaluate and recommend changes to government programs. Such panels have been used several times over the last few years to evaluate the energy R&D programs of the U.S. Department of Energy (DOE). Several recent evaluations include:

- National Research Council (NRC) in 2001
- President's Committee of Advisors on Science and Technology (PCAST) in 1997
- Secretary of Energy's Advisory Board, Energy Task Force on Applied R&D chaired by Daniel Yergin (Yergin) in 1995

The Yergin report also sparked a review by the General Accounting Office (GAO) in 1996 of program benefits claimed in an appendix to that report (DOE's Success Stories).

In varying degrees, all four of these reviews sought to understand the benefits of federal investment in energy efficiency R&D, relative to the costs of those investments. That is, these panels (or in the case of GAO, auditors), sought to answer the question "Do the public and private benefits (typically the value of energy saved) of this research exceed the cost of Federal R&D expenditures?" The panels also typically address whether the research goals align with true national needs, and what, if any, is the appropriate federal role.This paper will discuss the basic steps, and subtleties and difficulties in deriving estimates of R&D program impacts. We will compare and contrast the alternative methodologies and results used by these panels and the GAO to evaluate the R&D programs of the Office of Building Technologies, State and Community Programs (BTS) within DOE's Office of Energy Efficiency and Renewable Energy (EERE). We will conclude with a succinct overview of lessons learned, to help prepare others for similar program evaluations.

Brief Overview of DOE's Buildings Applied R&D Program

The Office of Energy Efficiency and Renewable Energy (EERE) has been around in one form or another since the establishment of the U.S. Department of Energy (DOE) in 1977. In fact, some aspects of the buildings research program predate DOE; as such efforts existed in the Energy Research and Development Agency (ERDA), a predecessor to DOE. EERE has generally been organized around end-use markets, like buildings, transportation, industry, and utilities (renewable energy), with special emphasis in areas like low-income weatherization, Federal energy use, and assistance to States.

The Office of Building Technologies, State and Community Programs (BTS) has overseen a broad program to address energy use in the building sector. This has included, and still includes research and development (R&D), regulation, and market deployment and transformation activities. For over 25 years, and despite several major policy shifts as Administrations changed, the R&D program has maintained a lead role in advancing key building technologies. Several of these technologies have entered the marketplace, assisted in no small part by the market deployment and transformation, and regulatory activities of DOE and other Federal and State agencies.

The R&D program focused on the major end-uses in buildings; particularly where large opportunities for energy savings were apparent. These especially included:

- Heating and Cooling (advancing both envelope and equipment technologies)
- Lighting (light sources, luminaries, controls, daylighting)
- Appliances (refrigeration, etc.)
- System or "Whole Building" Approaches (design, construction, operation and maintenance)

The program emphasis shifted over the years, but at various times included one or more of the following principal strategies:

- Long-range, high-risk basic or enabling research with National Laboratories, small business R&D firms and universities.
- Applied research and development of systems, software and methods for improving the design of building components, subsystems and whole buildings.
- Development of new products in collaboration with industry and/or small business firms through competitive solicitations.

That is, the program has focused on all three phases of the classic R&D model; basic research, applied research, and product development.

The DOE annual budgets for the buildings R&D program are shown in Figure 1. DOE has invested a total of about $1.3B in buildings R&D since the program's inception in 1976. Of this, roughly 75% has been for applied R&D or technology development. The balance is split for between more basic research, at one end of the spectrum, and technical feasibility and commercial viability studies, at the other.

Although the buildings R&D budget was by far the smallest among the EERE programs, it has generally been recognized, by internal and external reviews, as producing a disproportionate share of DOE's overall R&D accomplishments, in terms of products purchased in the market and resultant energy saved.[1] The most significant, or at least the most often cited, technologies are:

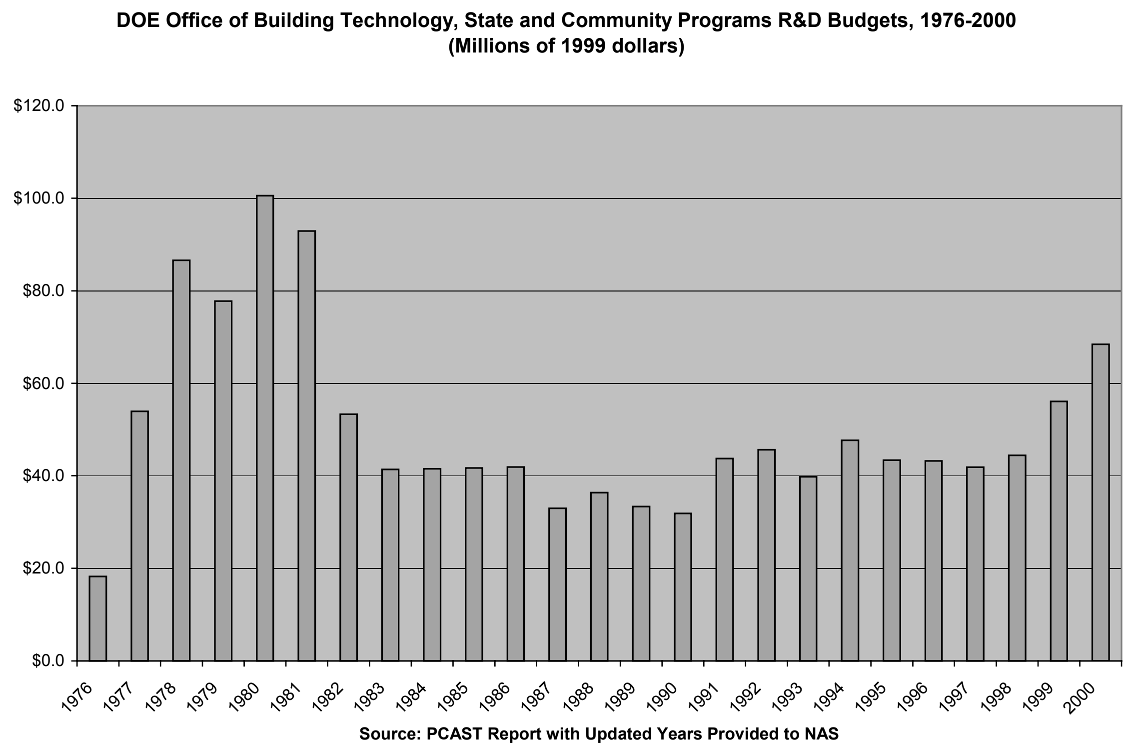
- Materials for condensing gas furnaces
- Electronic ballasts for fluorescent lights
- Efficient refrigerator compressors
- Low-emissivity (Low-E) coatings for windows

[1] In the most recent evaluation, the NRC found that "[b]y an order of magnitude, the largest apparent benefits were realized as avoided energy costs in the buildings sector in energy efficiency…", Energy Research at DOE: Was It Worth It? Energy Efficiency and Fossil Energy Research 1978 to 2000, National Research Council, Pg. 6.

- DOE-2 (now Energy Plus) building design energy analysis software
- Flame retention head oil burner
- Modulating supermarket refrigeration system.

Most of the reviews of the building R&D program have included some combination of these technologies as part of the evaluation.

Figure 1. DOE/BTS R&D Budgets, 1976-2000



History and Overview of "Blue Ribbon" Reviews of DOE's Applied Research Portfolio

Over the last decade, DOE's applied research portfolio has frequently been evaluated by external organizations or by "Blue Ribbon" panels of experts. The buildings R&D program; as part of DOE's extensive applied R&D portfolio has been included in these high-level, and highly visible reviews. Most recently, these have included

- National Research Council (NRC) in 2001
- President's Committee of Advisors on Science and Technology (PCAST) in 1997
- Secretary of Energy's Advisory Board, Energy Task Force on Applied R&D chaired by Daniel Yergin (Yergin) in 1995

In addition, the General Accounting Office (GAO) later reviewed the claimed benefits from energy efficiency research contained in an appendix to the Yergin report (DOE's Success Stories) in 1996.

These reviews usually originate in two different ways, which also tends to impact the nature and conduct of the review:

1. The Executive (Administration) branch or DOE directly tasks an expert panel to review DOE's portfolio and to make recommendations about the content and direction of that portfolio (Yergin and PCAST).
2. The Legislative (Congress) branch of government tasks an organization to review DOE's portfolio, and in essence to audit its claims of program impacts (NRC and GAO).

Administration-initiated reviews occur every two to three years.[2] While these types of reviews have somewhat different emphases, the general essence is well captured in the charge to the Yergin Task Force:

> *"Provide a top level review of our civilian energy R&D programs…assess the DOE's current R&D portfolio against our strategic goals…and national needs…and provide recommendations for changes in our programs and plans."*

The questions posed to DOE by the Yergin Task Force also illustrate the nature of this kind of review:

- What are the specific technical goals of the R&D program?
- If the goals were met, how would the results contribute to larger strategic goals?
- What successes have you had?
- What benefits have been derived?
- How do the benefits compare with the costs?
- What is your overall "bottom line" as an R&D enterprise?

While these types of reviews, in general, include analysis, or at least information on the benefits and costs of these programs, they also provide recommendations for future research. When included, this cost and benefit information is usually generated by DOE and then accepted into the report. For instance, the Yergin report contained a DOE submission entitled "DOE Success Stories" that provided DOE-generated estimates of consumer energy bill savings; e.g.:

> *"Use of the DOE-2 software accounts for $1.9 billion in energy savings for buildings constructed through 1993."*

Claims such as this bring us to our second kind of review; Congressionally initiated evaluations or audits of claimed program benefits. The goal of this type of review is more finely tuned to answer the question of whether the public expenditures on R&D programs are worth the costs. With respect to the *Success Stories* in the Yergin report, the Chairman of the House Committee on the Budget asked the General Accounting Office (GAO):

[2] Most recently (March 2002), the Office of Energy Efficiency and Renewable completed and submitted to the Secretary of Energy a Strategic Program Review (SPR) of the EERE research and development programs. This was in response to a recommendation in the DOE National Energy Plan (2001).

"…to determine whether the claims DOE makes in the (Yergin's Success Stories) report are valid."

A more recent example of this kind of review is the National Research Council (NRC) report "Energy Research at DOE: Was It Worth It?" In this case, a House Appropriations subcommittee directed the NRC to evaluate the benefits from energy R&D programs including those in EERE and consequently buildings research. The goal of the review, as revealed by the title, was to give a sense to appropriators whether past R&D had yielded benefits, as a guide to future investments.

To answer the "Was It Worth It?" question, the NRC applied a detailed and novel methodology framework (developed specifically for the study) to assess the benefits associated with the DOE program. The NRC also a case study approach, as did the GAO audit. The authors of this paper, following NRC guidelines, developed the initial benefits estimates for a number of buildings R&D successes (the NRC also asked DOE to identify failures). The benefits presented in these case studies were in turn evaluated and critiqued by the NRC, and formed the basis of the conclusions in the final report.

Here is an important distinction between these two types of evaluations (Administrative reviews versus Congressional audits). The objective of Administrative reviews is usually to evaluate program goals and performance against those goals, and to chart future direction for programs. The objective of Congressional audits is much more narrow focusing just on performance, as measured by the benefits (program outcomes) less the cost of producing those benefits (both government R&D and private sector expenditures).

All Audits Are Not Created Equal

In essence, the GAO and NRC both audited the benefits of EERE programs, albeit using different methodologies. On the surface, estimating the benefits of these programs is accomplished by simply adding up the energy savings (BTUs/Unit times the number of unit), multiplying by the value of those savings ($/BTU), and subtracting out the R&D and incremental cost to the consumer (price differential) to purchase the enhanced product. However, each step in this process requires a series of assumptions, simplifications, and other adjustments due to lack of data or dealing with a host of other unknowns. The willingness of auditors to accept, reject, or modify these assumptions can greatly impact the results.

For illustration, the energy savings of a new window technology depend on how many and where the windows are installed (geographically and in the building itself). These savings have to be compared to some baseline that could be the typical (average) stock, or the next best technology. But the baseline could also be impacted by behavior (such as drawing curtains). Next one would need to know the energy used to compensate for window losses (heating loss in winter) or gains (increase in summer cooling requirements) and the cost of that energy. The next value in the equation is the increase in consumer costs, which is not the same thing as product price as producers can either mark up these products by a higher, lower, or the same margin as the "baseline" technology. Finally, there is the question as to how much credit is due the government program. That is, what share of the net savings is attributable to government efforts (some, all, half?) and the answer to the unknowable question as to whether the private sector would have, eventually, developed the technology absent the government program?

The results of the GAO and NRC audits, and conclusions drawn, did not uniformly agree. This difference in results occured though the technologies, and even source data reviewed were similar. This is a powerful indication of the challenges inherent in the complex process of deriving national scale benefits from case studies.

Table 1 below summarizes the findings for buildings technologies evaluated by both the GAO and NRC. The GAO summary "findings" were very terse and mostly appeared in a summary table sent to Congress. Since the GAO report lacked sufficient detail, it is difficult to determine precisely how they arrived at their findings. Conversely, the NRC's findings are discussed in substantial detail allowing for a more in-depth look at the thought process. As a result, the comparison that follows provides only a summary sense of the differences in the findings.

A look at the findings and evaluation of the BTS DOE-2 model illustrates how much these evaluations may diverge. While overall the GAO "found problems with the analyses DOE used to support the benefits cited in 11 out of the 15 cases we reviewed," one of the only exceptions was DOE-2. Here the GAO completely accepted DOE's estimate, data, and underlying methodology without further comment. The NRC, by contrast, had several significant problems with DOE's underlying assumptions regarding DOE-2.

Table 1. GAO and NRC Audit Findings Relative to DOE Energy Cost Savings Estimates

	***GAO* Findings**	***NRC* Findings**
DOE-2 Energy Analysis Software	"Adequate Support for all benefits claimed" ($1.9B)	"Substantial but indeterminate" benefits.
Advanced refrigerator compressors	"Adequate Support for all benefits claimed"	Validated. ($9.0B was DOE estimate, $7.0B was NRC)
Low-Emissivity Windows	"Economic Analysis Problem"	Validated but reduced. ($23 billion for DOE, $8 billion for NRC)
Electronic Ballasts for Fluorescent Lamps	"Economic Analysis Problem"	Validated. ($13.0B was DOE estimate, $15.0B for NRC)

When two august bodies, such as the General Accounting Office and the National Research Council, arrive at very different conclusions about a technology's impact on energy use in the marketplace, we have dramatic evidence that calculating of benefits may not be a straightforward or simple calculation. The NRC itself stated as much in the final report:

> *"In theory, evaluating the benefits and costs of DOE's research program should be relatively straightforward. It would require the adding up of the total benefits and costs of research conducted since 1978, determining what proportion of each is attributable to DOE funding, and calculating the difference between the DOE contributions and the cost of achieving them. In practice, of course, methodological challenges abound."*[3]

So, methodological issues are unavoidable in estimating the benefits and costs of the R&D programs.

[3] NRC, pg. 3

Methodological Issues in Estimating Actual Energy & Cost Savings

This section describes many of the methodological challenges inherent in estimating energy savings impacts. These challenges are highlighted and illustrated based on our experience responding to the NRC, GAO and other Blue Ribbon panels. However, the discussion below goes well beyond the first order questions posed by these review panels to reveal further nuances and subtleties in producing such estimates.

We focus first on *energy* savings, and then turn to economic (cost) savings, which pose additional challenges, next attribution (who gets credit?), and finally conclude with a discussion of "additional benefits" which are still more difficult to quantify, let alone economically valuate.

Energy Savings

Conceptually, the methodology for determining energy savings consists of three simple steps:

- Determine the energy savings of a single unit of the new technology versus the technology being replaced
- Multiply by the number of units sold each year
- Add the annual savings over an appropriate period

This assumes of course no change in behavior (take back effect) as consumers realize that they can increase services (having a warmer house), while still realizing som savings due to increased efficiencies of a the new technology (e.g., furnace).

Since the goal of these evaluations is to document historical savings, the methodology appears to avoid any forecasting issues and simply relies on historical data on sales and efficiency levels. Hopefully the data are obtainable. Gains in efficiency can be verified (though actual performance may vary) while units sold may be available from government or industry sources. The "replaced" technology may be a bit contentious, but at least the assumption can be transparent. Then, one has only to run the numbers.

While the basic approach is indeed simple, executing the methodology in practice raises a number of operational questions. In order to highlight these issues, we break down the estimation process into several steps. Each of these steps is discussed in detail below.

Determine the Energy Savings of a Single Unit of the New Technology

Technology baseline. The first question that arises is: energy savings relative to what baseline? That is, what technology is being replaced? There are basically three options for establishing the baseline:

- <u>Average of existing stock</u> (acknowledging the fact that specific units replaced spanned a range of vintages, in both new and retrofit situations)
- <u>Typical new installation</u> for new or replacement units (assuming current market trends and technology choices. Standards of course also play a role in defining the baseline.)

- Next best available alternative technology (reflecting the viewpoint that energy-conscious buyers would have bought the best available technology)

When considering hardware these options are fairly easy to understand. For example, 1200 kWh might be the stock efficiency for refrigerators. A typical new installation would be 650 kWh, though upcoming standards reduce that further. An Energy Star refrigerator with still lower consumption would be the next best available technology.

However, the choice of options becomes more perplexing when considering technologies such as software or other enabling technologies that do not save energy directly, but have the "direct" effect of changing how people *think* about their energy technology choices. The question then becomes: what logic chain can be validated to establish a ***credible and defensible*** link with actual energy savings?

For example, the NRC challenged the DOE's assumption that the DOE-2 program was twice as effective as the next best simulation program. In the view of the NRC panel, the incremental value of DOE-2 as compared with the next best technology needed to be documented with a statistically valid survey of practicing design engineers and architects (GAO accepted the less formal survey data available from DOE). The NRC also noted the lack of a necessary causality between using DOE-2 to analyze alternative building designs, and the actual construction of a more efficient building. For the NRC, these questions weakened the case for incremental energy savings attributable to DOE-2 and, in contrast to the GAO acceptance of these assumptions, caused the NRC to list the energy savings as "probably substantial" but unquantifiable.[4]

Comparability of unit efficiency data. The next question is what is the relative efficiency gain versus the baseline technology. The nuances behind the answer to this question are often ignored. Is the energy efficiency data for both the conventional and the new technology equally reflective of what happens in actual practice? Are test procedures reasonably accurate? Do new appliances have added features (e.g., electronic controls with standby losses) or usage patterns which are not reflected in published efficiency data? Are the more complex energy-efficient technologies more likely to be installed improperly or do they have degradation modes and rates different than conventional technologies?

To add a real world example to these hypothetical questions, high efficiency air-conditioners and heat pumps with capacity control have longer running times because they operate at lower capacity most of the time, compared to single capacity units. While the unit itself has higher efficiency when operating at low capacity, the longer running time results in much greater distribution (duct) losses in many installations. The added duct losses almost certainly negate some of the increase in unit efficiency.

The NRC in particular voiced concern over degradation of performance in its analysis of DOE's calculations for low-emissivity windows. Specifically the NRC stated: "the windows may not last 30 years in the field. If they lasted on average only 20 years, the net life-cycle cost savings would be reduced to $5.1 billion, about $3 billion less." As can be seen from this calculation, the assumptions underlying the answer to this question can have real impacts.[5]

---

[4] NRC, pg. 103
[5] NRC, pg. 117

Another aspect of actual vs. theoretical performance of technologies is the potential for users to "take back" some of the energy savings. As an example, the NRC noted: "one might speculate that reducing the thermal conductance of windows encourages people to design more glass into new residences than they had before low-e glass."[6] Such behavioral changes are very difficult to quantify and remain an inherent weakness for analyzing some technologies.

Another aspect of the unit efficiency question is whether the analytical support for energy savings for the new technology is sufficiently robust. In the case of building envelope or HVAC technologies, the potential for significant variance in unit energy efficiency and heating and cooling loads due to climatic variations imposes very complex analytical requirements on the analysis. Comparing seasonal energy efficiency ratings (SEER values) or laboratory test data is probably not sufficient. Obtaining the necessary regional efficiency and sales data can be a significant challenge.

Finally, the baseline should almost certainly not be held constant. The "baseline" technology would almost certainly be evolving except in the cases of very mature and relatively simple devices (e.g., windows). As a result, baseline energy use may be going up (added features) or down (increased efficiency). This can be handled to some degree by limiting the period of which benefits are claimed (see Sections 3 and Section B-2 below).

Site vs. source. How one chooses to consider energy also impacts the benefits analysis. In the opinion of the authors, comparing energy savings based on source energy consumption (allowing for consumption of fossil fuels to create electricity) is highly preferable. This approach tends to put electric and gas technologies on an equal footing, plus it reflects the ultimate energy resources at which the programs are targeted.

Unfortunately, electricity supply factors can severely complicate the analysis. The most difficult issue is determining the marginal fuel and heat rate of the electricity actually being saved by more efficient technology. Most analyses to date have not addressed this issue and have simply used annual national average heat rates or heat rates reflecting the newest generating plants (yet alone the issue of what plant is actually operating during the time of day or season of operation of the technology). In comparing a new energy technology against a replacement, this issue only changes the absolute savings. In evaluating an entire program with multiple technologies and energy sources, this site-source issue can change the relative impact of technology programs, and hence impact portfolio choice.

Multiply by the Number of Units Sold Each Year

Data availability. Data quality and appropriateness for the analysis being conducted raises another series of questions. Are statistically valid data on shipments of conventional and new technology publicly available? Are the data available with the required differentiation by product type, region, efficiency level, new vs. retrofit application, etc? Have assumptions related to the data been validated and documented?

As an example, regional data are available for heat pumps sales. However, the regions are rather large geographical areas covering different climate zones. There is also little information (data) on whether the products are used for retrofit or new construction. As

[6] Ibid.

such, assumptions are required about the likely climates, and hence loads served by these heat pumps. Models of building types and geographic locations can help to resolve this issue, but this introduces a new complexity; the evaluation of models, and modeled data, as opposed to sample or census data.

Market data for enabling technologies. If a new DOE technology enables energy savings (like software), units sold may not be an accurate measure of impact. What information exists to show how widely it is used and for what types of projects? If the enabling technology leads the building designer or retrofitter to use better components, how is double counting avoided? That is, how is credit apportioned between the enabling technology and the new DOE component technology?

Again in the case of the DOE-2, the NRC voiced considerable skepticism over DOE claims. In this instance, the NRC panel honed in on the "market data" proxy of "units sold" of DOE-2. NRC questioned the extent to which building designers actually use DOE-2 to design buildings, even if they have the program available to them. The NRC independently interviewed three major building design engineers to shed light on this question.[7] While three engineers is likewise less that a statistically valid sample against which to draw conclusions, the NRC highlighted the potential weakness in underlying data and assumptions.

Add the Annual Savings over an Appropriate Period

The key word in this step is "appropriate." Due to the lifetimes of the products being considered (e.g., refrigerators and boilers), DOE developed technologies can potentially impact the market for a period of 20 or more years. The baseline technology and the DOE-developed technology will certainly evolve considerably during this period. Incremental energy and cost performance can be expected, especially with the newer technology. But perhaps the most important consideration is "at what point in the time horizon would the technology have emerged in the absence of the DOE program?" This issue is discussed in more in "Attribution to DOE."

Cost Savings

Incremental cost. What is incremental cost? It would seem that the incremental addition to first cost over the baseline technology would be quite obvious. After all, prices should be readily available from equipment catalogues or retail sales outlets. However, several factors add complexity even in this seemingly straightforward situation.

Is incremental cost the price of the product (which may include varying markups) or the additional cost of production? Does incremental cost include installation costs and training associated with new products, or is it just the price of the product? Should the incremental cost be allowed to vary over the analysis period, as it almost certainly would as the technology penetrates the marketplace, becomes better understood, and manufactures find better, and cheaper ways to produce the product?

Pricing of new products often includes higher markups. Manufacturers will also try to differentiate markets and early technology adopters usually pay a price premium as

[7] Ibid, pg 101.

compared to standard "commodity" products. New products may also include bundling of additional, non-efficiency features as compared to standard products.

Another factor is installation. With newer technologies, installation may be more complex and costly. Installers may increase prices to cover real or imagined risks of unfamiliar technology. One would expect that these price premiums (for early adopters, and uncertainty) would decrease over time as the technology becomes more commonplace.

Finally, some technologies are embedded as a component in other products. As such, catalog prices may not be an accurate reflection of the incremental difference of the efficiency improvements. Examples include low-emissivity glass that is embedded in window products and improved appliance insulation that is integral with the refrigerator structure. For enabling technologies such as DOE-2, the purchase cost is often minor and the major incremental cost is in learning and using the technology, which is may be difficult to document.

Attribution to DOE

The methodology discussed above provides quantitative energy and cost savings estimates for implementation of new technology. The methodology does <u>not</u> tell us whether, or to what extent, these savings are the result of DOE actions as distinct from the private sector role in brining the technology to the marketplace. This is a thorny question as the number will be greater than zero, but less than 100%, and the ability to argue one number as opposed to another can depend upon a number of factors.

Review panels have focused considerable attention on the causality between technology market performance and research and innovation funded by DOE programs. GAO for instance requested that BTS revise its benefit estimate for the Flame Retention Head Oil Burner (oil furnaces). GAO argued, and DOE agreed, that for this particular technology, the private sector would probably have adopted or developed the technology within 5 years. As a result, DOE only takes credit for the first 5 years of sales of this product. Energy savings continue to accrue, though, as products purchased and installed in this first five years have operation lifetimes that extend well beyond that period.

Counterfactual History

The ultimate question boils down to: "how much of the energy and cost savings would have resulted in the absence of any DOE effort whatsoever? Upon close inspection of the historical record, nearly all unbiased observers would conclude that DOE action had <u>some</u> impact on the market introduction and market penetration of the new technologies. Determining <u>how much</u> impact requires that the analyst construct a non-DOE hypothetical scenario. The non-DOE scenario has to answer several very difficult questions:

- When would the technology have been developed without DOE R&D support?
- When would competitive market pressures warrant introduction of the technology?
- What level of product support would have been provided to nurture the new technology?

Building the non-DOE scenario has to account for the rate of technical development and market strategies internal to private firms as well as the market's demand for new technology and the competitive environment. The scenario building also has to factor out the inherent imprimatur that accompanies a technology developed by DOE, including DOE's performance validation and confidence-building efforts that have influenced the historical record.

This question, of what would have happened if DOE had not conducted the research, of course, like all counterfactual arguments, impossible to answer; one can only speculate. In lieu of speculation, there are alternate methods of limiting this number.

Dollar contribution. One approach is to use the government share of the total R&D as a simple multiplier. For instance, if the government contributes half of the cost of R&D, then it would get one-half the credit for energy savings. This simple rule of thumb is appealing in it's simplicity, but is fraught with problems.

For instance, government funding might "crowd out" (substitute for) private spending. That is, the private sector would have been willing to spend money for this research, but is more than happy that the government is willing covering the cost. In this situation, the benefits should actually be higher, as the government expenditures should have been zero, or at least less than they were. But, perhaps there should be no government credit at all since the private sector would have done it anyway?

At the other extreme, a relatively small contribution can come at a critical point in time or in the R&D. This might be the case where a technology is being brought from one application and market to another. Another situation might be where government funding solves a critical issue that benefits more than one technology. Therefore, the dollar, or dollar equivalent contribution may understate the relative contribution of government.

Setting aside R&D, there are many more steps required to get a product into the marketplace. Product testing and evaluation (in addition to demonstration) will have to take place to increase confidence in the product. There will be scale up costs to bring the technology into production. Finally, advertising and other market conditioning activities will be necessary to increase consumer awareness or create the market for the product. The cost of the actual R&D (and the government share thereof) might be a very small fraction of the total product cycle development. Therefore, take half the credit for energy savings for a product, based on half the cost of the R&D, might once again seriously overstate the contribution.

Finally, outside the Federal government, there may be State and local activities in R&D as well as deployment. Other Federal agencies may have a role in developing the technology. And, there may be international firms and governments pursuing like technologies. So, the amount of funding provided by DOE and the private sector partner may understate the total investment to bring the technology to market.

The NRC 5-year rule. The NRC adopted an approach to the non-DOE scenario that put a premium on transparency and consistency across the various technology case studies it reviewed. In each case study, the NRC assumed that, in the absence of DOE, every technology would have been introduced to the market 5 years after the actual introduction date and the historical rate of growth of market penetration would have been replicated exactly 5 years later.

The disadvantage to this "one size fits all" approach is that it ignores the very significant differences among the technologies. For instance, the government intervention (R&D) may take place very earlier in a very long product development cycle. A five-year cutoff would not recognize that the product might still have several years of (private and/or public investment). As another example, the government R&D might increase the eventual level of efficiency and, in the case of related government programs, the level of adoption.

However, the NRC did not claim this approach resulted in a credible alternative history. In fact, the five-year rule has no basis of support in the literature. This rule was simply created in establish a "…uniform, conservative standard for the analysis…"[8] The approach also appears methodologically "cleaner" because on the surface it has few subjective assumptions.

The x-year rule. A more reasonable approach, though admittedly more subjective, is to modify the NRC approach for each case. This allows the analyst to take into consideration the overall length of the product development cycle and the point at which government R&D was conducted. It could be reasonably argued that the private sector, which has a shorter time horizon, would take longer to "do it anyway" as one moves earlier into the product development cycle. The difficulty in this approach is that it requires many more subjective assumptions that can be challenged.

The shape of the market penetration curve. Market penetration of any product is often described as an "S" curve where time is plotted along the horizontal axis, and units sold, market share, or dollar volume is measured up the vertical axis. Typically, unit sales start relatively low, but as market adoption increases, sales accelerate sharply, and then level off as the market saturates.

The NRC assumed that the rate of market adoption, and eventual efficiency level, would be the same with or without DOE. However, DOE programs consist of a number of activities that support the marketing of a new technology. In some cases, these activities have been extremely effective in changing the shape of the penetration curve.

The flame retention head oil burner illustrates this point very clearly. DOE did not develop this technology. Instead, DOE's contribution was entirely downstream of the development process. The flame retention head burner was a very low cost advancement that had been available as an option to oil furnace technicians for at least 10 years, yet it had virtually no acceptance in the marketplace due to a number of technical and institutional barriers. This remained true despite dramatic increases in fuel costs of heating oil that one would expect to move this cost-effective technology into the market.

DOE initiated a program to address all of the critical barriers preventing the use of this technology. Specifically, DOE's program carefully assessed safety and reliability issues through validated laboratory and field tests. The program also aggressively promoted the technology with both consumers and oil heater manufacturers and dealers. It also created a knowledgeable infrastructure by training technicians.

Taken together, these efforts appear to have been critical in causing the quick adoption of this technology. The flame retention head oil burner captured 100% of the new burner market within 5 years after completion of the $7 million DOE effort in 1981. In this case, the curve is steeper, <u>and earlier</u>, than would have been the case without DOE.

[8] NRC, pg. 19.

Nevertheless, the GAO was highly skeptical of DOE claims of causality for a technology that already existed prior to DOE's involvement:

> *"DOE cites energy savings of $5 billion attributable to the use of the flame retention head oil burner technology over the past 15 years, although DOE's contribution was limited to testing and publicizing the technology. ....We do not believe that the total savings resulting from consumers' use of this technology should be linked directly to DOE's activities."*[9]

Another general argument would be that the eventual level of efficiency is higher than would have been without DOE intervention. Once again, it can be reasonably argued that, as is its mission, that DOE would focus more resources and attention at getting the maximum level of efficiency achievable while still being cost effective. A private firm, on the other hand, might end the research earlier (good enough) or focus R&D elsewhere to optimize some other product attribute. Once again, though, this requires additional subjective assumptions that, though intuitively reasonable, are difficult to justify to skeptical audiences.

Attribution of related impacts. Some DOE technologies have had significant impacts in related areas. For example, there is a logical, conceptual linkage between the availability of more efficient refrigerator compressors (result of DOE research) and the efficiency levels established by the DOE appliance standards. There is a clear linkage in the use of the DOE-2 energy analysis program, which has been used extensively to develop new building performance standards, such as the ASHRAE 90 series. However, the alternative path in the absence of DOE-2 is unclear. Even less clear is the actual reduction in energy use that has resulted from the ASHRAE standards. But, a robust analysis of the total benefits of a technology should consider these spin-offs.

Other Benefits

One clear contribution by the NRC is the consideration of a broader spectrum of benefits. Typically, the benefits are stated in terms of observed energy savings and the dollar impact in the reduction of pollutants. NRC took this further to note that R&D efforts include the creation of "Options" and "Knowledge" benefits.

In the case of options benefits, the NRC formally recognized that some R&D does not result in immediate product development. Instead, these technologies are shelved as options should the situation arise that energy prices, or some other driver (like climate change) make these options feasible. The dollar valuation of these options is not possible without also postulating alternative scenarios of the future.

In the case of knowledge benefits, NRC also recognized that R&D "failures" also provide knowledge. At a minimum, a failure reveals a dead-end so that future research down that line of investigation is avoided. But, while the research may be a dead-end for the application to the particular market under consideration by the government program, it may have applications in other markets. The NRC recognized both of these.

---

[9] GAO, pg. 5

Lessons Learned and Summary Observations

The DOE R&D program, and particularly the Energy Efficiency program, has been the subject of numerous external reviews over the last several years. These reviews have included attempts to quantify the impacts of these programs in terms of a cost benefit analysis. Using several recent reviews, the authors have attempted to show that estimating the benefits of R&D programs, while seemingly simple to calculate, are difficult to develop in practice.

Nuances involved in developing benefits estimates can be overwhelming. Data limitations and complexities in analysis require simplifying assumptions to make the analysis tractable. However, with each successive twist and turn, assumption and aggregation, the results become more and more subject to criticism.

Examples were shown where external reviewers came to differing estimates, and conclusions about BTS program impacts. These differences arose even when looking at the same technology and supporting data. This variation arose because of challenges to data, difference in underlying assumptions, and application of different methodologies.

The paper discussed the subtleties at each point in the calculation process when these estimates are being developed. Simplifying approaches were discussed, including those developed by the National Academy of Sciences. The benefits and potential pitfalls of using such simplification schemes were highlighted.

The authors have participated in several of these reviews by performing background analyses and to requests for information from review panels. The authors recommend that program managers calculate and document anticipated and actual benefits as an integral part of the program, not retrospectively.

These internal and external "Blue Ribbon" panel evaluations are cyclical and it is anticipated that such requests for retrospective benefits estimates will arise again and again. R&D is a long-term process, but the collection of data and evaluation of program impacts does not have to wait until program completion. To attempt to reconstruct this story line and data years after the fact is both difficult and expensive.

References

National Research Council (NRC). 2001. Energy Research at DOE: Was It Worth It? Energy Efficiency and Fossil Energy Research 1978 to 2000. Washington, D.C.: National Academy Press.

General Accounting Office (GAO). 1996. DOE's Success Stories Report. Washington, D.C. GAO/RCED-96-120R.

President's Committee of Advisors on Science and Technology (PCAST). 1997. Federal Energy Research and Development for the Twenty-First Century. Washington, D.C.: Report of the Energy Research and Development Panel.

Secretary of Energy Advisory Board, U.S. Department of Energy (Yergin). June 1995. Energy R&D: Shaping Our Nation's Future in a Competitive World. Washington, D.C.