Conference on
Estimating the Benefits of Government-Sponsored Energy R&D[1]
March 4 and 5, 2002

SUMMARY OF DISCUSSION IN WORKSHOP A: PROSPECTIVE BENEFITS – CROSS-CUTTING ISSUES[2]

James C. Wolf, Independent Consultant
May 1, 2002

Framework for Defining the Benefits of R&D Programs, and its Use for GPRA and R&D Planning and Evaluation

This session focused on discussion about the National Research Council (NRC) matrix that had been presented in the plenary session that morning. Overall, many workshop participants regarded the matrix as being very helpful and useful, but some concerns were expressed about it.

The concerns are reflected in the following questions and comments made by some of the workshop participants:

- Isn't there still a knowledge benefit even if it is not "realized?" What of training, learning by doing, and other factors that may not be considered a specific "knowledge" benefit in technical reference terms?

- How does the matrix for an individual program reflect the overall portfolio? How do you get from a program assessment to a portfolio assessment?

- Is the matrix too complex – does it help a decision maker with marginal decisions? Should there be weightings or screens reflected on the matrix to reflect overall goals of the department? Or mandates or priorities? Whether Congressional, Presidential or others.

- There is a need to separate out prospective benefits of the new funding being proposed from the prospective benefits that will be achieved from past work already completed that is part of an ongoing program – both will have expected future benefits.

[1] Organized by Oak Ridge National Laboratory and sponsored by the Office of Energy Efficiency and Renewable Energy, Office of Fossil Energy, Office Nuclear Energy, Science and Technology, and Office of Science of the U.S. Department of Energy. Information about the conference is available on the conference web site, www.esd.ornl.gov/benefits_conference; in the white paper distributed prior to the conference, "Ideas on a Framework and Methods for Estimating the Benefits of Government-Sponsored Energy R&D;" and in the report summarizing the conference proceedings, "Synthesis of Conference Discussions."

[2] This document is believed to be a reasonably accurate summary of discussions in Workshop A of the conference on "Estimating the Benefits of Government-Sponsored Energy R&D;" but the accuracy is not guaranteed by the workshop rapporteur, Oak Ridge National Laboratory, UT-Battelle LLC, or the U.S. Department of Energy. Furthermore, the opinions expressed by those at the conference are their own and therefore nothing in the reporting of the discussions in Workshop A or of the conference proceedings should be construed as government policy.

- Are the expected future benefits probability-discounted? Or is the approach assuming that the project will be successful? It was suggested a column be added to reflect “technical risk” – the NRC didn’t have to do this since it was retrospective only – the NRC knew what had happened. Risk could be assigned on a 1 to 4 basis ranking- 4 being the highest technical risk. It was noted that a process needs to be established so that all projects are not ranked the same – no scores of all 4’s, Olympic 10's or whatever. Differences among programs need to be noted or the scores are of no use to policymakers – despite the understandable tendency of program managers to resist rankings.

- There is a need for a column or indication on the matrix to reflect timing – is the project expected to be complete in the near-, mid-, or long term? Should the timing issue be when the project is complete or when significant benefits are realized? This question was not resolved.

- Projects being considered are at very different stages – some just beginning while others many years in development. How can the matrix reflect different stages of a program’s development so “apples are being compared to apples and not oranges”? Shouldn’t projects in the same stages of development be compared to each other and not to others? This question was not resolved.

Defining and Calculating the Projected Baseline on which New Technologies are to be Introduced

Formal presentations were made by Frances Wood, James Turnure and Robert Hirsch. Their presentations are on the conference web site.

There was extensive discussion about the proper definition of "baseline." Discussions centered about differences between evolutionary, revolutionary, most likely and business as usual cases.

There was a discussion about the concerns of using models to project a baseline or to estimate the effects of a government R&D program. The concerns included the suggestion that the effects of an individual, relatively small R&D program are not reflected in large integrated or maco-models. Another concern was how the models deal with ongoing R&D project interactions. If EERE and FE both have projects that are targeting the same market segment, how do the models reflect the potential competition? Questions arose about whether the models accounted for any project interactions on the positive side? Do the models consider activities ongoing at the state level, whether regulatory or on R&D?

The need to make all the *assumptions transparent and clear* for whatever model is being considered – as well as scenarios developed – was also expressed. Models could be small micro models, larger macro models, “integrated” models, or even a process of interviews with independent experts.

Many workshop participants agreed that the projected baseline for the purposes of our discussion would:

- be an evolutionary one – with some progress in increasing efficiency – a business as usual case,
- include most likely changes in the market and technology,
- include expectations about fuel prices,
- be based only on existing laws and regulations – no new laws or regulations are assumed.

The baseline was described by some participants as our best “guesstimate” of what is likely to happen. It was discussed that different “models” could be used for different projects – there was not a "one size fits all" solution; but if different models are used, their respective assumptions should be made consistent to the maximum degree possible.

For many workshop participants, the market under consideration for the projected baseline was the relevant one to the technology under consideration – a hypothetical “microwave toaster” was used as an example. Baselines could be developed with interviews of independent, non-involved experts in the field of the R&D project; competitors, NGOs (non-government organizations), and others. This approach was an alternative to the use of mathematical modeling techniques or as a complement to them.

Concerns were expressed that the baseline definition was dealing with only individual projects and not the portfolio; and that the models were less useful for the longer term than for the shorter focused projects – there was simply too much uncertainty in the longer term.

The discussion began to overlap with that proposed for the next session on projecting the effects of a government R&D program on technology deployment and use.

The Energy Information Administration's (EIA's) National Energy Modeling System (NEMS) model and the assumptions underlying the Annual Energy Outlook (AEO) were deemed by many to be the best ones to start with in terms of models used for the baseline. This approach was also favored by many workshop participants for projecting the market penetration of new technologies. Many viewed that a good approach was one starting with NEMS, and then using other models, such as Markel, as needed if NEMS were not suitable. Many workshop participants thought the other models needed to be calibrated to NEMS to assure consistency. Data limitations were raised as a limitation for all these approaches, as was the large cost of too much or too refined an analysis.

The EIA NEMS base case scenario was viewed as the most appropriate to use as a projected baseline. However, this scenario includes projects underway at DOE so one should be careful to net out the R&D program being considered when trying to assess its benefits against the EIA baseline if that project was in the baseline already.

Scenarios

Discussion began about how to evaluate different scenarios. Several different approaches were suggested:

- One could construct different scenarios for different potential policy goals or objectives: e.g., no carbon increase; decrease oil imports by X%; etc.

- Another approach is to test different market assumptions: e.g., dramatic increase in oil prices (outside of normal sensitivity range) etc.

As an alternative to "option" categorization in the matrix, it was suggested that all scenarios should be run and then ascertain the benefit of each technology under each scenario.

There was a discussion about differences between sensitivity analysis and scenario analysis. It was not resolved. Was conducting a sensitivity analysis varying an "external" event – prices, etc. – as opposed to scenarios where something we could control – policies – was varied? Or were sensitivities the case where only one factor is varied, and scenarios the case where several different factors are varied. Whatever term is used, the definitions should be clear.

It was viewed among many workshop participants that *independent peer judgment* needed to be exercised over all assumptions underlying the models, scenarios, and market penetration results. This review would enhance the credibility of the assessments.

The Effect of a Government's R&D Program on Technology Deployment and Use

There were presentations by Ric Cheston, Thomas Pelsoci and Rosalie Ruegg. The presentations are on the conference web site.

The presentations focused on assessing the impact of a government R&D program. It was noted that the effect could be much broader than the 5-year acceleration assumed by the NRC. The impact could include work on the technology starting earlier, or progressing faster, or even increasing the scale and scope of the project and its having wider applications.

The view of many workshop participants was that the impact of the government varied on a case-by case basis, and that an across the board "5-year rule" provided some benefits (easy comparisons), but was probably incorrect in reality. The importance of other factors influencing market penetration was stressed, apart from the R&D – policy, market conditioning activities, and the other considerations.

Factors to consider when evaluating both the impact of the government R&D program as well as the ultimate market penetration of the technology were listed by workshop participants:

- The economics of the technology/product –both in terms of its scale (a $50 technology/product versus a $500 million technology/product); and its comparison to the next best alternative.

- The percentage of government funding in the total R&D. It was cautioned that focusing on this ratio might result in an incentive to increase the amount of government funding to raise the benefits attributable to the government. This situation would be counter to another objective, which might be to increase private investment.

- The product development cycle.

- The nature of the industry – fragmented; highly investing in R&D or little investment; competition (both foreign and domestic); etc.

- The nature of the installed base of the product which the new technology is seeking to replace.

- Non-economic customer attributes affecting demand for the product.

- Whether the market the product was entering was a competitive or highly regulated.

- The likelihood and reasons for early adoption by “early adopters” of technology.

- The nature of the market failure which justified a government role.

- The compatibility of the product with other products – does it enhance any other products?

- The risk of the research – how risky is it?

- The stage of product development.

- The timing of when the benefits are projected to be realized.

- The precise actions of the government. Where is it sponsoring research -- in private industries, or universities, or national laboratories?

If the technology is ten or more years from development and deployment, some of these considerations were viewed by many as much less important – e.g. the reasons for adoption by early adopters.

The discussion proceeded to consider how to allocate benefits to the government program. Some workshop participants noted that private industry is typically extensively involved and provides funding in most cases – ranging from the R&D itself to the commercialization of the product. Since there are 100's if not 1000's of projects in many cases in the relevant offices, many workshop participants thought that this analysis should be done on a program and not a project basis.

Two cases were considered: 1) where the government involvement was necessary but not sufficient to develop the technology, and 2) where the government had an important effect but perhaps was not necessary for the development of the technology. In the first case, many workshop participants thought the government should get more credit for the technology development. In order to allocate benefits, the need to clearly and precisely identify what the government was doing at different project stages, and what the private sector was doing, was stressed.

The NRC, as part of its study, had tried to categorize the government's role in the most important energy technology developments of the past 20 years. The NRC committee retrospectively categorized the government's impact as dominant, influential, minimal or absent.

Looking prospectively, if the government is involved, then "absent" would not be an appropriate estimate of the impact of the government program. It was suggested that another 4th category could be added to make it a 1 to 4 a "forced" choice about the significance of the government's proposed program on the technology development – the category could be "Very Important" and be in between Dominant and Influential.

A qualitative assessment that considers the factors listed previously could then be made, and the prospective significance of the government program could be either:

- Dominant,
- Very important,
- Influential, or
- Minimal.

If needed, different percentages could be assigned to each qualitative rating to allocate numerical benefits to the government program. The percentages assigned should be the same across all programs for the same qualitative rating.

Several participants cautioned that independent panels should be used to make this qualitative assessment, and once again that not all the projects should get the same rating. Judgments and distinctions were important despite the reluctance of program managers to make them.

Overall, five distinct approaches were seen as methods to allocate benefits:

1) Based on percentage of government funding. The cautions noted earlier about this approach were expressed again.

2) Using "traditional" market penetration models and comparing the effects with and without a government program.

3) Rules of thumb. Several were discussed:

    a) The 5-year rule.

    b) Time to deployment of the technology. If the time were 0 to 10 years, then the acceleration or government impact would be set between 4 to 5

years. If the time for deployment were 10 years or more, then the acceleration effect would be greater, say 7 to 10 years. This difference recognized that the longer the time horizon to deployment, the less the private sector would be interested in investing in R&D and developing the project.

c) Peer review – forming independent panels with objective criteria to judge importance. The four qualitative ratings discussed above could be used in this process.

d) Random sampling of projects to obtain an idea of the importance of the government. This approach would involve field interviews and expert advice as well. The process could focus on only “big” projects if deemed appropriate.

4) Government payback - Return to the government. This approach would estimate returns achieved by the project and have the government involved in ways different from the usual grant or contract approach. The involvement could be in the form of loans or even equity investments where the government would be receiving a return on its investment, in addition to the public benefits that are presumably achieved by the technology. It was noted that NYSERDA is in part making such investments.

5) Use whatever methods were productive when successfully evaluating projects that had been completed. Whatever methods were successful for "backcasting" could be used to allocate benefits prospectively.

The interaction of R&D with other government policies – tax credits for deployment, market transformation programs, etc, was stressed.

As with the discussion on market penetration, it was recognized that some assessments only needed an order of magnitude or gross rankings, while others might require more precision. It was important to recognize the difference between these situations and to determine which types of analysis required what degree of detail.

Using Estimates of Benefits in R&D Program Planning and Evaluation

John Mortensen and Helen Kim made presentations, which are on the conference web site.

The importance of interim milestones – and decision making to terminate a project, change direction, or change the benefit assessment if the milestones are not met – was highlighted. Benefit assessment is critical to starting, stopping or changing programs. Changes could include expansion, contraction or a change in program focus or direction.

The importance of portfolio management was also highlighted. The need to cut across DOE stove pipes for portfolio benefit-assessments was discussed. This DOE-wide portfolio assessment was viewed as particularly necessary where different office technologies might in effect be competing in the same market, e.g., for base load capacity.

Several workshop participants considered that the portfolio process must be sustained to be effective. Not only do common methodologies have to be employed but one should also look for program synergies. To evaluate a portfolio, one should involve a variety of stakeholders and interested parties, but not just those who are part of the proposed program. Independent and "disinterested" perspectives and reviews are also needed.

One of the formal presentations included a visual representation of a hypothetical portfolio against various factors such as time frame, program cost, and size of projected benefits. This visual representation of portfolio mapping was highly praised. It was noted that a portfolio could similarly be mapped against other factors as well, such as technical risk, etc.

Several workshop participants also noted the importance of complementary activities to the R&D, such as market transformation activities or even demonstration (if not considered as part of the R&D), was highlighted. The impact of these activities can not be separated or ignored if the evaluation is limited to the "traditional" R&D activities.

Concluding Remarks

The workshop concluded with each participant being asked to present his or her most important "take-away." No consensus was sought. The individual take-aways included:

- Visual representation of the data is very useful
- Whatever is done needs to be Transparent, Relatively Simple, Credible – perhaps peer reviewed by independent parties; Fair to all technologies, program managers and interested parties, and Affordable.
- People are as or more important than whatever process is adopted. More than one approach may need to be used across a portfolio of programs.
- Results are only as good as the data. The data feed the process, which is run by people. All should be vetted.
- The portfolio approach is more important than individual project assessments.
- There is a tension between case-study intensity and top-down generality.
- Don't be too mechanistic. Judgments are important. Just make it apparent when judgment is being exercised.
- Focus on how the information you are seeking will be used. That should determine the process.