Norman Metzger
National Academy of Sciences

Norman Metzger (center)

Last November the National Academy of Sciences (NAS), the National Academy of Engineering, and the Institute of Medicine sponsored a workshop* on the federal role in research and development. In part that workshop dealt with our capacity to measure economic return on federal R&D investments. I'm going to review the conclusions of that workshop, as well as a second document, "Research Funding as an Investment; Can We Measure the Return?," that has just been published by the Congressional Office of Technology Assessment (OTA).

Basically I'm going to focus on what these reports say about our capacity to measure economic returns. I should warn you that I am reviewing other people's materials. I'm not an economist and in no sense an authority in the field. However, having orchestrated the workshop, having read the papers, and recently having

*The workshop report has now been published and is available while supplies last. Write to: Norman Metzger, National Academy of Sciences, 2101 Constitution Ave. N.W., Washington, D.C. 20418.

to review the OTA report, I keep being reminded of that old story about a person coming upon a drunk on his hands and knees obviously searching for something under a streetlight. He asks, "What are you looking for?" "My wallet," the drunk replies. "Is this where you lost it?" he queries. The inebriated gentleman answers: "No, but the light is better here." Looking at some of these methods for economic evaluation, I think we may be looking for the money in the wrong place for the right reasons.

There are four methods for judging the consequences or results of investments in R&D. There are economic methods, which I will discuss briefly. There are science indicators, which use various measures, such as personnel statistics, comparative funding levels, etc., to assess the vitality of a research enterprise. There is bibliometrics, which is a method which has gained some force recently, which uses scientific publications to evaluate the research output. Finally there is peer review. This is easily the dominant method in government and in industry for judging basic research investments and results.

Peer review can be handled in various ways: it can be formal or informal, it can be purely technical, or it can be blended with other perspectives from engineering, marketing, development, and the like. In any case, peer review is the dominant method, especially at the basic research level. In areas closer to applied research and development other methods begin to pop up.

The OTA report points out that economists using the economic method have been able to show quite healthy returns on private R&D investments. According to the OTA (some people dispute this) they have not been able to show comparable returns, and have at times been unable to show *any* returns, on federal R&D expenditures excepting some applied research programs in agriculture, aeronautics, and energy. Why? What are the difficulties?

A principal one, as you all know, is that federal investments in basic research are rarely ever made for economic purposes. They are made principally to serve the public purposes that we understand but can't easily measure. That difficulty saps the reliability of the various available investment models. The reason is that these models embody an assumption that a hypothetical decision-maker can estimate the dollar value benefits of potential investments and gauge the probability of achieving those benefits. As this audience knows, neither of these assumptions is applicable to government-funded research except in special cases. Basic research by its nature creates new, often unexpected understanding which cannot be valued economically. It cannot be evaluated in the usual investment sense, because it is not appropriable. It is available to everyone, and by that token, of value to all, but to no one in particular.

An example: In the early '70s, physicists began to talk about new kinds of semiconductor structures. At the same time, surface science was having a rejuvenation. Concurrently, high-vacuum methods were improving, in part driven by the needs of the high-energy physicists. As a result, people began to create epitaxial techniques for growing thin-layer crystal structures. The dreams of the physicists in the early '70s were realized in actual structures. Fascinating structures were created including quantum well lasers, integration of optical with electronic components monolithically on a single chip, etc. These are very interesting developments, but how would one have valued the work that was going on in the '70s? How would one have estimated what the results and the economic consequences were going to be?

Another difficulty is that the relationship between research and innovation is rarely linear. Most economic models depend on an input and an output, and a production function that links the two. If you put in a production function you are making an assumption about the way the world works. You are making an assumption about how the research system works, and how innovation happens. But, in fact, what I said earlier about the relationship between basic research and innovation is hardly linear. It is long term, indirect, unpredictable.

Innovation also depends on a lot of things other than research. Interest rates, anti-trust policy, patent policy, the quality of the work force, what other countries are doing, on and on. If you still want to attempt to do an economic evaluation, then you have several difficulties ahead of you. Dealing with inputs, you have to decide where does R&D end and a commercial program begin? For example, is the cost of the Manhattan Project part of the R&D leading up to nuclear medicine? What R&D is relevant to a commerical product? How quickly does the R&D start to depreciate? Does one measure gross or net R&D stock? How are the numerous contributors to innovation valued? What weight is assigned to private versus federal R&D? Should one assign different values to federal research done in government laboratories, universities, or companies? Is a government's definition of R&D the one that should be used? Are overhead, training, information dissemination, data collections and the like, which are often included in agency R&D budgets, in fact really R&D expenses?

Those kinds of questions are mirrored when one tries to also think about what the outputs are. For example, how is the value of such non-economic outputs as national security measured? How are intermediate outputs, such as a mathematical theory that may eventually contribute to the development of new products, to be valued? Where do the outputs of a given R&D project end? How are spillovers in an unrelated field identified or evaluated? Is there an acceptable measure of output in service industries? For example, does a count of a physician's work and the number of hospital beds occupied reflect the output of the healthcare industry accurately? Do hours available for work measure the value of health? What does one do with an output such as the maintenance of the scientific enterprise, which involves real costs but can easily be factored into innovation or productivity gains?

The difficulty, again, is that typical federal R&D produces things that have no market value that economists can begin to measure. There is no market price for most health advances or a strong national defense. Overall, then, to quote Peter Reiss of Stanford University, "Current economic measures of returns for federal R&D at most provide crude historical statements about the contributions of that R&D investment." Further, there are no easy shortcuts that can dramatically improve our methodologies. At the NAS workshop, Harvey Brooks of Harvard commented that the simpler the methodology, the less valid the results, and that some studies are little more than propaganda.

Finally, to quote the OTA report, "The metaphor of research funding as an investment, while valid conceptually, does not provide a useful practical guide to improving federal research decision-making or to judging results."

What happens then? Well, we fall back on the anecdotal, on case histories, on retrospective analysis, and the like. By and large they serve reasonably well. I remember Harvey Brooks calling me up last summer when he had been asked to do a paper for the workshop on evaluating the returns on investments in the physical sciences. He said he couldn't find any quantitative data that he trusted or that he thought was reliable. As a result, he fell back on anecdotal case history methodology. Brooks also talked about the advances in superconducting technology, ultra-high vacuum technology, minicomputers - all driven by investments in high-energy physics.

Overall, just to repeat an earlier point, peer review dominates in industry and government. No federal agency has in place a research or evaluation system which appears to have moved substantially beyond the use of informed technical judgement. There are some more recent efforts by the National Institute of Health and other agencies, to try to apply bibliometric techniques to research evaluations. Those are still going on.

Let me add a final caveat which is more personal than anything else. It is based on a comment that Eric Sevareid once made, that the chief cause of problems is solutions. Obviously, the pressures to make an economic case for basic research will remain and probably intensify. I think there are reasonable responses to those questions, and I suspect we will hear some from Steinmueller. But what I think needs to be avoided at all costs is a number. That could be, for example, the per cent contribution to the GNP of high-energy physics, or anything else, or science as a whole. For the reasons I outlined very briefly, any number is at best questionable and open to attack. I would only repeat the caution given by a participant at the NAS workshop, Hendrik Hortzfeld. He was talking about returns on federal investments in space research, but I think his comments apply to other fields. He said no economic study should attempt to put a "bottom ratio" on returns, on space

R&D investment. There is no such number in existence. It lives only in the uncharted world of general equilibrium theory. All such numbers are products of economic models with many limiting assumptions. Even when these assumptions and qualifications have been carefully laid out, the existence of the number is an attractive bait for those who need to justify space R&D. Once a total returns number is used, it quickly finds its way into misuse.

Finally, the use of an economic model for R&D success may skew decision-making so that we may find that policy makers will pay too much attention to the simple economic measure at the expense of the more important but less quantifiable criteria for federal R&D.