OECD Global Science Forum
Workshop on Science of Science Policy:
Developing our Understanding of Public Investments in Science
Helsinki, Finland
July 12, 2006
Keynote Address

John Marburger
Director, Office of Science and Technology Policy
Executive Office of the President

Good morning. I wish to thank the organizers of today's Workshop for inviting me to speak at the outset of your discussions on this important topic. In 2002 when President Bush met for the first time with his Council of Advisors for Science and Technology (PCAST), he questioned whether the large sums the United States invests in research and development are allocated properly. "Are we investing in the right things?" he asked.

We duly formed a panel on "Federal Investment in Science and Technology and its National Benefits" which held hearings, commissioned a study on historical trends, compared U.S. investments with other countries, and produced a report in October 2002 entitled *"Assessing the U.S. R&D Investment."* That report, which you can find on the U.S. Office of Science and Technology Policy website (www.ostp.gov), made three recommendations. The first suggested re-balancing funding among different fields of science. The second proposed a major fellowship program to address perceived technical workforce shortfalls. The third called for increased attention to assessment, monitoring, and analysis of federal R&D investments with respect to national needs, international developments, and the national science and engineering workforce. I imagine the senior science policy advisor of every country represented here today has received a similar request and followed a similar path to similar advice in recent years.

The panel did its work, and well. Eventually it led to a Presidential initiative, the *American Competitiveness Initiative*, announced earlier this year in President Bush's State of the Union message and currently enjoying favorable reception by Congress for funding beginning in fiscal year 2007. This initiative is popular, I think, because it aligns well with conventional wisdom. While I am satisfied that the President's initiative is consistent with the best analysis available, I wish we did not have to rely on conventional wisdom to carry the day. Indeed the third PCAST panel recommendation was an expression of uneasiness about the adequacy of analysis underlying science policy decisions in general, which is what we are here to discuss today. Two months ago I spoke at a similar workshop at the Georgia Institute of Technology in Atlanta, and today I will share with you the somewhat philosophical remarks I made then.

The deepest human question is *how to act?* And indeed the ability to ask the question at all, and not simply to act from instinct, is a defining human quality. Asking it acknowledges the possibility that there may be an answer beyond blind, intuitive reactions to the challenges nature presents to our existence.

Some parts of our experience are obviously ordered, which makes prediction possible – most dramatic are the cycles governed by planetary motions: the rising and setting of the sun and moon and the sequence of the seasons. Primitive people came to rely on these to plan their affairs and gradually effect the profound transition from hunting and gathering to agriculture. Astronomical regularities are dependable because the sun and planets possess a huge inertia which makes them nearly impervious to the stochastic and unpredictable processes that influence events of smaller scale. From a physicist's perspective the real world is a system with an inconceivably large number of degrees of freedom governed by laws that themselves contain the seeds of uncertainty and chaos. Here the ability to predict the future depends entirely on the possibility of distinguishing a small set of phenomena whose link to all the rest is either negligible or easy to describe. The struggle for predictability consists of a search for such sets of phenomena – subsystems of the whole – that are relatively decoupled from all the rest. Practical applications always rely on artifices that describe the effects of messy nature by a few parameters like the physics concepts of temperature, friction, or mean fields. Without these approximations, practical professions like architecture and engineering would be impossible.

As our power of abstract reasoning developed through the centuries, particularly through the tools of mathematics and computation, the complexity and scale of predictable systems has grown. Today we can consider for the first time how to anticipate the effects of disturbances to the intricate machinery of our own bodies. Dare we begin to hope that we can anticipate the responses of societies to policy? I want to stress that prediction is different from understanding. The process of reduction to elemental components provides a powerful framework for understanding, but it does not suffice for prediction. It is one thing to know a bee is made of atoms, but quite another to guess which way it will fly when it leaves the hive. Scientists nevertheless have abstracted from the bee's behavior certain patterns that make such predictions possible. In a well-defined domain of bee society a set of phenomena has been isolated from all the rest that shows a consistency in repeated trials that allows at least statistical forecasting under defined conditions. Closer to our topic today, it does not suffice to know how laws are made, budgets passed, or how Congress responds to its constituents, if our goal is to understand how society at large responds to our policies.

In this connection, it is worth remarking that the predictions scientists make when conducting experiments differ strongly from the forecasting of actual conditions or events in uncontrolled natural settings. When doing science we attempt to simplify the conditions of experiments to match the limited predictive capability of theory. When we try to forecast actual events we aim for a simulation of "the world" that takes all relevant variables into account. This distinction is a central point of the useful book *"Prediction: Science, Decision Making, and the Future of Nature"* edited by Daniel Sarewitz, Roger A. Pielke, Jr., and Radford Byerly, Jr. (Island Press, 2000). Less ambitious simulations are a useful means of gaining insight into complex issues, and they have a long history in the social sciences, beginning with the imaginative work on game theory by von Neumann and Morgenstern. Forecasts (as opposed to limited simulations) of complex events such as environmental, economic, or public health conditions, are exceptionally

difficult because of the very large number of relevant variables, and the stochastic nature of human affairs.

Difficult as it may be, the predictability of human affairs is a precondition for effective policy. Policies are irrelevant if we have no idea how society will react to them. Or perhaps I should admit that policies can be used as tools for advocacy or for leadership even when they have no empirical foundation. A policy proposal can be a goad, and the policy itself an instrument for change even if it has only symbolic significance. That is not the kind of policy of interest to us here, or at least to me. I am interested in policies that drive the investment of substantial resources, including the expenditure of public funds. Public officials need more than advocacy to make responsible decisions. I view this kind of policy-making almost as a branch of engineering whose effectiveness derives from the science of human behavior. In this science, as in physics or biology, prediction requires identifying relatively decoupled systems, and a logical machinery of empirically validated models and associated analytical power commensurate with the complexity of the systems. I do not know whether today's rapidly developing information technology is yet up to the task of dealing with the massive complexity likely to be encountered in useful policy forecasting, but I am pretty sure it has outpaced our progress in identifying systems, and building and evaluating models that are necessary for good science policy.

The problems of predicting responses to policy actions go well beyond model-making and data analysis. They include the practical and expensive process of obtaining and validating data once you decide what you need. Some data may be culturally sensitive, or controlled, or widely dispersed. Historical data series whose significance has been overtaken by changing practices persist because of a kind of inertia that always assigns value to time-series with a long history. All these issues are grist for the academic mill, but the core issue is the discovery of measures that are causally linked in such a way that the achievement of societal objectives can be confidently traced to policy actions. I have not been able to convince myself that we have such measures today.

More than a year ago I expressed this concern at the annual policy forum of the American Association for the Advancement of Science (April 21, 2005). That forum focuses mostly on science funding, and asks (among other things) whether we are funding all the R&D we need to defend ourselves, improve and sustain our quality of life, and compete with other nations in a globalized high-technology economy. At the time I said that I do not know of any reliable way to answer this question short of developing a massive econometric model for the world's economies and workforces, and exercising it with various scenarios. I have learned since then not to pin all my hopes on econometrics, but there are many ways to model social behavior, and some of them would illuminate issues of great interest to science policymakers. Modern information technology brings much power to the task. But we do not have such models now.

My advice to students of science policy was that "It is well to keep in mind how primitive the framework is that we use to evaluate policies and assess strength in science and technology. In the absence of models that link inputs like federal R&D investments to outputs like Gross Domestic Product per capita, we collect annual data and fit straight lines to them to forecast future conditions. We try to interpret the data by taking various

ratios, plotting the results in different ways – on semi-log graphs, for example – and then talking about the results based on our intuitions about what it all means." The interpretation of numbers requires comparison with other numbers, and the choice of what to compare implies a model of social behavior – often a crude and almost certainly incorrect model. Other fields of social science, such as studies of the behavior of markets or consumer choices, do much better than this and have a richer variety of scholarly investigations to inform practitioners.

Some of the results of what might be called the naïve approach are useful for advocacy. They wake us up to changes so rapid they have to be important somehow – the rate of production of engineering degrees in China, for example, or rates of publication in technical journals, or government investments in different fields. But these "benchmarks" are not of themselves reliable guides for policy-making. I particularly deplore the use of federally funded R&D per unit of GDP as a measure of appropriate R&D investment. No rationale exists for selecting any value of this ratio as ideal, or for maintaining it at a particular value over time, or for adopting the same ratio as a target for different states or nations. The misuse of this and other "benchmarks" in widely publicized advocacy documents has distorted the public discourse on science policy and is likely to be counterproductive in the long run. I expressed this concern in a Science magazine editorial a year ago titled *"Wanted: Better Benchmarks."* (Science **308,** 1087 (2005))

In the U.S. current best practice is embodied in the NSF Science and Engineering Indicators Program that produces an outstanding series whose volumes are full of analysis as well as data. These are objective, high quality documents full of excellent insights. Unfortunately the indicators are based on an old data structure that is not well adapted to how R&D is actually conducted today. And they are not linked to an overall interpretive framework that has been designed to inform policy. These problems and recommendations about how they might be addressed are discussed in the National Research Council report *"Measuring Research and Development Expenditures in the U.S. Economy"* (National Academies Press 2005) which found that "The NSF research and development expenditure data are often ill-suited for the purposes to which they have been employed. They attempt to quantify three traditional pieces of the R&D enterprise – basic research, applied research and development – when much of the engine of innovation stems from the intersection of these components, or in the details of each. … [T]he data are sometimes used to measure the output of R&D when in reality in measuring expenditures they reflect only one of the inputs to innovation and economic growth. It would be desirable to devise, test and, if possible, implement survey tools that more directly measure the economic output of R&D in terms of short-term and long-term innovation. Finally, the structure of the data collection is tied to models of the R&D performance that are increasingly unrepresentative of the whole of the R&D enterprise."

The growing importance of R&D within our society, and its strong association with national priorities, demands much more than the kind of improvements recommended in the NRC report. My perception of the field of science policy is that it is to a great extent a branch of economics, and its effective practice requires the kind of quantitative tools economic policy makers have available, including a rich variety of

models, and a base of academic research. Much of the available literature on science policy is being produced piecemeal by scientists who are experts in their fields, but not necessarily in the methods and literature of the relevant social science disciplines needed to define appropriate data elements and create econometric models that can be useful to policy experts.

During the past year NSF has developed a 5-year plan for the Social, Behavioral and Economics Sciences Directorate to develop the data, tools and knowledge needed to develop a new science of science policy. One track of this program will focus on data and improve and expand science metrics. The other track will focus on producing knowledge from data through the development of models, econometric and statistical tools aimed at increasing our understanding of how innovation occurs and why some investments pay off while others do not. Other agencies too are taking steps to develop new models to perform systematic evaluations of federal R&D programs, to search journal and grant abstracts so as to analyze the status of various fields and identify research gaps, and to understand the long-term outcomes of basic R&D investments.

Under the aegis of the National Science and Technology Council, my office has established an interagency process to foster a new "science of science policy" for improved assessment of the impact of R&D investments, identifying new metrics, understanding the effect of globalization on science and technology, and improving the basis for national science policy decisions. The Atlanta conference I mentioned earlier, and this one, are examples of the kind of activity we want to encourage.

Since my talk last year I have learned much about the field of science policy that I wish I had known when I came to Washington in 2001. Scientists, even those active as advisors, do not hear much about scholarly investigations of science policy. Science policy scholars are rarely included on science agency advisory panels, and are not highly visible even at science policy forums organized by science and engineering professional societies or advocacy groups. Politicians, prominent scientists, and government officials have higher visibility in public discussions of science policy than the men and women who devote their lives to studying this fascinating topic. I hope that conferences such as this one occur more frequently and increase the visibility as well as the size and effectiveness of the community of social scientists who study science policy.

Thank you for the opportunity to bring these concerns to your attention.