Scientists bring an exacting, rigorous analytical approach to their research, but don't always bring that rigor to other pursuits. In the case of science policy, the lack of rigor is partly due to gaps in knowledge and metrics. We're fortunate to have Julia Lane in the ASA community. Julia is the program director of the Science of Science and Innovation Policy program at the National Science Foundation, a program to address the gaps in science policy. I'm delighted that Julia was willing to pen this column to inform and engage the statistics community. What better way for statisticians to influence science policy than through the science of science policy?

~Steve Pierson, ASA Director of Science Policy

The Science of Science Policy: Opportunities and Responsibilities for Statisticians

Julia Lane, National Science Foundation

Federal investments in science and technology have had an enormous effect on innovation, economic growth, and social health and well-being. As recognized by the America Creating Opportunities to Meaningfully Promote Excellence in Technology, Education, and Science Act (or the America COMPETES Act), future investments by the federal government will be critical in many arenas, such as mitigating the consequences of global climate change, exploring new energy sources, defending against external threats, and maintaining international competitiveness. At the same time, the rest of the world is vastly increasing its scientific investments, and there is increased foreign competition for scientific ideas and talent. The United Kingdom has established a Department for Innovation, Universities, and Skills; Saudi Arabia has invested $6 billion to establish a University of Science and Technology; and almost every developed European and Asian country is aggressively investing in and competing for scientific talent.

Given the importance of science policy, it is imperative that science policy decisionmakers have at their disposal the most rigorous tools, methods, and data to develop sound investment strategies. Unfortunately, science policy discussions are frequently dominated by advocates for individual scientific fields who argue for their particular interests, but leave policymakers with little ability to objectively discriminate between investment options. Policy decisions may be based on past practices or data trends that may not always accurately reflect current conditions.

The federal government has taken two steps to respond to this need. First, the National Science Foundation established an interdisciplinary program: the Science of Science and Innovation Policy (SciSIP), which you can learn more about at *www.nsf.gov/funding/pgm_summ.jsp?pims_id=501084*. This program has three major goals:

advancing evidence-based science and innovation policy decisionmaking

building a scientific community to study science and innovation policy

leveraging the experience of other countries

In addition, the statistical arm of NSF, Science Resource Statistics (SRS), has taken a number of steps to measure R&D and innovation better. One is survey redesigns, notably the Business R&D and Innovation Survey and the Academic R&D Survey (presently on hold during a continuing resolution). SRS also is working with the Bureau of Economic Analysis (BEA) to develop R&D satellite accounts and linking NSF Industrial R&D data to BEA data on foreign direct investment. SRS is working to improve information about scientific human resources by developing a data collection system to increase the quality and quantity of information on post-doctoral students and early career scientists and engineers, as well as obtaining information about field of degree on the American Community Survey.

> "The challenge to statisticians is clear, as the rapid advances in cyber infrastructure mean new ways of collecting data exist."

Second, the National Science and Technology Council's Subcommittee on Social, Behavioral, and Economic Sciences commissioned an Interagency Task Group (ITG) on Science of Science Policy. The ITG published a road map (*http://ostp.gov/cs/nstc/documents_reports*), "The Science of Science Policy: A Federal Research Roadmap," that represents the first organized description of the emergent field of the science of science policy, outlining scientific theories and defining terms that encompass efforts in the field so far. It highlights the potential to greatly increase the knowledge base and provide needed insights to improve the data, tools, and methods that would enable a more rigorous and quantitative basis for science and technology policy.

Opportunities and Responsibilities

Some of the opportunities and responsibilities for statisticians involve providing input into the appropriate use of existing sets of models and tools as identified by the ITG to address science policy questions. Others involve bringing statistical tools to bear on new ways of analyzing and describing complex relationships. Yet others involve providing statistically valid approaches to collecting and disseminating data.

By way of illustration of the first set of opportunities, the road map identified a set of models, tools, and metrics most useful for federal agencies in addressing the scientific questions, particularly whether it is possible to "predict" scientific discovery or "predict" the impact of scientific discovery. These included many approaches familiar to the statistical community, such as deterministic models (econometric, risk modeling, options modeling, cost benefit, cost effectiveness) and stochastic models (agent-based modeling and system dynamics). A few agencies, such as the Department of Energy and National Institutes of Health, have begun to experiment with dynamic modeling and options modeling as ways to describe the effect of discovery. Statisticians could help provide guidance about the quality and reliability of the resulting inferences.

Other identified approaches that are less familiar included visual analytics, scientometrics, and network analysis. These approaches offer intriguing possibilities for tracking the impact of investments in science. The possibilities range from tracing the path from basic research discoveries to patents and innovation to the changing structure of scientific disciplines, and from examining the importance of social networks to the dispersion of scientific innovations to comparators of international performance in science. However, before such a vision is achieved, many statistical questions about the robustness, validity, and usability of the visualization tools remain to be answered. How robust are different taxonomies to different mapping algorithms? How robust are the apparent relationships to different distance metrics? What do visual relationships among different scientific units of analysis mean? What do changes in the visual relationships mean? What statistical models can be applied to visualization algorithms to validate relationships and predictability of how they are likely to evolve? How replicable and generalized are the results of visualization techniques?

A necessary component for developing an evidence-based platform for science policy decisionmaking is the development of an appropriate (i.e., statistically valid) microdata infrastructure. The ITG identified four key areas in which such an infrastructure is necessary: measuring and tracking federal funding of science, measuring and tracking the scientific work force, measuring and tracking scientific outcomes, and measuring competitiveness. It also recognized the importance of providing analytical access by researchers and federal government agencies. The challenge to statisticians is clear, as the rapid advances in cyber infrastructure mean new ways of collecting data exist. These include web scraping, text and video data mining, and new uses of administrative data, but the analytical reliability of such sources is relatively unknown.

Engagement of the Statistical Community

The statistical community can engage with the Science of Science Policy effort in a number of ways. Statisticians are encouraged to submit proposals to the NSF in the area of Science of Science and Innovation Policy. The SoSP Interagency Group established an electronic mailing list and wiki to engage the SoSP community. The Science of Science Policy web site, *http://scienceofsciencepolicy.net*, serves as the central point for data sharing and information dissemination and to communicate new events. Statisticians also can use the site to provide input via the polling and assessment tools or wiki once they have logged in. Those who are interested in participating should send an email to Julia Lane at *jlane@nsf.gov* for information about how to log in to the web site. Statisticians also can join the electronic mailing list by sending a blank email to *subscribe-scisip@lists.nsf.gov* and then replying to the automatic response email.

To summarize, the burgeoning interest in creating an evidence-based platform for science policy decisions will require the input of the statistical community to ensure high-quality decisionmaking. We look forward to the involvement and activity participation of statisticians in developing the science of science policy. ■