This product is part of the RAND Corporation conference proceedings series. RAND conference proceedings present a collection of papers delivered at a conference. The papers herein have been commented on by the conference attendees and both the introduction and collection itself have been reviewed and approved by RAND Science and Technology.

CONFERENCE PROCEEDINGS

The U.S. Scientific and Technical Workforce

Improving Data for Decisionmaking

Terrence K. Kelly, William P. Butz, Stephen Carroll, David M. Adamson, Gabrielle Bloom, editors

CF-194-OSTP/SF

June 2004

Prepared for the Office of Science and Technology Policy and the Alfred P. Sloan Foundation

SCIENCE AND TECHNOLOGY

The research described in this report was conducted by the Science and Technology Policy Institute (operated by RAND from 1992 through November 2003) for the Office of Science and Technology Policy under Contract ENG-9812731 and the Alfred P. Sloan Foundation.

ISBN: 0-8330-3651-3

Published 2004 by the RAND Corporation
1700 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138
1200 South Hayes Street, Arlington, VA 22202-5050
201 North Craig Street, Suite 202, Pittsburgh, PA 15213-1516
RAND URL: http://www.rand.org/
To order RAND documents or to obtain additional information, contact
Distribution Services: Telephone: (310) 451-7002;
Fax: (310) 451-6915; Email: order@rand.org

Contents

Part III
Summary and Recommendations

CHAPTER THIRTEEN

Rapporteur's Summary

William P. Butz[1]

This conference afforded an extraordinarily rich cross talk among experts who make policy that affects the STEM workforce, analysts who study it, and statisticians who design and collect the data that support both policy and analysis. Despite the diverse perspectives represented here, there was an underlying consensus on the purpose of these deliberations: to identify specific actionable improvements in the data system for decisionmaking.

Keeping that aim in view, this section summarizes the conference's dominant themes. It first proposes that the questions addressed can effectively be distilled to three: Who are the decisionmakers that need the data? What data do they need? What data can suppliers provide? In addition, a set of crosscutting themes relevant to more than one of these questions emerged.

Decisionmakers in the Markets for STEM Workers

Who are the decisionmakers to be informed? Conferees identified five groups of decisionmakers. In addition, the discussion pointed to a sixth group with special needs: researchers.

Policymakers in Science Funding Agencies

These agencies, NSF and NIH principally among them, strongly influence the numbers of STEM students in the educational pipelines. At the same time, they provide much of the support in many fields for the job positions and laboratories that will or will not await these students when they enter the job market. When deciding, for example, whether to increase or decrease fellowship and postdoctoral support, policymakers should balance trend information on numbers of positions by field against data on numbers of students becoming available to fill them.

Program and Human Resource Administrators in Federal and Other Public Agencies That Employ Scientists and Engineers

These officials infer STEM personnel requirements from prospective changes in their agencies' responsibilities and in the science that underlies their work. In satisfying these requirements, they must balance uncertain staff retirement and retraining prospects against the possibility of hiring replacements and new types of specialists from the labor market. Program and budget planning typically requires human resources departments to specify these hiring

[1] William P. Butz is President and Chief Executive Officer of the Population Reference Bureau in Washington, D.C., and a consultant to RAND.

plans several years ahead of the need. Some agencies mitigate some of the risk involved by contracting work requiring STEM personnel to the private sector.

Private-Sector Employers, Colleges, and Universities

These entities face similar balancing acts as they forecast the demand for their products and services, the staff on hand to produce them, and the availability of suitable personnel to fill replacement and new positions. The increasing possibility of outsourcing work to other countries does not diminish the requirement for planning and balancing; to the contrary, it extends the geographic domain for which data on the STEM workforce are valuable.

Members of the STEM Workforce

Whether driven by scientific discovery, technical development, market demand for the products of technology, or federal funding, wide swings in employer demand for successive science and engineering specialties have characterized U.S. labor markets since the mid-1960s. Most American STEM workers now undergo some form of mid- or late-career retraining—in colleges, employers' laboratories, or on the fly. The more accurately workers can discern the changing demand for their specialty or for other specialties not far removed, as well as the changing supply of workers in these fields, the better they can identify approaching job trouble or opportunities. Families, employers, and the nation will suffer less economic deprivation, production cost increases, and unemployment, respectively.

Students and Their Advisors

A high school senior working toward a technical associate degree only needs to look ahead two or three years to assess what jobs will be available to reward that training. A high school senior working toward a doctorate and a postdoctoral appointment, however, must try to look ahead many more years—at least a dozen, on average, in biology, perhaps a year or two less in other sciences. College seniors must also look ahead several years in deciding among careers in science, on the one hand, or in medicine, dentistry, veterinary science, law, or business, on the other. Thus, long-term information about job markets, trends, and prospects can be invaluable to career decisions.

While no crystal ball is available, current information about trends in job offerings by degree field, the number of students already in the pipeline, and underlying changes in government funding can at least suggest which fields have unattractive near-term employment prospects.

Labor Market Researchers

The estimates and insights of the economists, sociologists, and psychologists who study the labor market help others make informed decisions. This work clarifies the influences on student choice of college major and graduate field, on worker decisions about job choice and retirement, and on employer decisions about job characteristics and recruitment. Such research requires data beyond counts of workers and jobs.

At every level—from science policy administrators to guidance counselors to high school students—there are decisionmakers. When federal statistical agencies or industry associations do not provide objective and timely data on STEM jobs, job holders, and job seekers, responsible decisionmakers will instead collect anecdotes, impressions, and whatever else they can find to inform one of life's most consequential decisions. Moreover, when these

responsible organizations do not provide credible analyses of these data—analyses that turn the numbers into useful information—decisionmakers perform their own analyses, however implicit and simplistic.

Crosscutting Themes

The conference discussions yielded a set of specific data requests from the researchers and decisionmakers represented, as well as responses from the statistical agencies represented, about priorities and practicalities. Cutting across these specific data items were some general themes and issues that spanned the different decisionmaking and data domains and connected many of the otherwise diverse topics.

The first of these—general STEM competency—is ideally a national goal that underpins maintaining a quality STEM workforce. The next three themes—revisiting assumptions about STEM careers, data coordination and collection, and identifying workforce shortages and surpluses—are means of achieving a range of goals. The final theme, the global dimension, is an emerging variable that must be factored into STEM workforce decisionmaking.

A Broader Context: General STEM Competency

It is important to bear broader contexts in mind when considering STEM workforce data issues. Among the most important of these is promoting the goal of STEM competency among the general public. STEM competency is important at all levels from the "ordinary citizen" through those holding science doctorates. Information about the effectiveness of K–12 science and mathematics education is important because it both lays the foundation for entering and succeeding in scientific fields and contributes significantly to the success of informed citizens in any occupation.

Revisiting Key Assumptions About Training and Careers

Data collection and analysis are often influenced by assumptions about typical career paths and job progressions, which may not accurately reflect shifting and complex occupational realities. Many of these assumptions need to be revised. For example,

- Industry employment is quantitatively important in most STEM fields. While this may seem evident, some data collection and analyses still proceed as if newly minted scientists move solely or mainly into college teaching and as if professors do not commonly work in both worlds, simultaneously or sequentially.
- Life-cycle progressions—starting with choice of college major, moving through career advancements and choices, and ending with retirement—are poorly understood for people in STEM professions. Even for the archetypal career path into college teaching, we lack basic descriptions of these progressions, the key decision points, and the influencing factors.

Data Coordination and Collection Issues

The conference discussion focused primarily on federal data collection and statistical analysis. This point raised several issues:

- Further coordination of statistical concepts and definitions and productive sharing and linking of data across federal statistical agencies would contribute substantially to the completeness and usefulness of data for STEM decisionmakers of every type.
- Although the data that federal statistical agencies collect command by far the most attention from researchers and STEM decisionmakers, considerable data are also available from federal and state administrative agencies and from professional associations and other private sources. For some important concepts, the latter organizations are the only current sources. Coordinating and documenting these diverse collections of data, presumably by the federal statistical system, could provide a considerable payoff.
- Expanding the access nonfederal researchers have to data would also increase the useful information available to all STEM decisionmakers. Directed research and coordinated federal effort could mitigate the effects of confidentiality provisions and the technical and bureaucratic difficulties of linking related data sets.
- When possible, it would be best to satisfy the requirement for new or different data by adding question modules to existing survey programs, rather than beginning new surveys.

Identifying Workforce Shortages and Surpluses

Existing data suitable for monitoring possible shortages or surpluses among STEM workers by field are not sufficient for understanding and improving observed patterns:

- People trained in STEM fields but working in other occupations must be accounted for to document and understand the existence of shortages or surpluses in STEM fields. The progression of scientists to general administration is only one example.
- Cross-sectional and longitudinal data are both critical. In particular, labor-market research requires data on worker and employer characteristics and on the community context, as well as a longitudinal survey structure that documents life-cycle changes for the same samples of workers. Such data are essential for estimating the labor supply elasticities that characterize student and worker responsiveness to changing job availability and earnings.
- In addition, quick-turnaround data systems can and should be designed to identify STEM fields at risk of emerging shortage or surplus. These relatively crude indicators of job openings, job applications, salaries, graduate student enrollments, and job placements, for example, would direct more substantial ongoing data collections toward particular STEM fields, revealing whether the concern is justified. In the meantime, students contemplating fields of study would have tentative indicators several years earlier than they do at present, a substantial difference at that life stage.
- Data on workers are many times more plentiful than data on their employment conditions. Expending the time and money necessary to document the demand side of the scientific labor market—job slots, job offers, characteristics of facilities and equipment that complement or substitute for labor input, characteristics of demand for the products or services the employer produces—will have a big payoff in elucidating the sources of STEM worker shortages and surpluses.
- Additional data on how firms adjust to occupational labor shortages are critical to understanding whether shortages reflect a short-term lack of equilibrium or long-

term structural shortages. To the extent that rising wages do not elicit a sufficient supply of workers in the short, medium, or even long runs, firms may adjust in other ways, often affecting working conditions and the quality of STEM labor. These adjustments can include changes in overtime hours, increased use of immigrants, lowering educational attainment requirements, or the use of offshore or contract labor.

- Finally, forecasting is an important issue. Forecasts of labor market conditions received little attention and emphasis at the conference because most participants considered them too inaccurate to be useful. However, if such forecasts could be made more accurate, they would be valuable.

The Global Dimension

A new variable that must be taken into account in considering these is the global dimension. Changes in the numbers of foreign students enrolled here and the foreign scientists employed here; in the international movement of U.S. scientists during the course of their work; and, most recently, in the outsourcing of technical and even scientific work to less-expensive workers abroad—these trends have integrated U.S. scientific institutions and their employees with foreign counterparts, in many cases quite inextricably. Decisions based solely on U.S. data about Americans may increasingly paint an incomplete and possibly misleading portrait of the market for STEM workers.

What Data Do Decisionmakers Require? What Data Can Producers Supply?

Workshop participants from the research community and from organizations that employ STEM workers identified particular data items and data changes that would improve analysis and decisionmaking. In turn, participants from the federal statistical system addressed each of these requirements, pointing out some that are already being satisfied, others that could be met with more or less difficulty, and still others that might be met by substituting available proxy information. Indeed, an immediate payoff of the workshop was that data users discovered that some requested data are already available, either as described or in the form of a useful proxy. Below are descriptions of the data that the users—the organizations and researchers—requested, alongside statistical agency indications of current and potential availability. The workshop participants' presentations in Part II provide the background justification and context for most of these requests.

- Education and training
  - Collect median and variance of length of graduate program and probability of completion data, by field. Not available apart from statistics kept by several professional associations.
  - Collect data on numbers and characteristics of persons who drop out of a STEM educational process, by discipline, or who leave STEM occupations while still employed, by occupation entered. Not available or planned.
  - Collect data on numbers and characteristics of "false starts" in the educational process, by level and field. Not available or planned.

  - Compile clearinghouse of demonstrations and experiments designed to improve training, retraining, job search, or career advancement of STEM workers. Not available or planned.
  - Compile information on alternative educational paths to STEM careers. Not available or planned.
  - Collect data following foreign students through the U.S. educational process. Student and Exchange Visitor Information System (SEVIS) data, still being developed, are a big improvement.
  - Define an "understanding of science" in the citizenry and articulate how this is measured. Not available or planned.
  - Collect data on numbers of students by discipline, degree program, stage, and citizenship status, reported within three months of the beginning of the academic year. NCES and SRS collect a subset of these data.

- Degree holders and postdoctorates
  - Collect data on numbers and characteristics of persons holding postdoctoral positions. SRS is beginning methodological research on administrative data that NIH and NSF have collected.
  - Collect data on holders of multidisciplinary degrees comparable to data available for traditional disciplines. SRS is conducting methodological research.
  - Study numbers and characteristics of foreign STEM degree holders in the United States. Not available or planned.
  - Compile data on STEM workers with degrees and certifications below the baccalaureate level—for example, technical associate in arts degrees and software-manufacturer training certifications. SRS and NCES are working collaboratively to develop these data.
  - Add STEM workers with clinical degrees and without doctorates to the Survey of Earned Doctorates (SED) and the Survey of Doctorate Recipients (SDR). Not available or planned.
  - Include doctorates from foreign institutions in the sample frames for federal surveys of STEM workers. Not available or planned.

- STEM workforce: jobs and occupations
  - Collect employment data that provide greater detail on occupations at local, state, and national levels. Available at BLS.
  - Collect data on job vacancies and turnovers, by occupation. Not available or planned.
  - Compile data on newly emerging occupations. Not available or planned.
  - Compile detailed data on net job replacement requirements, by occupation. Available at BLS.
  - Study job competencies that employers seek, apart from degrees and certifications. Not available or planned.
  - Compile data on hiring at venture capital firms, by occupation, as a leading indicator of needed skills. Not available or planned.
  - Compile data on the number of job offers, by occupation. Not available or planned.

  - Follow up on the 30 percent of SED respondents who report "no definite plans." Not available or planned.
  - Track STEM workers in industry employment. Not available or planned.
  - Create longitudinally matched data on employers and STEM employees in the U.S. Census Longitudinal Establishment Data (LED) file. This should not be limited to the Census LED but should be broadened to include linked employer-employee establishment data files. Not available or planned.
  - Assemble characteristic data on large samples of STEM workers more frequently. The Census Bureau's American Community Survey will produce such data on very large samples between censuses.
  - Collect data on job offers, acceptances, and salaries in aggregated occupational categories, reported three months after close of the survey period. Not available or planned.

- Metrics and Databases
  - Create a national STEM database. Not available or planned.
  - Develop a common set of concepts and definitions across federal statistical agencies. Talks among agencies are continuing.
  - Develop measures of lifetime risk of unemployment and career change, by STEM discipline. Not available or planned.
  - Create a unified bibliography of publications and citations regarding the STEM workforce. SRS is working on methodological improvements to existing bibliographies.
  - Maintain and improve response rates for STEM surveys. Statistical agencies maintain substantial research on this challenge.
  - Facilitate data-sharing across federal statistical agencies to produce matched files for analysis. OMB and the statistical agencies are working on this challenge.
  - Make it easier for university researchers to use federal microdata. OMB and the statistical agencies work continually on this challenge.
  - Construct an omnibus sample frame and a related program of "snap" surveys to produce very quick turnaround data on particular policy-related questions. Not available or planned.

- International issues. Nearly all the other recommendations have an international dimension that needs to be documented. We therefore list the international data issues separately to highlight their importance, even though they relate closely to many of the other items:
  - Assemble internationally comparable data on numbers of STEM workers and of STEM jobs. Not available. NCES is working with Statistics Canada, OECD, and UNESCO to develop comparable concepts and definitions. NCES is working with the Department of Homeland Security (DHS) to obtain information on degrees and occupations of STEM visa holders working in the United States.
  - Collect information on international mobility of foreign STEM workers educated in the United States. SRS is conducting methodological research based on DHS data.
  - Collect information on STEM jobs outsourced to other countries by U.S. employers. Not available or planned.

– Survey employers on why and under what circumstances they move jobs overseas. Not available or planned.

Concluding Observations

Again and again, the conference discussion returned to the connection between data and decisionmaking and reiterated the basic point that decisionmaking does not grind to a halt in the absence of adequate data. It simply proceeds with inadequate data. Employers and managers who lack a credible information base produced by statistical experts may base decisions on information and analyses that they themselves have produced, often on the fly, or that are produced by others lacking statistical expertise. Administrators of science funding agencies who lack such information may base funding allocations across scientific disciplines on judgments about where the science is most exciting or where other support is lacking, to the detriment of students encouraged toward fields, however exciting, without waiting jobs. Moreover, without understanding key decision points for STEM students and workers, universities and their science funders cannot efficiently design interventions to affect such decisions. Among all these decisionmakers, students and workers are the most disadvantaged, for they typically command insufficient resources to uncover any but the most rudimentary information about trends in potentially interesting fields. And yet, ironically, it is they who bear the largest burden of mistaken decisions—lengthy training and uncertain outcomes, job insecurity, and potential disillusionment with the scientific enterprise.

Of the numerous data requirements that surfaced at the conference, the eight highest priority needs are presented in detail in Chapter Fourteen.