

A Toolkit for Evaluating Public R&D Investment

Models, Methods, and Findings
from ATP's First Decade

Prepared for
Economic Assessment Office
Advanced Technology Program
National Institute of Standards and Technology
Gaithersburg, MD 20899-4710

By
Rosalie Ruegg
Managing Director
TIA Consulting, Inc.
ruegg@ec.rr.com

Irwin Feller
Senior Visiting Scientist
American Association for the Advancement of Science
Professor Emeritus
Pennsylvania State University
iqf@psu.edu

July 2003



U.S. DEPARTMENT OF COMMERCE
Donald L. Evans, Secretary

TECHNOLOGY ADMINISTRATION
Philip J. Bond, Under Secretary of Commerce for Technology

NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY
Arden L. Bement, Jr., Director

Contents

Note from Project Manager **xiii**

About the Authors **xv**

Acknowledgments **xvii**

Abstract **xix**

Executive Summary **xxi**

Evaluation Underpinnings xxii

Multi-Faceted Methodological Approach xxii

An Emerging Body of Findings xxiv

Firm/Industry Effects xxv

Collaboration Effects xxv

Spillover Effects xxvi

Interfaces with State Programs and Comparison with Counterpart
 Programs Abroad xxvii

Overall ATP Performance xxviii

Recommendations for Future Directions xxviii

Acronyms and Abbreviations **xxx**

Introduction **1**

The Political Economy of the Advanced Technology Program 3

The Role of Evaluation 8

PART I: EVALUATION FRAMEWORK

Chapter 1. Evaluation Fundamentals 13

Objectives of Evaluation	13
Mapping Evaluation to Mission and Stages of Implementation:	
A Generic Evaluation Logic Model	17
Evaluation Steps	19
Best Practices	20

Chapter 2. Choosing Methods of Evaluation 27

Analytical/Conceptual Methods for Modeling and Informing	
Underlying Program Theory	28
Survey Method	29
Case Study: Descriptive	34
Case Study: Economic Estimation	35
Sociometric/Social Network Analysis	45
Bibliometrics: Counting, Citing, and Analyzing Content of Documents	47
Historical Tracing	50
Expert Judgment	52
Suggested Readings on Evaluation Methods	53

Chapter 3. ATP's Evaluation Program 61

Background: Evaluation Drivers	61
ATP's Evaluation Logic Model	63
Conceptual Tests of ATP's Success	67
ATP's Approach to Evaluation	68

PART II: APPLYING EVALUATION METHODS TO ATP

Chapter 4. Modeling and Informing

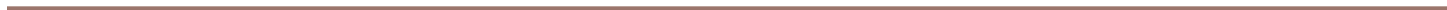
Underlying Program Theory	85
Concepts, Models, Metrics, and Paths Connecting Program Activities to Intended Impacts	91
Conditioning Expectations through Studies of Private-Sector Behavior	103

Conditioning Expectations through Studies of Other Public Sector Programs	116
Testing Dominant Paradigms	127
Summary of Research Informing Underlying Program Theory	135
Chapter 5. Survey Method	139
Gaining Early and Broad Perspective of a Program's Effects	142
Extending and Deepening Survey Data on Program Effects	146
Establishing Routine Project Reporting by Electronic Survey	149
Soliciting Feedback by Survey on Customer Satisfaction and Marketing Issues	156
Using Survey for Case Study and to Address Research Questions	158
Summary of ATP's Use of the Survey Method	166
Chapter 6. Case Study Method	169
Economic Case Study of Individual Projects	169
Project and Portfolio Assessment Using Multiple Cases Studies with Uniform Collection of Key Indicator Data	202
Explicating Program Features and Exploring Program Dynamics	210
Summary of ATP's Use of the Case Study Method	214
Chapter 7. Econometric/Statistical Method	217
Testing ATP's Leveraging Effects on Advanced Technology Development ...	219
Modeling Impacts of Public-Private Partnerships on Firm Productivity ...	222
Spillovers, Appropriability, and Firm Productivity	224
Analyzing the Role of Universities in Public-Private Partnerships	228
Modeling the Impact of Publishing by Industry Scientists on the Quality of Innovative Output	231
Investigating Characteristics and Impacts of Joint Ventures	236
Summary of Econometric/Statistical Methods	247
Chapter 8. Other Evaluation Methods Used by ATP	251
Expert Judgment	251
Bibliometrics	264

Emerging Methods: Using Social Network Analysis to Identify Knowledge Spillovers	270
Emerging Method: Using the Cost Index Method to Estimate Social Benefits	279
Emerging Method: Using a Composite Performance Rating System Constructed from Case Study-Derived “Indicator” Data	284
Summary of Other Evaluation Methods	290

PART III: AN EMERGING BODY OF KNOWLEDGE

Chapter 9. A Crosscutting Look at Study Findings	295
Direct Private Firm Effects	295
Collaboration Effects	320
Spillover Effects	336
State and International Technology Programs	343
Measures of Overall ATP Performance	349
Summary of Crosscutting Findings	364
Chapter 10. Conclusions and Recommendations ...	367
Program Evaluation: An Essential Tool for ATP	367
Opportunities and Future Directions	370
Closing Note	377
References	379
About the Advanced Technology Program	inside back cover



Figures

Generic Evaluation Logic Model	xxiii
ATP’s Evolving Use of Methods Over its First Decade	xxiv
Highlights of Main Topics	xxix
Figure 1–1. Evaluation Logic Model	18
Figure 3–1. ATP’s Evaluation Logic Model	65
Figure 3–2. Intensity of ATP’s Use of Evaluation Methods	69
Figure 4–1. Private and Social Returns to R&D: Pure Market Spillover, Plus Pure Knowledge Spillover, Plus Interaction of the Two	93
Figure 4–2. Three Dimensions of ATP and the “Journey to Q”	96
Figure 4–3. Direct and Indirect Paths to ATP’s Impacts	99
Figure 4–4. Social Benefits from Product Innovation that Reduces Costs of Industries Using It	105
Figure 4–5. Reducing Inaccuracies in Profitability Estimates Over Time	106
Figure 4–6. Roadmap of Financing Options	115
Figure 4–7. Strategic Mind Map for a Rich Technology Platform	116
Figure 4–8. Conceptual Time Path for ATP’s Technology Diffusion and Expected Impacts	126
Figure 4–9. A Social Interaction View of Innovation	132
Figure 6–1. Elements Determining Social Return on Public Investment and Social Return on Investment	180
Figure 6–2. Valuing Per-Patient Changes in Health Outcomes	183
Figure 6–3. Historical Adoption Rates of Technologies Enhancing Fuel Efficiency, and Implementation Rates used by Ehlen in the Case Study	192
Figure 6–4. Market Drivers for the Further-Processed Food Industry	197
Figure 6–5. Distribution of 27 Completed Projects at Small, Single Applicant Companies by Percentage of Employment Change	206
Figure 7–1. Relationships across Industry R&D Intensity, Intra-Industry R&D Information Flows, and Appropriability Mechanism Effectiveness	226

Figure 7–2. Patenting and Publishing in the Scientific Literature, Gene Therapy: Patents and Publications	233
Figure 8–1. The Classic Bass Diffusion Curve	258
Figure 8–2. Interview Guide Used by RTI Researchers to Compile Judgments from Physicians Who Are Expert in Treating the Relevant Diseases	259
Figure 8–3. Distribution of Projects by Number of Publications and Presentations	267
Figure 8–4. Distribution of Completed Projects by Number of Patents Filed	268
Figure 8–5. Illustrative Patent Tree for ATP	269
Figure 8–6. Conceptual Framework of the R&D Network	272
Figure 8–7. All Citations are not Equal: A One-Way Flow Illustration	273
Figure 8–8. Top Twenty Optical Waveguide Regions	277
Figure 8–9. Derived Demand for New Technologies: Illustration of Net Surplus Change	281
Figure 9–1. Change in the Nature of Industry R&D	302
Figure 9–2. Importance of Timing	307
Figure 9–3. Employment Change at 31 Small Companies Receiving a Single-Company Award	319
Figure 9–4. Key Factors Influencing Collaborative Success	329
Figure 9–5. Illustration of Two Project Structures with Alternative Collaborative Relationships	331
Figure 9–6. Costs Attributed to Collaboration	336
Figure 9–7. Comparison of the Focus of ATP and State Technology Development Programs	349
Figure 9–8. Distribution of Terminated Projects by Reason for Termination	357
Figure 9–9. Distribution of Projects by Overall Performance Score	359
Figure 9–10. Distribution of Projects by Overall Performance Score and Technology Area	360
Figure 10–1. ATP’s Evolving Use of Methods Over its First Decade	369

Tables

Table 1–1. Why Evaluate?	14
Table 1–2. Steps in Evaluation	20
Table 2–1. Overview of Evaluation Methods	30
Table 2–2. Steps in Survey-Based Statistical Studies	33
Table 3–1. ATP’s Mission, Mechanisms, and Features	64
Table 3–2. Select ATP Studies Commissioned and Completed, 1991–1995	71
Table 3–3. Select ATP Studies Commissioned and Completed, 1996–1999	73
Table 3–4. Select ATP Studies Commissioned and Completed, or Delivered in Draft, 2000	76
Table 4–1. Twenty-Two Studies and Papers Modeling and Informing Underlying Program Theory	86
Table 4–2. Classification of Spillovers	92
Table 4–3. Factors Increasing the Likelihood of Spillovers	94
Table 4–4. Recommended Metrics for ATP and Timing of Data Collection	100
Table 4–5. Interview Questions Investigating Determinants of Success of Joint Ventures	102
Table 4–6. Volume of Venture Capital Activity	109
Table 4–7. Categorizing State Technology Assistance Program by Type of Challenge and Stage of Technology Development and Commercialization	120
Table 4–8. ATP Strategies Designed to Promote Technology Diffusion	131
Table 5–1. Eight of Eleven Studies Using Survey Method Represented*	140
Table 5–2. Program Effects Treated in the First Survey of ATP Projects	145
Table 5–3. Illustrative Survey Questions on Commercial Progress	150
Table 5–4. Illustrative Survey Questions Relating to Knowledge Spillovers	151
Table 5–5. Illustrative Survey Questions about Collaborative Partners	152
Table 5–6. Components of ATP’s Business Reporting System	154

Table 5–7. Examples of Variables for Comparing Small and Larger Firms Using BRS Data	155
Table 5–8. Competitive Position of Member Companies in the World PWB Market	159
Table 5–9. Survey Questions Probing the Counterfactual: With and Without ATP	162
Table 6–1. Ten of Sixteen Studies Featuring Case Study Represented	170
Table 6–2. Summary of Survey Findings on Partial Early-Stage Economic Impacts	176
Table 6–3. Overview of ATP Projects Included in this Study	178
Table 6–4. Impact of ATP Funding on the Development of Medical Technologies for Seven Tissue Engineering Projects	181
Table 6–5. Comparison of QALY Utility-Weights for Different Health States	185
Table 6–6. Social Return on Investment and Social Return on Public Investment: ATP Projects in Tissue Engineering for a Single Preliminary Application	186
Table 6–7. Composite Private Returns: ATP Projects in Tissue Engineering for a Single Preliminary Application	187
Table 6–8. Annual Impact on U.S. Macroeconomy of Near-Term, Five-Year Implementation Path: Year 2004	195
Table 6–9. Secondary Market Opportunities for CCAR Technology	198
Table 6–10. Base Case Cash Flows from Improved Quality, Yield, and Production Rates and from Reduced Refrigeration Costs from Application of the CCAR Technology for Food Processing, millions 2001 dollars	200
Table 6–11. Base Case Net Present Value, Internal Rate of Return, and Benefit Cost Ratio (Calculated from the Cash Flows in Table 6–10) for the CCAR Technology	201
Table 6–12. Additional Qualitative Benefits from Using the CCAR Technology	202

Table 6–13. ATP’s Role	207
Table 6–14. Status Report Data	209
Table 7–1. Eight of Ten Studies Using Econometric/Statistical Methods Represented*	218
Table 8–1. Ten Studies Using Expert Judgment, Bibliometrics, and Emerging Methods, including Cost Index, Social Network Analysis/Fuzzy Logic, and Composite Performance Rating System	252
Table 8–2. Data Sample Collected from Expert Physicians and Used to Estimate the Bass Technology Diffusion Model	261
Table 8–3. Dissemination of Non-Proprietary Information from ATP-Funded Projects	266
Table 8–4. Structured Interview Guide for Collecting Data Needed to Exercise the Model	283
Table 9–1. Studies Informing the Impacts of ATP on Private Firms	297
Table 9–2. Illustrative Metrics Showing ATP’s Impact on Firms’ Innovation Goals	303
Table 9–3. The Extent to Which Non-Winners Pursue the Proposed R&D Project Without ATP Funding	308
Table 9–4. Estimates of Economic Value of a One-Year Reduction in Applied Research Cycle Time, in Order of Decreasing Value: \$5 Million to \$6 Million Median Value	310
Table 9–5. ATP Effects that Helped Interviewees to Reduce Cycle Time	311
Table 9–6. Carryover of Cycle-Time Improvements to Other Projects	312
Table 9–7. Effect of ATP Funding on Completed Projects	313
Table 9–8. A Comparison of Small and Larger Firms in ATP	316
Table 9–9. Progress of Participating Companies in Commercializing New Technologies	317
Table 9–10. Examples of Products and Processes from the First 50 Completed ATP Projects	318
Table 9–11. Impact on Industries of Near-Term, Five-Year Implementation Path for Flow-Control Machining Technology	320

Table 9–12. Studies Extending Knowledge about Collaboration Activities . . .	321
Table 9–13. Summary of Study Findings on Frequency of Collaboration . . .	322
Table 9–14. Propensity to Collaborate with Other Organizations and Form New Partnerships	324
Table 9–15. Collaborative Activity of the First 50 Completed Projects	326
Table 9–16. Membership Changes in the Printed Wiring Board Research Joint Venture, 1992–1996	328
Table 9–17. Specific Benefits of Collaborations	333
Table 9–18. Studies Informing Spillover Effects	337
Table 9–19. Characteristics of R&D Networks Generating Pre-Competitive, Enabling Technologies	341
Table 9–20. Spillovers Imputed by Comparing Composite Social Returns, Public Returns, and Composite Private Returns on Seven Tissue Engineering Projects	343
Table 9–21. Studies Comparing State and International Programs	344
Table 9–22. Comparative Features of ATP and Its Analogues	346
Table 9–23. Studies Providing Broader Assessments of Program Performance	350
Table 9–24. Measures of Economic Benefit from Six Detailed Case Studies	361

Note from Project Manager

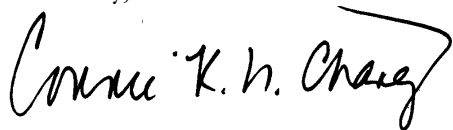
The seed for the idea of constructing a toolkit of evaluation tools for program managers came from attending a workshop I co-organized on “Financing and Managing Technology Development, in Veracruz, Mexico in April 2001. This workshop was held during Mexico’s 13th International Congress for Science and Technology and brought together program managers, officials, and consultants from Norway, the European Union, Canada, the United States, and Mexico to share their experiences. Presenters addressed the financing methods and time-frames for public support of technology development; program management, project selection, and oversight; and benchmarking and evaluation of impact.

Rosalie Ruegg, the co-author of this report, gave a presentation at this workshop on evaluation fundamentals that led to a lively exchange over the most effective evaluation tools and methodologies to achieve policy goals. It was during the plane ride home that the idea for a toolkit for evaluating public R&D investment took root.

The timing of the project was perfect because it came on the heels of a National Research Council report (*The Advanced Technology Program: Assessing Outcomes*, 2001) that found that ATP “set a high standard for assessment involving both internal and independent external review.” Over the past thirteen years beginning in 1990, the Advanced Technology Program through its Economic Assessment Office has been sponsoring rigorous and groundbreaking scholarship that has advanced the understanding of the process of technology-based innovation.

It is with the hope of sharing what we have learned and to educate those new to program evaluation that we bring to you this toolkit. The authors and I welcome your comments and suggestions.

Sincerely,

A handwritten signature in black ink that reads "Connie K.N. Chang". The signature is written in a cursive style with a large, sweeping initial "C" and a distinct flourish at the end.

Connie K.N. Chang

Supervisory Economist

Leader, Policy Research & Analysis Group

ATP Economic Assessment Office

connie.chang@nist.gov

About the Authors

Rosalie Ruegg, managing director of TIA Consulting, Inc., specializes in the economic assessment of new technologies. Recent clients include the National Academies of Science, Harvard University, the Department of Energy, the State of Indiana, Finland's national technology program, Johns Hopkins University, other agencies of government, and other private and non-profit consulting companies. Prior positions include director of the ATP's Economic Assessment Office, senior economist in NIST's Center for Applied Mathematics, and financial economist for the Federal Reserve System's Board of Governors. Among her more than 60 publications is a case-study guide for science managers, proceedings of an international workshop on evaluation, and an economics textbook. She has served on various editorial boards, most recently as economics editor of Macmillan's *Encyclopedia of Energy*. A former member of the Federal Senior Executive Service, she received the Department of Commerce Gold Medal Award for excellence and the Institute of Industrial Engineers' 2001 Wellington Award for contributions to the field of engineering economics. She received degrees in economics from the Universities of North Carolina (Phi Beta Kappa, cum laude) and Maryland (Woodrow Wilson Fellow), an M.B.A from The American University, and extensive executive training.

Dr. Irwin Feller is Senior Visiting Scientist, American Association for the Advancement of Science and Professor Emeritus of Economics, The Pennsylvania State University, where he served on the faculty from 1963–2002, and as director, Institute for Policy Research and Evaluation from 1977–2002. Dr. Feller's current research interests include the evaluation of federal and state science and technology programs, the economics of academic research, and the university's role

in technology-based economic development. He is the author of *Universities and State Governments: A Study in Policy Analysis*, and over 100 refereed journal articles, final research reports, and book chapters. He has been a consultant to the President's Office of Science and Technology Policy, National Aeronautics and Space Administration, the Carnegie Commission on Science, Technology, and Government, the Ford Foundation, the Cleveland Foundation, National Science Foundation, National Institute of Standards and Technology, U.S. General Accounting Office, U.S. Department of Education, the U.S. Department of Energy, and to several state governments and universities. He has served on National Academies committees on international benchmarking of U.S. science and manufacturing modernization, and chaired the American Association for the Advancement of Science's Committee on Science, Engineering, and Public Policy. He chairs the National Science Foundation's Advisory Committee for Social, Behavioral, and Economic Sciences, and is a member of its Advisory Committee on the Government Performance and Results Act, the Institute of Medicine's Committee for Assessment of Excellence Programs at NIH, and the National Research Council's Transportation Research Board's Research and Technology Coordinating Committee.

Acknowledgments

We express a special thanks to Connie Chang, ATP supervisory economist and the project manager, for her meticulous attention to all aspects of the endeavor, inspiration, and valuable comments and advice throughout. We also thank Stephanie Shipp, Director of ATP's Economic Assessment Office; Jeanne Powell, Shawn McKay, and Prasad Gupte, also of the Economic Assessment Office; Barbara Cuthill, Group Leader for Information Technology and Applications, ATP; and John Hewes, former ATP program manager, for their careful reviews and many helpful comments. Elissa Sobolewski, Acting ATP Deputy Director, and Brian Belanger, former ATP Deputy Director, reviewed the final version of this report. We acknowledge the central role played by the many evaluators who conducted the body of 45 evaluation studies from which our report draws. We also acknowledge the excellent editorial assistance provided by Sheila Turnage, technical writer.

Abstract

Evaluation is an essential component of publicly funded R&D programs, both in support of program management and public policy. The Advanced Technology Program (ATP) has emerged over its first decade as a leader in evaluation, engaging nationally prominent evaluators to apply new and existing methods in building an analytical and empirical basis for ATP's operations and performance.

This report draws from a body of 45 studies commissioned by ATP between 1990 and 2000 and analyzes the methods and techniques used and examines the findings of those studies. These studies have increased understanding not only of ATP but also of the dynamics of innovation systems and the relationships between public and private sector funding of R&D. The findings examined are organized around five major themes: firm/industry effects, collaboration effects, spillover effects, interfaces and comparisons with other programs, and measures of overall program performance.

The extensive toolkit of evaluation methods presented in the report illustrates how those methods can be used to answer a variety of stakeholder questions. Methods include survey, descriptive and economic case study, bibliometrics, historical tracing, econometrics, expert judgment, social network analysis, cost index, and a composite performance rating system constructed from indicator metrics. Additionally, the use of analytical and conceptual modeling to explore a program's underlying relationships and process dynamics is considered. The political economy of ATP is discussed, and an evaluation framework and an overview of evaluation best practices are provided.

The report integrates and condenses a large body of related research and thus provides ATP with a convenient reference work, toolkit, and planning guide. For those administrators of other programs, public policy makers, and evaluators, the report also serves as an evaluation toolkit by providing a logical framework

for program evaluation, illustrating the use of evaluation methods and techniques, providing an overview of evaluation principles and practices, organizing a body of knowledge on how public-private partnership programs function, and contributing to an understanding of what evaluation is and how it is practiced in the field of R&D.

Keywords: Advanced Technology Program, assessment, economic evaluation, evaluation methods, impact analysis, logic models, public policy, public-private partnership program, R&D, spillovers, technology

This research was conducted between July 2001 and December 2002.

Executive Summary

Evaluation is a powerful tool for decision makers, but only if it is correctly structured, managed, and applied. Among federal and state agencies interested in science and technology, the Advanced Technology Program (ATP), located in the U.S. Department of Commerce's National Institute of Standards and Technology (NIST), has emerged as a leader in the effective use and development of evaluative tools. Over its first decade, ATP followed a multi-faceted approach to evaluation, providing a mosaic of findings about how the program works and its impacts.

This report assembles a large body of ATP's evaluation studies into a coherent framework, making the studies more accessible and understandable to a diverse audience. An expected benefit is better utilization of past evaluation and increased efficiency and effectiveness in planning future evaluation. In effect, the report provides an evaluation "toolkit" for ATP that will also be useful to others who operate public technology programs. The toolkit provides an evaluation framework; a directory of evaluation methods, tools, techniques, principles, explanatory information, and best practices; an account of ATP's use of evaluation models and methods over its first decade as revealed in a body of 45 selected studies; a cross-cutting compendium of findings; and recommendations for future work. The report addresses the science, craft, and art of evaluation in the context of ATP. It shows how a program established in a climate of political and conceptual debate can use evaluation techniques to answer questions about its fundamental rationale, design features, and economic impacts.

Part I provides a general framework for evaluation, discussing evaluation fundamentals and methods, best practices, and an evaluation logic model to describe ATP's evaluation program. Part II demonstrates the use of evaluation methods by drawing on ATP evaluation studies. Part III presents the emerging body of knowledge from studies of ATP—knowledge about firm behavior, collaboration, spillover effects, interfaces with state and international technology programs, ATP's performance at large, and knowledge about evaluation

itself. Part III also presents the authors' conclusions and recommendations. Other features include a glossary of terms, methods bibliography, and a quick reference guide to evaluation models and methods, ATP studies cited, and study findings on program impacts.

Evaluation Underpinnings

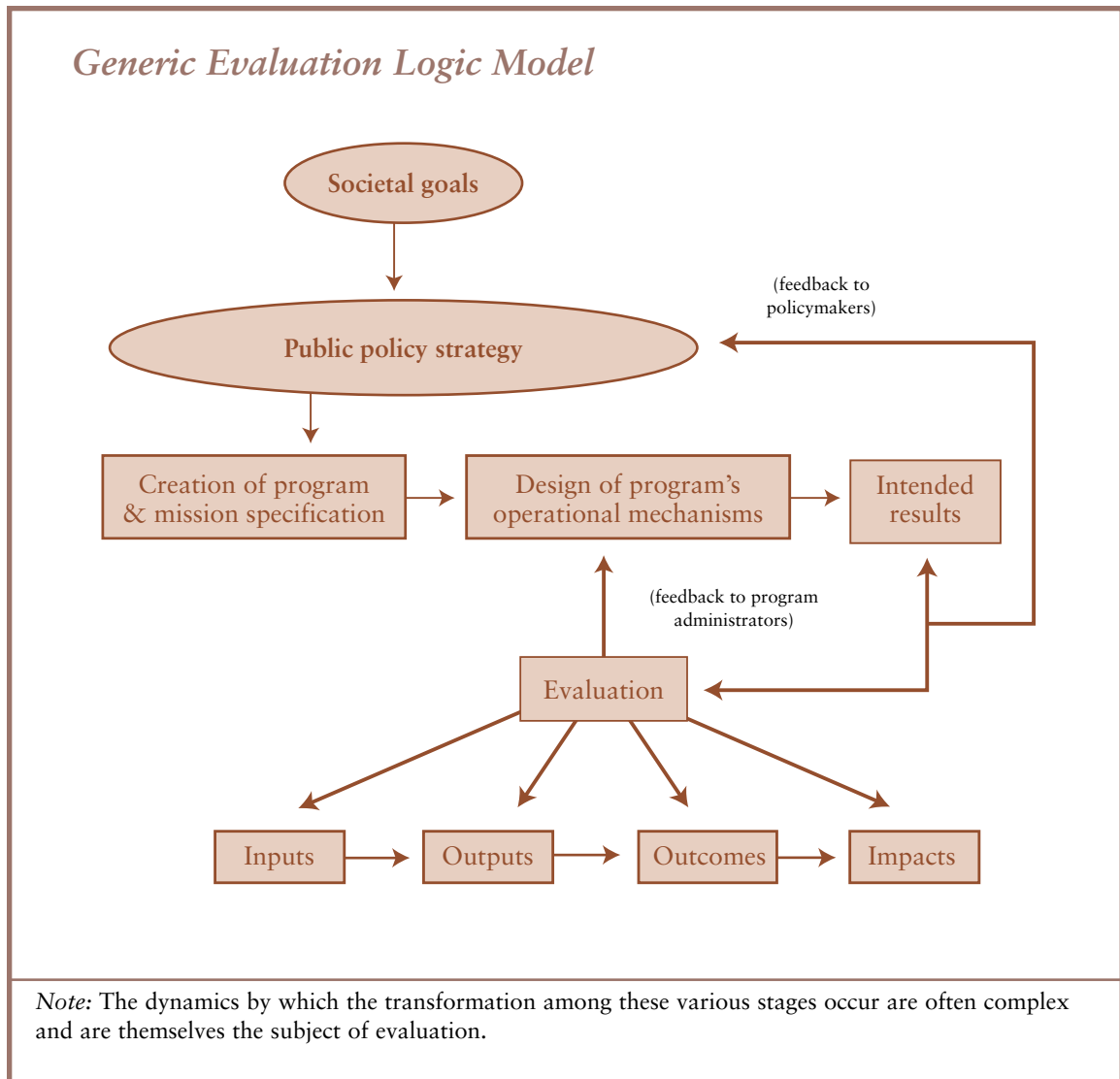
For a public sector program like ATP, evaluation seeks to measure change and to determine if the change is attributable to program intervention. An effective evaluation program should investigate change in terms of a program's mission-driven goals, and should compare its findings against intended results. As a point of departure, this report starts with a generic logic model of program evaluation, depicted below, and fleshes out the model using ATP as illustration.

The report summarizes the major analytical themes economists and others use to explain the rationale for ATP: (1) Global economic competition is increasingly driven by technological advance; (2) enabling technologies tend to generate large spillovers; (3) high level of technical risks contribute to an R&D funding gap in the private sector; (4) many advanced technological development projects require multi-disciplinary and multi-organizational collaborative efforts; and (5) the nation's capacity for economic competitiveness and prosperity depends in large part on its innovative capacity, which can be strengthened through public-private partnerships.

ATP's evaluation program has emphasized modeling its underlying program theory—exploring basic concepts, developing underlying causal maps, developing and refining analysis models, and investigating paths connecting program activities to intended impacts. Findings from ATP's studies in turn have shaped its program and evaluation design in numerous ways.

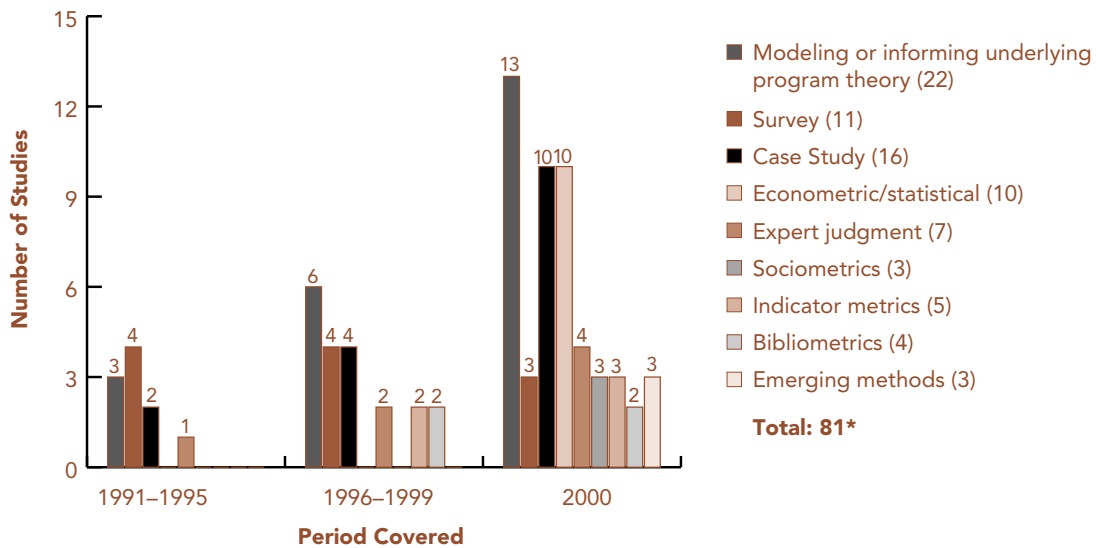
Multi-Faceted Methodological Approach

Evaluators use a variety of evaluative methods, each with its advantages, disadvantages, and specialized purposes. In a multi-faceted approach, like that used by ATP, methods are chosen for their appropriateness to the question at hand, to cost and administrative feasibility, and to a purposeful mixture of methodological



paradigms. Three dominant characteristics of ATP’s evaluation program have been the care with which methods and techniques have been matched to the questions being posed, the evolution toward more rigorous tests of causal relationships between ATP activities and observed outcomes, and the development of new tools when existing tools were not up to the task. The result is an extensive and increasingly sophisticated toolkit of methodologies available to evaluate ATP and other technology programs. The figure below depicts the major methods used by ATP over its first decade, and the changing intensity of their use over time. For

ATP's Evolving Use of Methods Over its First Decade



*These 81 methods are employed in the 45 ATP studies commissioned between 1990 and 2000 that are examined in this report.

example, the use of case study increased from 2 to 4 to 10 between the period 1991-1995, 1996-1999, and 2000, respectively.

An Emerging Body of Findings

Throughout its history ATP has had to demonstrate that its operations added to, rather than displaced, the actions of the private sector in assembling the capital necessary to nurture high-risk, enabling technological innovations. It also has had to prove that ATP assistance produces economic benefits that extend beyond the direct recipients of ATP awards to generate broad benefits for the nation.

Evaluation has provided descriptive and analytical information on program recipients and program outputs to ATP and NIST officials, to key executive and congressional decision makers, and to other stakeholders, including the

general public. The body of evaluative work conducted over ATP's first decade has answered central questions arising from ATP's mission.

A crosscutting analysis of the evaluation studies reviewed revealed much information that bears directly on ATP's mission-driven goals. This analysis is organized around the following major themes: (1) firm/industry effects, (2) collaboration effects, (3) spillover effects, (4) interfaces and comparisons with other programs, and (5) measures of overall ATP performance, including portfolio analysis, social returns on investment, and impacts on competitiveness. Taken as a body of work, these studies have also contributed to enhanced understanding of the dynamics of the U.S. innovation system, particularly the characteristics of productive R&D relationships between the public and private sectors.

Firm/Industry Effects

Findings on private firm effects, drawn from 13 studies, indicated that ATP substantially expanded and enhanced the R&D activities of the companies examined. The studies provided a growing body of evidence that ATP funding is complementary to, not a substitute for, private sources of R&D funds. They indicated that ATP funding leverages and accelerates R&D, refocuses R&D on more technically challenging problems and enabling platforms of technologies, and fills a significant funding gap. One study concluded that the median time-savings per project was three years and the median economic value to the company per year saved was \$5–\$6 million. Two other studies estimated significant program-induced increases in patenting by ATP award recipients, indicating a positive impact of ATP on firm research productivity. With regard to the participation of small firms in ATP, the research showed robust participation rates and strong project performance relative to companies of larger size.

Collaboration Effects

The report drew findings on collaboration from 10 studies. One recurring conclusion was that there are high rates of collaboration in ATP projects, including formal joint venture members and extending strongly to single-applicant companies. For example, 84% of the first 50 completed projects entailed collaborative relationships, ranging from R&D partnerships with other firms, universities, and non-profit labs, to alliances with other firms to pursue commercialization. These

studies found that ATP successfully encouraged applicants to propose projects entailing collaboration, frequently with entirely new partners. Collaborations of firms with universities was a topic of several of the studies on collaboration. Findings were that collaborations with universities were frequent and that they enhanced the research capabilities of the firms and provided an avenue of knowledge diffusion from and through the universities.

Interestingly, studies also found that the collaborations were frequently fluid, with changes among collaborators occurring during a project's life cycle. Some of these changes may be positive, keeping true to ATP criteria, while others may represent deviations, such as the loss of key participants or a retreat from the more challenging research goals, requiring ATP managerial intervention. The studies suggest that by monitoring projects throughout their lives, ATP is able to respond to and manage change.

According to a study of joint venture participants, ATP contributes to joint venture success by: (1) accelerating the development of high risk technologies, (2) increasing project stability, (3) getting projects through particularly difficult periods in their life cycles, (4) overcoming barriers to collaboration, and (5) increasing up-front planning. Project participants identified specific benefits (particularly a positive effect on creativity), and costs (primarily increased administrative burden) associated with collaboration. Almost all project participants involved in collaborative arrangements indicated that their experience with ATP has stimulated them to plan additional collaborations. Among factors important to the success of collaborative relationships, the study corroborated other work that found establishing an environment of trust to be critical.

Spillover Effects

The concept of economic spillovers occupies a central place in the case for a public sector program like ATP and has helped shape many of ATP's program design features. Findings from 10 of the studies increased understanding of ATP's success in generating spillovers. The studies provided considerable evidence that ATP-funded projects generate outputs—publications, patents, patent citations, collaborative linkages, and products—that will potentially lead to knowledge and market spillovers. The potential of network spillovers was also identified, but not yet measured.

One study concluded that ATP selects projects with attributes conducive to generating large knowledge spillover effects. Those attributes included linkages to other organizations, and a positive attitude of award winners toward information sharing. Several studies concluded that the degree to which a funded company is embedded in organizational networks is a major factor in knowledge spillover potential.

Study results also indicated that ATP selects projects whose firms have more extensive ties to other businesses, and, hence, are better positioned to realize commercial success and related market spillovers. In the studies examined, quantitative estimation of the economic value of spillover benefits was limited to market spillovers. Where estimated, market spillover benefits appeared large, and far in excess of private benefits. Among the body of work examined, none of the studies estimated the economic value of both market spillovers and knowledge spillovers.

Interfaces with State Programs and Comparison with Counterpart Programs Abroad

This report draws on five studies for data on the interactions between ATP and state programs and on ATP-counterpart programs in other countries. One study's major conclusion, based on analysis of existing state technology programs, was that state technology programs span the research and development continuum, but cluster around the downstream applied/commercialization segment rather than the upstream research segment of the continuum. The study found that ATP, in contrast, centers its activities on technical challenges, supporting work primarily in the concept and development phases. A collection of case studies highlighted the possibilities of firms combining support from both ATP and state government programs, and illustrated how ATP and the state programs can augment one another. With regard to counterpart programs in other countries, one study offers a framework for standardizing the comparison of ATP with foreign counterpart programs. This systematic approach has helped ATP meet its mandated requirement to test for eligibility of foreign-owned companies for awards and has allowed ATP to learn from the experience of other programs.

Overall ATP Performance

Thirteen studies provided findings on ATP's impact on national industrial competitiveness and the national capacity to innovate, its ability to deal appropriately with failed projects, its contribution to social benefits, and its overall effectiveness. Prospective case studies provided evidence that the benefits of the program far exceed its costs. These studies collectively attributed to ATP more than \$15 billion in expected present value social benefits from just a few projects, much greater than the total amount spent by the program. As expected, not all of the projects are strong performers, but several years after project end an estimated 16% of completed ATP-funded projects showed strong progress toward creating and disseminating knowledge and commercializing projects and processes, and another 26% also showed substantial progress. Five to 6% of all funded projects failed to start or were terminated prior to completion for a variety of reasons. In a major independent assessment, the National Research Council concluded that ATP is effectively meeting its legislative goals.

Recommendations for Future Directions

This report concludes by proposing future directions for ATP's evaluation program, taking into account stakeholder questions, gaps in coverage, past accomplishments, and promising research opportunities. The authors provide 10 recommendations, in no particular rank order, as follows:

- Increase retrospective, market-data-based analyses
 - Incorporate both direct- and indirect-path analysis in benefit-cost case study, including estimates of both market and knowledge spillovers
 - Continue status reports of completed projects and, on a sample basis, repeat them further out in time
 - Update information on state and foreign counterpart programs
 - Further develop several of the promising new evaluation techniques
 - Deepen analysis of knowledge spillovers beyond patent-only-based studies
 - Identify and address new questions that arise as ATP is modified
 - Pursue analysis of failures and successes
 - Continue an effective mix of in-house and external evaluation studies
 - Take greater advantage of evaluation results in decision-making processes
-

In sum, evaluation has provided an objective analytical and empirical basis for assessing ATP's operations and impacts during its first decade of operations. Cumulatively, these evaluations highlight the value of applying multiple evaluation methods to complex problems, building a body of credible evidence over time that ATP is achieving its objectives.

Main topics covered in the report are highlighted below.

Highlights of Main Topics

MODELS AND METHODS

- ✓ ATP's evaluation logic model
- ✓ Generic treatment of evaluation methods: list, definitions, examples of use
- ✓ Chronological listing of 45 ATP evaluation studies commissioned (1990–2000), with principal and secondary methods used by ATP
- ✓ ATP's use of evaluation methods*
 - Modeling or informing underlying program theory (22 supporting studies covered in the report)
 - Survey method (8 of 11 supporting studies covered in the report)
 - Case study method (10 of 16 supporting studies covered in the report)
 - Econometric/statistical methods (8 of 10 supporting studies covered in the report)
 - Expert judgment method (5 of 7 supporting studies covered in the report)
 - Sociometrics (3 supporting studies covered in the report)
 - Indicator metrics (5 supporting studies covered in the report)
 - Bibliometrics method (3 of 4 supporting studies covered in the report)
 - Emerging methods (3 supporting studies covered in the report)
 - Cost index method
 - Social network analysis/fuzzy logic
 - Composite performance rating system

*Some studies used multiple methods. Not all studies are referenced in each chapter.

continued on next page

Highlights of Main Topics (Cont'd)

CROSSCUTTING FINDINGS

- ✓ Impact on private firms
 - Financing gap
 - Halo effect
 - Acceleration
 - Firm productivity
 - Small firm participation
 - Commercialization, company growth, and private returns
- ✓ Collaboration
 - Activity, structure, formation, and attribution
 - Changes in relationships
 - University representation and roles
 - Determinants of success
 - Benefits and costs
- ✓ Spillover effects
 - Market spillovers
 - Knowledge spillovers
- ✓ State and foreign programs
 - State program interfaces
 - Foreign program comparisons
- ✓ Overall ATP performance measures
 - ATP's contribution
 - Improving competitiveness of the United States and its businesses
 - Fostering the national capacity to innovate
 - Dealing with failed projects
 - Measuring progress, social benefits, and overall effectiveness

Acronyms and Abbreviations

ATP	Advanced Technology Program
BRS	Business Reporting System
CAFE	Corporate average fuel economy
CCAR	Closed-cycle air refrigeration
CPRS	Composite performance rating system
DCCT	Diabetes Control and Complication Trial
DDS	Digital data storage
DSG	Diamond Semiconductor Group
EAO	Economic Assessment Office
FCM	Flow-control machining
GAO	General Accounting Office
GDP	Gross domestic product
GPRA	Government Performance and Results Act of 1993
IIH	Information Infrastructure for Healthcare
IRR	Internal rate of return
MEMS	Micro-electromechanical systems
NAS	National Academy of Sciences
NBER	National Bureau of Economic Research
NSIC	National Storage Industry Consortium
NIST	National Institute of Standards and Technology
NPV	Net present value
NRC	National Research Council
OIG	Office of Inspector General
PDM	Pit depth modulation
PWB	Printed wiring board

QALY	Quality Adjusted Life Year
REMI	Regional Economic Modeling, Inc.
RTI	Research Triangle Institute
STEP	Science, Technology, and Economic Policy
SWAT	Short-wavelength sources for optical recording



Introduction

Evaluation is a powerful tool for policy and decision making, but only if it is correctly structured, managed, and applied. At the technical level, this means having a clear set of objectives, a logical framework, and valid methods and findings. At the program level, this means having knowledgeable evaluators and program administrators who can work with evaluators to structure relevant questions, relate general findings to specific agency settings, and communicate effectively with diverse audiences.

Evaluation involves methodological science and craft and organizational art. Evaluation involves the selection and implementation of systematic, valid, and appropriate methodologies. Evaluation also involves the organizational establishment, management, deployment, and dissemination of a portfolio of studies and associated findings that provide defensible and relevant information to decision makers.

Since at least the 1960s, program evaluation and its close companion, policy analysis, have become institutionalized aspects of congressional oversight and agency management of federal programs.¹ Institutionalization, however, is not synonymous with acceptance, quality, credibility, or impact.

Many federal and state agencies struggle to implement evaluation programs, which may be mandated but may, in fact, be unwelcome appendages to program operations and/or organizational decision making. Evaluation efforts are some-

¹Joseph S. Wholey, John W. Scanlon, Hugh G. Fukumoto, and Leona M. Vogt, *Federal Evaluation Policy, Analyzing the Effects of Public Programs* (Washington, DC: The Urban Institute, 1970); A. Meltzner, *Policy Analysis in the Bureaucracy* (Berkeley, CA: University of California Press, 1976).

2 / A Toolkit for Evaluating Public R&D Investment

times undertaken only when required by outside forces, such as legislation. These realities make the evaluation program of one of the nation's public-private partnership programs, the Advanced Technology Program (ATP), all the more striking. The design and implementation of its evaluation program offers a creative laboratory for learning more about evaluation, particularly in the field of science and technology.

This report addresses the science, craft, and art of evaluation in the context of ATP, a program within the U.S. Department of Commerce, National Institute of Standards and Technology (NIST). In terms of science and craft, this report describes the evolving set of methodological techniques ATP has used to monitor and assess programmatic impacts, and reports methodological and empirical advances generated by ATP. The report assembles a large body of past work into a coherent framework, making it more accessible to a diverse audience.

In terms of art, the report describes the creation and evolution of ATP's evaluation program. It describes how the program has used evaluation techniques to answer questions directed at its fundamental rationale, design features, and economic impacts.

This report also highlights ways in which early challenges to the program led to an evaluation program noted for its methodological variety, recourse to nationally prominent scholars, and a distinctive emphasis on disseminating its findings in peer-reviewed literature and in policy-relevant presentations for agency officials and political constituencies. Finally, the report shows that in politically contested arenas, methodological rigor and empirically grounded findings are necessary but not sufficient to protect or advance a program. For even as ATP drew national recognition in the evaluation community for its systematic, and rigorous evaluation program, congressional critics of the program repeatedly charged that the program lacked adequate evidence of its impacts.

The report is presented in three major parts, plus this introductory section. The remainder of the introduction presents the political underpinnings of ATP and discusses further the role of evaluation. Part I provides a general framework for evaluation, and in the context of the framework, discusses evaluation

fundamentals and methods, best practices, and ATP's evaluation program. Part II demonstrates the use of evaluation methods by citing, both through direct quotation and paraphrase, the contents of a selection of ATP evaluation studies. Part III presents the emerging body of knowledge from the studies of ATP over the past decade—knowledge about evaluation, firm behavior, and ATP's performance. Part III also presents conclusions and recommendations, summarizing key points and identifying remaining questions, issues, obstacles, and challenges, and promising opportunities to learn more through evaluation.

The Political Economy of the Advanced Technology Program

ATP is an outgrowth of a national policy dialogue directed at redressing systematic gaps in the market settings that link scientific advances to technological innovation. ATP was designed to fill those gaps by a project selection and funding process that fosters and enhances new and intensified modes of collaboration among private, not-for-profit, and public sector organizations engaged in high-risk research, all with the objective of accelerating the development and commercialization by U.S. firms of enabling technologies.

ATP was established in 1988 under Title V (Technology Competitiveness Act), Subtitle B of the Omnibus Trade and Competitiveness Act (P.L. 100-418), and received its first appropriations in fiscal year 1990. The program is a response both to a specific period in United States international economic competitiveness and to a longer-term historic perspective of the federal government's contribution to the development and commercialization of potentially significant technological advances of a broadly enabling nature.

The specific historic backdrop for ATP's creation was a pervasive concern in the United States throughout much of the 1970s and 1980s, documented by a variety of key economic indicators, that the nation's slow rate of economic growth and worsening international trade balance were attributable in part to its loss of technological competitiveness. The U.S. innovation system was widely seen as exhibiting structural flaws. Points of concern included a loss of international and domestic markets in technology-intensive products, a failure to gain

4 / *A Toolkit for Evaluating Public R&D Investment*

market position for products U.S. firms had helped research and develop, and a faltering standing in R&D “races” to exploit the commercial significance of emerging scientific advances.²

The flaws were seen as products of an emphasis in U.S. policies on scientific leadership, mission-directed R&D, and breakthrough discoveries, paired with a lack of attention to technology development and deployment. Critics believed an overemphasis in these areas was made at the expense of diffusion-oriented strategies and programs that would assist U.S. firms and laboratories to gain technologically and economically when their scientific advances were converted into new and improved products and processes.³ Critics argued that the United States needed an “innovation policy” rather than a “science-only policy.”

These flaws were described in both absolute and relative terms. In an absolute sense, the flaws were seen as reflecting specific forms of market failure, primarily the lack of adequate incentives for private firms and private capital markets to fund “high-risk, high-return research on broadly enabling, precompetitive technologies for which appropriability was at issue.”⁴ In a relative sense, the challenge to the United States was that major economic competitors, most notably Japan and some members of the European community, were held to be far more willing to use national funds to support the development and commercialization of civilian technologies, particularly those seen as providing “first mover” advantages in new, strategic, or large commercial markets.⁵ Lack of

²U.S. Congress, Office of Technology Assessment, *Making Things Better: Competing in Manufacturing*, OTA-ITE-443 (Washington, DC: U.S. Government Printing Office, 1990).

³H. Ergas, “The Importance of Technology Policy.” In P. Dasgupta and P. Stoneman, eds., *Economic Policy and Technological Performance* (New York: Cambridge University Press, 1987), pp. 51–96; R. Florida and M. Kenney, *The Breakthrough Illusion* (New York: Basic Books, 1990).

⁴L. Branscomb and G. Parker, “Funding Civilian and Dual-Use Industrial Technology.” In L. Branscomb, ed., *Empowering Technology* (Cambridge, MA: MIT Press, 1993), pp. 64–102.

⁵U.S. Congress, Office of Technology Assessment, *Competing Economies: America, Europe, and the Pacific Rim*, OTA-ITE-498 (Washington, DC: U.S. Government Printing Office, 1991).

comparable government programs in the United States was held to place U.S. firms at a disadvantage, because they had to bear the full burden of supporting costly and risky R&D projects with questionable appropriability of profits.

The design of ATP also drew upon emerging perspectives about the character of competition among firms. Knowledgeable observers no longer saw firms that conducted business in the same product lines as engaged exclusively in Darwinian struggles for survival, and they no longer identified firms engaged in buyer-seller relationships only as seeking to maximize profits/minimize costs from one-time transactions. Instead, both theory and increased documentation of business practice pointed to numerous forms of collaboration between such pairs.⁶

In particular, with respect to R&D, it became increasingly evident that firms sharing a common interest in selected “generic” technologies could increase their competitiveness through collaboration. IBM’s agreement in 1992 to join with Toshiba of Japan and Siemens of Germany to develop memory chips illustrates what was seen as a new way of doing business, especially in R&D-intensive and technology-driven sectors.

Collaboration spread the capital costs of large-scale R&D projects among investors. It offered a way to explore otherwise out-of-reach technological frontiers, both as part of a firm’s offensive strategy of finding new products and markets, and as part of a defensive strategy of better predicting and understanding the rise of disruptive technologies that threatened a firm’s corebusinesses.

The growing acceptance of inter-firm and inter-sector collaboration, especially in pre-competitive, generic (or enabling) R&D, was reflected in several legislative changes, such as the National Cooperative Research Act of 1984,⁷ that relaxed antitrust bars to collaborative R&D programs. The evolving framework also saw a relaxation of the analytical and ideological categories that had previously been

⁶D. Mowery, ed., *International Collaborative Ventures in U.S. Manufacturing* (Cambridge, MA: Ballinger, 1988).

⁷A. Link and L. Bauer, *Cooperative Research in U.S. Manufacturing* (Lexington, MA: Lexington Books, 1989).

6 / *A Toolkit for Evaluating Public R&D Investment*

used to define the boundaries of federal government and private sector roles in supporting R&D directed at civilian-oriented technology.⁸

Under the framework that dominated policy thinking following World War II, the federal government had responsibility for funding basic research and R&D related to mission-oriented national objectives such as defense, while the private sector had responsibility for R&D directed at civilian-oriented products and processes. Reinforcing the hold of this paradigm was the checkered political and technological history of efforts by the federal government to promote specific technological innovations.^{9,10} Over time, however, the paradigm began to yield to a more complex but nuanced appreciation of both the overlap and holes in these seemingly fixed boundary markers.¹¹ In particular, the emerging model emphasized government support of technically risky projects that had prospects for technological and commercial importance; significant amounts of industry cost sharing; sunset requirements for government funding; selection of projects for funding in a fair and open competitive process free of political influence; large spillover effects; and collaboration with other firms and organizations.

Despite growing support for public-private partnerships emphasizing high-risk, enabling technology, opposition continued.¹² Opposition embodied several

⁸B. Smith, *American Science Policy Since World War II* (Washington, DC: Brookings Institution, 1990).

⁹See L. Cohen and R. Noll, *The Technological Pork Barrel* (Washington, DC: Brookings Institution, 1991) for examples of inefficient government funding projects and a discussion of factors behind the inefficiencies.

¹⁰See Charles W. Wessner, ed., *The Advanced Technology Program: Challenges and Opportunities* (Washington, DC: National Academy Press, 1999), pp. 1–4, for a summary description of the instrumental role played by the federal government since its earliest history in the development of new production techniques and technologies. Examples include a government contract in 1798 with Eli Whitney to help lay the foundation for the U.S. machine tool industry, government funding to demonstrate the feasibility of Samuel Morse's telegraph, and government support for the development of the radio.

¹¹National Academy of Sciences, *The Government Role in Civilian Technology* (Washington, DC: National Academy Press, 1992).

¹²C. Hill, "The Advanced Technology Program: Opportunities for Advancement." In L. Branscomb and J. Keller, eds., *Investing in Innovation* (Cambridge, MA: MIT Press, 1998), pp. 153–173; P. Hallacher, "Effects of Policy Subsystem Structure on Policymaking: The Case of the Advanced Technology Program and the Manufacturing Extension Partnership Program," Ph.D. Dissertation, Pennsylvania State University, 2000.

reinforcing ideological and theoretical positions.¹³ It started with the premise that the national government contributes most effectively to long-term economic growth and technological innovation when it adopts a minimalist approach to involvement in the economy, focusing its activities on the enforcement of private contracts, the provision of stable monetary aggregates, the maintenance of certain and low taxes, and the provision of a small, carefully delimited set of selected public goods.¹⁴ Opposition was also based on propositions that public sector support cannot increase the rate of commercially productive R&D because of the phenomena of moral hazard, adverse selection, rent-seeking behaviors by firms applying for support, and bureaucratic entrepreneurship on the part of agency managers pushing for overly rapid development of untested and economically problematic technologies.¹⁵

Opponents of federal domestic technology development programs also point to further defects. First, they believe the programs may simply substitute public-sector R&D dollars for private-sector R&D dollars. Second, they say these programs may have untoward distributive effects, unfairly benefiting some U.S. firms and/or industries at the expense of others, such as large firms at the expense

¹³Opposition extended to earlier, contemporary, and subsequent similar programs that involved efforts by the federal government to stimulate acceleration of commercially oriented technological innovation through targeted selection of industries, firms, technologies, or projects. See B. Smith, *American Science Policy Since World War II*, 1990.

¹⁴Federal government support of basic research rests on a broad-based, bipartisan political consensus that it is a public good that would not be adequately provided by the actions of the private sector alone, and that it contributes in significant ways to national objectives in defense, health, space, and economic competitiveness.

¹⁵“Moral hazard” refers to the actions by individuals or firms to increase their risk taking behavior in response to the existence of insurance or other forms of compensation for costly outcomes. Thus, while ATP purposefully underwrites a portion of an R&D project’s risk to encourage private firms to take on more challenging—hence riskier—technical problems, the phenomenon of moral hazard raises the question of whether or not there is a side effect of promoting inefficient resource use. “Adverse selection” refers to a problem arising from asymmetry in the quality of information possessed by the applicant firms and ATP about the riskiness of prospective R&D projects. If ATP knows less about the risks than the companies, it will be at a disadvantage in its selection decisions. “Rent-seeking behaviors” refers to efforts directed at acquiring or making permanent a stream of payments, government program, or regulatory arrangement that yields returns to an economic group above that which they would receive under competitive market conditions. Such behavior would occur if awardees attempted continuation of their government awards beyond the scheduled end date or if political efforts were taken to maintain the program even if it were deemed not to produce its intended results.

8 / *A Toolkit for Evaluating Public R&D Investment*

of small firms. Additionally, opponents claim the programs may respond to pressures from bureaucratic and technological interests, and unduly push for rapid deployment of technologies before the technological feasibility or market demand has been demonstrated. They maintain that these programs may fail to select proposals with the highest likelihood of achieving technical success, commercial success, and large spillovers. They also maintain that funding a project creates vested interests that causes its funding to become self-perpetuating. Finally, opponents say the claims about the productivity of collaborative efforts, and the federal government's role in fostering them may be challenged as unproven.

The catalog of arguments on behalf of the benefits of public-private partnerships in general, and ATP in particular as well as the catalog of possible flaws are both mixes of normative perspectives on the role of the public sector, and theoretically and empirically testable propositions. The competing catalogs, in effect, describe domains of debate and decision making in which evaluation may contribute to improved public policy.

The Role of Evaluation

Evaluation has a recognized and well-understood role in the operations of most public sector agencies. As phrased by Mark, Henry, and Julnes:¹⁶

Evaluation assists sense making about policies and programs through the conduct of systematic inquiry that describes and explains the policies' and programs' operations, effects, justifications, and social implications. The ultimate goal of evaluation is social betterment, to which evaluation can contribute by assisting democratic institutions to better select, oversee, improve, and make sense of social programs and policies. (p. 3)

Evaluation has served all these purposes for ATP: to oversee, improve, and make sense of the program—both within and outside of NIST. ATP's evaluations have helped program administrators revise and refine the program so that the program is harmonized with legislative intent. Evaluation has been used to respond to

¹⁶M. Mark, G. Henry, and G. Julnes, *Evaluation: An Integrated Framework for Understanding, Guiding, and Improving Policies and Programs* (San Francisco: Jossey-Bass, 2000).

Congressional and OMB questions about program characteristics and impacts. ATP's evaluation program has had a broader span of coverage than typically encountered across federal agencies. Its studies have ranged from monitoring and surveying the characteristics of early grantees to sophisticated theoretical and econometric efforts designed to tease out differences between the performance of ATP awardees and non-awardees. In pursuit of these multiple objectives, ATP has strategically employed an evolving and broad set of methodologies.

Early ATP evaluation activities focused on measuring progress, generating information about awardees, and projecting economic impact. In the context of political opposition to ATP, the program's evaluation agenda has often been shaped in response to ongoing challenges, such as whether ATP awards were being used by firms to displace private sector funds. Over time, as some funded projects have had time to progress through technical research and commercialization stages, ATP's evaluation program has centered more on measuring impacts. These include both private impacts garnered directly by the firms participating in ATP projects, and spillover impacts that are one of the justifications for the program.

Finally, evaluation has served as a test of two of ATP's core and linked premises; that is, that a federal agency program can strategically and selectively generate additional and innovative modes of R&D collaboration between and among organizations, and that this collaboration among organizations will accelerate the rate of technological innovation in the nation's economy. ATP's evaluations have provided new insights into the characteristics of workable collaborations among various sectors, the character of gaps between scientific discovery and product development, the strategies and behaviors of firms as they seek to close or bridge these gaps, and the forms and networks through which knowledge spillovers diffuse within an economy.
