



## About FAS

The **Federation of American Scientists (FAS)** is an independent, nonpartisan think tank that brings together members of the science and policy communities to collaborate on mitigating global catastrophic threats. Founded in November 1945 as the Federation of Atomic Scientists by scientists who built the first atomic bombs during the Manhattan Project, FAS is devoted to the belief that scientists, engineers, and other technically trained people have the ethical obligation to ensure that the technological fruits of their intellect and labor are applied to the benefit of humankind. In 1946, FAS rebranded as the Federation of American Scientists to broaden its focus to prevent global catastrophes.

FAS can be reached at 1150 18th St. NW, Suite 1000, Washington, DC, 20036, [fas@fas.org](mailto:fas@fas.org), or through [fas.org](https://fas.org).

COPYRIGHT © FEDERATION OF AMERICAN SCIENTISTS, 2025. ALL RIGHTS RESERVED.

## Author

---

**Leya Mohsin** is the Manager, Government Capacity and Evidence at the Federation of American Scientists. In her role, she provides operational and project management support across FAS's government capacity portfolio, including its work on talent, digital services, and data policy. She has a particular focus on advancing the generation, synthesis, and uptake of research evidence into federal policymaking.

Prior to her work at FAS, Leya was part of the Evidence and Evaluation team at Arnold Ventures, where she supported the development and management of grant-making in rigorous impact evaluation and evidence-based policy. While her work today looks across a wide range of social policy issues, her background is in education, having started her career as a high school math teacher. Leya holds a Master's degree in Education Policy and Leadership from American University, and a Bachelor's degree in economics, policy studies, and civic leadership from Rice University.

## Acknowledgments

---

The development of this report was made possible by generous support from the Alfred P. Sloan Foundation and Coefficient Giving. We are also grateful to the 50+ participants and speakers at our February 2026 workshop, whose observations and insights paved the way for this agenda, as well as Dr. Erica Goldman, Daniel Correa, and Dr. Michael Stebbins for their leadership on this initiative.

## Contents

INTRODUCTION.....	1
METHODS.....	2
THE AGENDA IN CONTEXT: RESEARCH PRIORITIES ACROSS THE ECOSYSTEM.....	3
CONGRESS & OVERSIGHT COMMUNITY.....	4
FEDERAL AGENCIES & OMB.....	5
UNIVERSITIES AND THE RESEARCH ENTERPRISE.....	6
PHILANTHROPY AND THE PRIVATE SECTOR.....	7
THE RESEARCH AGENDA.....	8
FOUNDATIONAL QUESTIONS.....	9
QUESTION 1. HOW CAN THE RETURNS FROM R&D INVESTMENT BE RELIABLY MEASURED AND ATTRIBUTED, ACROSS FINANCIAL AND NON-FINANCIAL OUTCOMES?.....	10
QUESTION 2. WHAT WOULD HAVE HAPPENED IN THE ABSENCE OF A GIVEN R&D INVESTMENT, AND HOW CAN CREDIBLE COUNTERFACTUALS BE CONSTRUCTED FOR PUBLIC FUNDING DECISIONS?.....	11
QUESTION 3. WHAT STANDARDS OF EVIDENCE AND ACCOUNTABILITY ARE APPROPRIATE FOR EVALUATING PUBLICLY FUNDED R&D, AND HOW CAN THEY BE APPLIED CONSISTENTLY ACROSS AGENCIES AND PROGRAMS?.....	12
APPLIED QUESTIONS.....	13
QUESTION 4. HOW SHOULD R&D PORTFOLIOS BE CONSTRUCTED AND EVALUATED AS A WHOLE, AND WHAT ALLOCATION STRATEGIES MAXIMIZE RETURNS ACROSS PROJECTS WITH VARYING RISK, STAGE, AND FIELD?.....	14
QUESTION 5. WHAT IS THE DISTINCT CONTRIBUTION OF PUBLIC R&D INVESTMENT, AND HOW DOES THE CHOICE OF FUNDING INSTRUMENT AFFECT THE RETURNS IT GENERATES?.....	15
QUESTION 6. HOW DO THE BENEFITS OF R&D INVESTMENT SPILL OVER BEYOND THE DIRECT RECIPIENTS, AND HOW CAN THESE INDIRECT EFFECTS BE MEASURED?.....	17
QUESTION 7. TO WHAT EXTENT DOES PUBLIC R&D INVESTMENT CATALYZE ADDITIONAL PRIVATE AND PHILANTHROPIC FUNDING, AND WHAT PROGRAM DESIGNS ARE MOST EFFECTIVE AT CROWDING IN CO-INVESTMENT?.....	19
QUESTION 8. HOW EFFICIENTLY IS R&D FUNDING CONVERTED INTO ACTUAL RESEARCH ACTIVITY, AND WHAT REFORMS WOULD REDUCE FRICTION AND INCREASE THE SHARE OF RESOURCES REACHING RESEARCHERS?.....	20
STRUCTURAL QUESTIONS.....	21
QUESTION 9. HOW DO THE RETURNS FROM R&D INVESTMENT UNFOLD OVER TIME, AND WHAT EVALUATION FRAMEWORKS ALLOW FUNDERS WITH DIFFERENT TIME HORIZONS TO ASSESS IMPACT ON THEIR OWN TERMS?.....	22
QUESTION 10. HOW DO POLITICAL AND INSTITUTIONAL INCENTIVES SHAPE R&D INVESTMENT DECISIONS, AND WHAT GOVERNANCE STRUCTURES ALLOW LONG-TERM RESEARCH GOALS TO SURVIVE SHORT ELECTORAL AND BUDGET CYCLES?.....	23
QUESTION 11. WHAT CAN BE LEARNED FROM PAST R&D PROGRAMS – INCLUDING THOSE THAT FAILED – AND HOW SHOULD THAT KNOWLEDGE SYSTEMATICALLY INFORM FUTURE INVESTMENT DECISIONS?.....	24
CONCLUSION.....	25

## Introduction

The United States federal government [invests](#) nearly \$150 billion annually in research and development. Four times (\$937 billion in 2023, [as estimated by NSF](#)) this amount flows from state governments, private foundations, and corporate R&D budgets. These investments are made on the premise that they generate scientific, economic, and social returns. That premise is widely held. However, the supporting evidence generates wildly different estimates depending on the methods and available data.

That evidentiary weakness has always been a limitation. It is now, increasingly, a liability. Federal R&D budgets are subject to scrutiny of an intensity not seen in decades. Longstanding assumptions about the appropriate scale and direction of public investment in science are being actively contested across branches of government. In that environment, the inability to answer basic questions about what research investment produces – for whom, on what timeline, and compared to what alternative – leaves the field without the tools it needs to make its case. This agenda is designed, in part, to begin building those tools.

The literature on the returns to R&D investment has grown substantially over the past three decades. The work began in the 1950s with papers [published by Zvi Griliches](#) estimating the social rate of return to research activity. Economists have since [estimated](#) rates of return to federally funded basic research. Health economists have traced the pathway from NIH appropriations to [pharmaceutical innovation](#) to [reductions in mortality](#). Innovation scholars have [documented](#) knowledge spillovers across firms, sectors, and national borders. This body of work has been useful, and has provided important empirical grounding for sustained public investment in science. But it has also left critical questions underexplored, and has not kept pace with the demands that policymakers, funders, and program administrators now place on evidence and evaluation.

### THE GRILICHES QUESTION: WHAT IS THE RETURN ON INVESTMENT (ROI) OF R&D?

Three gaps are especially consequential. First, the existing literature is concentrated on a narrow band of measurable outcomes – financial returns, patent counts, publication rates, and selected health outcomes – while the full range of value that R&D investment generates remains largely unmeasured. Social benefits, distributional effects, environmental outcomes, and the intrinsic value of scientific knowledge itself are acknowledged in principle but rarely captured in practice. The result is an evaluation vocabulary that systematically understates what investing in public research produces, and that is poorly suited to the constituencies who most need to understand it. Further, the paucity of data on the returns on investment in research makes it difficult to make informed decisions on future funding levels and allocations.

Second, the existing literature has made limited progress on the foundational problem of causal attribution. Demonstrating that an investment generated value is not the same as demonstrating that it generated value that would not otherwise have existed. The counterfactual question – what would have happened in the absence of a given funding decision – is rarely, if ever, answered with the rigor that credible evaluation requires. Where causal methods have been applied, they have typically addressed narrow, well-defined questions; broader claims about program or agency-level returns rest on methodological assumptions that are often implicit and rarely interrogated.

Third, the producers and consumers of evidence on the ROI of R&D are poorly connected. Federal agency program officers, congressional appropriators, budget examiners, state science advisors, philanthropic funders, and university research administrators all face versions of the same underlying questions – about returns, timescales, counterfactuals, spillovers, and accountability – but they ask those questions in different registers, with different data needs, on different timelines, and with different tolerances for methodological uncertainty. Research that is designed without these customers in mind and engaged in the process tends to reach only the audiences already most predisposed to engage with it. This means that key decision-makers are often making hard choices with little-to-no high-quality evidence on hand to support their thinking.

This research agenda is a response to those gaps. It was developed through a structured process of stakeholder engagement that surfaced both the range of outcomes that R&D investment is understood to produce and the specific informational needs of the actors responsible for making, evaluating, and defending funding decisions across the public, private, and philanthropic sectors. The result is a set of eleven research questions, organized from foundational to applied, that together constitute a program of inquiry capable of substantially advancing the field's ability to measure, attribute, and communicate the returns to R&D investment.

The agenda does not promise that all of these questions can be fully answered. Several of the most important ones – including the construction of credible counterfactuals and the valuation of non-pecuniary returns – pose methodological challenges that current tools can address only partially. A transparent research agenda must acknowledge those limits clearly, both to set appropriate expectations and to direct methodological innovation toward the places where it is most needed. What the agenda does promise is that serious, well-designed inquiry into each of these questions would produce findings that are useful – to researchers, to practitioners, and to the policymakers who ultimately determine the scale and direction of public investment in science.

## Methods

The primary vehicle for developing this set of research questions was a one-day workshop hosted in February 2026. The workshop convened academics, policymakers, and policy experts to discuss where the literature stands today and where it should head in the future. It was also intended to help launch a new initiative: the [Pop-Up Journal](#). This initiative, whose first iteration will focus on the ROI of R&D (sometimes called the Griliches Question), will aim to answer many of the questions posed by participants and in this research agenda.

During the workshop, participants engaged in a series of exercises designed to elicit answers to two key questions:

- When we consider return-on-investment from R&D, what are the outcomes we should be measuring?
- What are the research questions that, if answered, are most likely to inform the decision-making of key stakeholders?

To answer the first question, we asked participants to identify a range of possible outcomes that might be important to someone trying to understand the return of their R&D investments.

To answer the second question, we assigned small groups different “personas” to consider – ranging from Congressional staff to federal agency leaders to universities and the private sector. They grappled with not just the types of research questions that are most novel, but the ones that were most likely to be useful to decision and policy makers.

To supplement the workshop, we also conducted a series of informal one-on-one interviews with experts – researchers who might one day seek to answer questions posed in this agenda, and policymakers who might use the evidence generated.



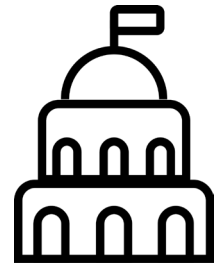
# The Agenda in Context: Research Priorities Across the Ecosystem

The research questions that follow were developed with a broad range of audiences in mind. Different actors in the R&D investment ecosystem – those who fund it, oversee it, execute it, evaluate it, and advocate for it – face different informational needs, even when they are asking related underlying questions. A congressional appropriator and a university vice president for research may both care about accountability, but they care about it for different reasons, on different timelines, and with different consequences if the answer is unsatisfying.

Derived from the “personas” exercise conducted during the workshop, this first section maps the agenda to those differences. For each research question, it identifies the stakeholder communities – including members of Congress and their staff, federal agency program officers and budget examiners, independent oversight and analysis offices, university research administrators, private sector partners, and philanthropic funders – whose core informational needs that question most directly addresses. The intent is not to suggest that other questions are irrelevant to each audience; the full agenda was designed to be useful across the ecosystem. Rather, it is to give each reader a clear entry point – a way of locating themselves in the agenda before engaging with it as a whole.

Readers who recognize their own role in one of these descriptions are encouraged to begin there, then follow the cross-references to the sub-questions and related outcome measures that bear most directly on their work. Those approaching the agenda from a research or funder perspective, rather than as a practitioner, will find the full question set and its methodological discussion in the sections that follow.





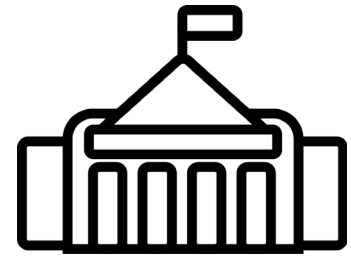
## Congress & Oversight Community

Members of Congress, committee staff, and the analysts and evaluators who support congressional decision-making share a common orientation toward accountability, but they approach it from meaningfully different vantage points. Legislators and their staff are principally concerned with whether federal R&D investments are delivering tangible benefits for constituents and whether agencies can justify their spending decisions when called to do so.

The oversight and analysis community – spanning the Government Accountability Office, the Congressional Budget Office, and the Congressional Research Service – brings a more methodological lens, asking not just whether programs worked but whether the evidence for that conclusion is credible, consistent, and replicable.

For this group, the most pressing research needs cluster around accountability standards, counterfactual construction, and the development of evaluation frameworks that can survive rigorous scrutiny and inform consequential budget decisions.

<b>MEMBERS OF CONGRESS</b>	<p>KEY QUESTION: HOW CAN THE RETURNS FROM R&amp;D INVESTMENT BE RELIABLY MEASURED ACROSS BOTH FINANCIAL AND NON-FINANCIAL OUTCOMES? (<a href="#">JUMP TO PAGE 10</a>)</p> <p>STORYTELLING IS KEY FOR SENATORS, REPRESENTATIVES, AND THEIR STAFF TO ADVOCATE FOR R&amp;D FUNDING. THEY NEED EVIDENCE ABOUT HOW INVESTING IN R&amp;D DELIVERS FOR THEIR CONSTITUENTS IN SPECIFIC COMMUNITIES, NOT JUST TOP-LEVEL NUMBERS.</p>
<b>APPROPRIATIONS COMMITTEE STAFF</b>	<p>KEY QUESTION: WHAT STANDARDS OF EVIDENCE AND ACCOUNTABILITY ARE APPROPRIATE FOR EVALUATING PUBLICLY FUNDED R&amp;D, AND HOW CAN THEY BE APPLIED CONSISTENTLY? (<a href="#">JUMP TO PAGE 12</a>)</p> <p>AN APPROPRIATIONS COMMITTEE HAS TO MAKE HARD DECISIONS ABOUT WHAT TO FUND. CLEAR, CONSISTENT STANDARDS FOR UNDERSTANDING WHEN INVESTING IN R&amp;D “WORKS” WOULD SUPPORT THEIR DECISION-MAKING.</p>
<b>GOVERNMENT ACCOUNTABILITY OFFICE</b>	<p>KEY QUESTION: WHAT STANDARDS OF EVIDENCE AND ACCOUNTABILITY ARE APPROPRIATE FOR EVALUATING PUBLICLY FUNDED R&amp;D, AND HOW CAN THEY BE APPLIED CONSISTENTLY? (<a href="#">JUMP TO PAGE 12</a>)</p> <p>MOST FEDERAL PROGRAMS, INCLUDING THOSE THAT FUND R&amp;D, ARE DESIGNED WITHOUT A CLEAR SET OF MEASURABLE GOALS. HAVING A FRAMEWORK FOR EVALUATING R&amp;D PROGRAMS ACROSS AGENCIES WOULD SUPPORT GAO’S OVERSIGHT OF SCIENCE AGENCIES.</p>
<b>CONGRESSIONAL RESEARCH SERVICE</b>	<p>KEY QUESTION: WHAT IS THE DISTINCT CONTRIBUTION OF PUBLIC R&amp;D INVESTMENT, AND HOW DOES THE CHOICE OF FUNDING INSTRUMENT AFFECT THE RETURNS IT GENERATES? (<a href="#">JUMP TO PAGE 15</a>)</p> <p>TO SUPPORT CONGRESSIONAL DEBATE, CRS NEEDS RIGOROUS, NON-PARTISAN RESEARCH. THE QUESTION THEY NEED ANSWERED TO INFORM LEGISLATORS IS FOUNDATIONAL: WHEN AND WHY IS FEDERALLY-FUNDED R&amp;D WARRANTED?</p>
<b>CONGRESSIONAL BUDGET OFFICER</b>	<p>KEY QUESTION: WHAT WOULD HAVE HAPPENED IN THE ABSENCE OF A GIVEN R&amp;D INVESTMENT, AND HOW CAN CREDIBLE COUNTERFACTUALS BE CONSTRUCTED? (<a href="#">JUMP TO PAGE 11</a>)</p> <p>BETTER COUNTERFACTUAL METHODS WOULD HELP STRENGTHEN THE CREDIBILITY AND ACCURACY OF CBO’S ANALYSIS.</p>



## Federal Agencies & OMB

Federal agency program officers and budget examiners at OMB occupy the operational center of the R&D investment system. They are the actors who design programs, allocate resources, set priorities, and ultimately determine whether evaluation evidence gets built into decision-making or treated as an afterthought. Their information needs are correspondingly practical: they want to know how to construct portfolios that deliver the greatest impact, how to set goals that are meaningful without creating perverse incentives, how to align timelines and metrics across agencies with different missions, and how to make the case for sustained investment in a budget environment that rewards near-term results.

For this group, questions about portfolio construction, accountability standards, and the governance structures that protect long-term research programs from short-cycle pressures are most immediately relevant.

<p><b>OMB</b></p>	<p>KEY QUESTION: WHAT GOVERNANCE STRUCTURES ALLOW LONG-TERM RESEARCH GOALS TO SURVIVE SHORT ELECTORAL AND BUDGET CYCLES? (<a href="#">JUMP TO PAGE 23</a>)</p> <p>WHILE THE FEDERAL BUDGET IS CONSTRUCTED IN ONE-YEAR CYCLES, RESEARCH PRIORITIES AND GOALS OFTEN TAKE MORE THAN ONE YEAR OF FUNDING TO ACHIEVE. OMB STAFF ARE CURIOUS ABOUT HOW THEY CAN DEVELOP APPROACHES TO SUPPORT R&amp;D THAT ACHIEVES LONG-TERM IMPACT.</p> <p>KEY QUESTION: HOW SHOULD R&amp;D PORTFOLIOS BE CONSTRUCTED AND EVALUATED AS A WHOLE, ACROSS AGENCIES WITH VARYING MISSIONS, RISK TOLERANCES, AND TIMELINES? (<a href="#">JUMP TO PAGE 14</a>)</p> <p>OMB, PARTICULARLY THE PROGRAM ASSISTANT DIRECTORS WHO OVERSEE BUDGET TEAMS, THINK ABOUT THE FEDERAL BUDGET ACROSS A PORTFOLIO OF MULTIPLE AGENCIES. WHEN THEY ARE CONSIDERING THE RETURN ON RESEARCH INVESTMENT, IT IS TO HELP UNDERSTAND HOW TO BALANCE FUNDING ACROSS THOSE AGENCIES.</p>
<p><b>SCIENCE AGENCIES</b></p>	<p>KEY QUESTION: HOW CAN THE RETURNS FROM R&amp;D INVESTMENT BE RELIABLY MEASURED, INCLUDING NON-FINANCIAL OUTCOMES? (<a href="#">JUMP TO PAGE 10</a>)</p> <p>KNOWING THE ANSWER TO THIS QUESTION HELPS STAFF AT AGENCIES LIKE NIH AND NSF MAKE THE CASE FOR THEIR PORTFOLIOS. WHERE POSSIBLE, THEY ARE LOOKING FOR DIRECT CAUSAL ATTRIBUTION – THE TYPES OF EVIDENCE THAT HELPS THEM SHOW THEIR INVESTMENTS ARE DELIVERING FOR TAXPAYERS.</p> <p>KEY QUESTION: TO WHAT EXTENT DOES PUBLIC R&amp;D INVESTMENT CATALYZE ADDITIONAL PRIVATE AND PHILANTHROPIC FUNDING, AND WHAT PROGRAM DESIGNS ARE MOST EFFECTIVE AT CROWDING IN CO-INVESTMENT? (<a href="#">JUMP TO PAGE 19</a>)</p> <p>PARTICULARLY FOR NON-TRADITIONAL R&amp;D FUNDING PROGRAMS LIKE ARPAS, A KEY GOAL IS THE TRANSITION PATHWAY TO FUTURE PRIVATE INVESTMENT AND COMMERCIALIZATION. THIS MEANS THAT AN IMPORTANT COMPONENT OF ROI TO DRAW OUT IS THE FUTURE PRIVATE AND PHILANTHROPIC INVESTMENT RELATED TO PROJECTS THEY FUND.</p>

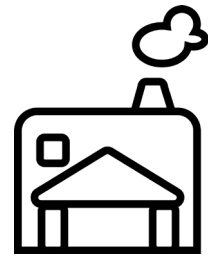


## Universities and the Research Enterprise

University research administrators and faculty leadership sit at the interface between federal funders and the researchers who do the work, navigating demands for accountability and demonstrable impact while simultaneously defending the conditions – autonomy, long time horizons, tolerance for failure – that make basic research productive. Their relationship to R&D evaluation is therefore inherently two-pronged: they need evidence that makes the case for sustained public investment in university-based research, and they need evaluation frameworks that do not inadvertently narrow research agendas or privilege near-term deliverables over fundamental inquiry over longer time horizons.

For this group, questions about how returns unfold over time, what accountability standards are appropriate for different kinds of research, and how the efficiency of the research enterprise itself can be improved without undermining researcher independence are most directly relevant.

<p><b>UNIVERSITIES</b></p>	<p>KEY QUESTION: HOW DO THE RETURNS FROM R&amp;D INVESTMENT UNFOLD OVER TIME, AND WHAT FRAMEWORKS ALLOW FUNDERS WITH DIFFERENT TIME HORIZONS TO ASSESS IMPACT ON THEIR OWN TERMS? (<a href="#">JUMP TO PAGE 22</a>)</p> <p>UNIVERSITIES NEED TO BE ABLE TO MAKE THE CASE FOR INVESTMENT IN THEIR RESEARCHERS AND INFRASTRUCTURE. ANSWERS TO THIS QUESTION WOULD HELP THEM MAKE THE CASE TO A RANGE OF AUDIENCES – LEGISLATORS, DONORS, AND INDUSTRY PARTNERS – EACH OPERATING ON DIFFERENT CLOCKS.</p> <p>KEY QUESTION: WHAT STANDARDS OF EVIDENCE AND ACCOUNTABILITY ARE APPROPRIATE FOR EVALUATING R&amp;D, AND WHO GETS TO SET THE TERMS OF EVALUATION? (<a href="#">JUMP TO PAGE 12</a>)</p> <p>THE TERMS OF EVALUATION OF R&amp;D INVESTMENTS ARE OFTEN DEFINED OUTSIDE THE RESEARCH ENTERPRISE. RESEARCH ON APPROPRIATE ACCOUNTABILITY STANDARDS WOULD GIVE UNIVERSITY ADMINISTRATORS A PRINCIPLED BASIS FOR PUSHING BACK ON EXTERNAL FRAMES THAT DON'T CAPTURE WHAT RESEARCH ACTUALLY PRODUCES.</p>
<p><b>RESEARCH ASSOCIATIONS &amp; OTHER ADVOCATES</b></p>	<p>KEY QUESTION: HOW DO THE BENEFITS OF R&amp;D INVESTMENT SPILL OVER BEYOND THE DIRECT RECIPIENTS, INCLUDING HEALTH AND DRUG DISCOVERY SPILLOVERS? (<a href="#">JUMP TO PAGE 17</a>)</p> <p>RESEARCH ASSOCIATIONS PLAY THE ROLE OF TRANSLATING COMPLEX SCIENCE INTO STORIES THAT MATTER TO A BROADER AUDIENCE. THE MOST PERSUASIVE CASE FOR INVESTING IN RESEARCH ISN'T THE NUMBER OF INFLUENTIAL JOURNAL ARTICLES PRODUCED; RATHER, IT'S THE LIVES CHANGED AND COSTS AVOIDED.</p>



## Philanthropy and the Private Sector

Private sector partners and philanthropic funders engage with the R&D evaluation agenda from outside the federal system, but their decisions – about where to invest, what to co-fund, and how to position their resources relative to public investment – are shaped by many of the same underlying questions that government funders and policymakers focus on. Industry partners and federal contractors want to understand where public R&D complements rather than competes with private investment, and what the evidence says about which funding instruments and partnership structures generate the highest returns.

Philanthropic funders, particularly those operating at the frontier of fields that federal agencies have been slow to support, face a more specific version of this challenge: determining whether their dollars are catalyzing durable public commitment or quietly substituting for it instead.

For this group, questions about crowding in, the public-private boundary, and what can be learned from past programs – including those that did not survive the transition from philanthropic seed funding to sustained federal support – are most directly relevant.

<p><b>INDUSTRY AND FEDERAL CONTRACTORS</b></p>	<p>KEY QUESTION: WHAT IS THE DISTINCT CONTRIBUTION OF PUBLIC R&amp;D INVESTMENT, AND HOW DOES THE CHOICE OF FUNDING INSTRUMENT AFFECT THE RETURNS IT GENERATES? (<a href="#">JUMP TO PAGE 15</a>)</p> <p>FOR THIS AUDIENCE, IT'S KEY TO UNDERSTAND WHEN THE GOVERNMENT IS DOING SOMETHING THAT THE MARKET WON'T. RESEARCH ON THIS QUESTION WOULD SHARPEN WHERE PUBLIC-PRIVATE PARTNERSHIPS ARE NEEDED, AND WHERE PUBLIC INVESTMENT IS A COMPLEMENT VS. A SUBSTITUTE. UNDERSTANDING THE ROI OF VARIOUS FUNDING INSTRUMENTS ALSO SUPPORTS PRIVATE INDUSTRY'S DECISION-MAKING ABOUT WHICH FEDERAL FUNDING OPPORTUNITIES TO PURSUE.</p>
<p><b>PHILANTHROPY</b></p>	<p>KEY QUESTION: TO WHAT EXTENT DOES PHILANTHROPIC R&amp;D INVESTMENT CATALYZE PUBLIC FUNDING, AND WHEN DOES IT QUIETLY SUBSTITUTE FOR IT INSTEAD? (<a href="#">JUMP TO PAGE 19</a>)</p> <p>WHILE THE GOVERNMENT WANTS TO UNDERSTAND IF PUBLIC INVESTMENT LEADS TO CROWDING IN, PHILANTHROPIC INVESTORS WANT TO UNDERSTAND THE OPPOSITE.</p> <p>KEY QUESTION: WHAT CAN BE LEARNED FROM PAST R&amp;D PROGRAMS – AND WHAT HAPPENS TO RESEARCHERS AND PROGRAMS WHEN PHILANTHROPIC FUNDING ENDS? (<a href="#">JUMP TO PAGE 24</a>)</p> <p>SIMILAR TO ABOVE, THIS IS THE VICE-VERSA OF A QUESTION THAT GOVERNMENT PROGRAM OFFICERS MIGHT ASK. FOR PHILANTHROPY, A KEY COMPONENT OF LONG-TERM ROI IS ATTRACTING FUTURE PUBLIC DOLLARS.</p>

## The Research Agenda

---

The eleven questions that follow represent the core of the ROI research agenda. They were developed through a structured process of stakeholder engagement, grounded in a systematic review of the outcomes that R&D investment is understood to produce – spanning social benefits, returns to investors, knowledge accumulation and spillovers, and the contextual factors that shape how those returns are realized and measured.

The questions are organized into three tiers that reflect both their logical relationships and their practical tractability.

**Foundational questions** address the measurement, attribution, and accountability challenges that underpin the entire field of research. Progress on answering these questions is a prerequisite for more confident answers elsewhere in the agenda, even though they are also among the hardest to resolve fully.

**Applied questions** are more bounded in scope and more amenable to near-term empirical progress. They address specific mechanisms, instruments, and effects where credible evidence can be generated and used within reasonable timeframes.

**Structural questions** concern the institutional and political conditions that shape how R&D investment decisions are made and sustained over time. They are not purely empirical questions, and research alone cannot resolve them, but rigorous inquiry can clarify where the leverage points are and what kinds of reform have worked in analogous contexts.

For each question, the sections that follow identify the relevant sub-questions, the outcome measures most pertinent to answering it, the audiences with the greatest stake in its resolution, and an assessment of what makes it tractable or difficult to study well.

## Foundational Questions

These three questions are simultaneously the hardest questions to answer and the most universally needed for decision-making for all key stakeholders. Making meaningful progress toward answering these questions gets to the core of the Griliches Question.

Parts of these questions may have no clean answer. The causal chains are too long, the counterfactual too remote, and some of the values at stake too contested to be resolved empirically. But that does not make them unproductive lines of inquiry. Producing a rigorous account of exactly where the limits of measurement and attribution lie – and being transparent about which findings will remain provisional as methods and data continue to develop – would give practitioners something they currently lack: a reliable guide to what the evidence can and cannot support, and where they should be cautious about claims made in its name.

### Questions

Question 1. How can the returns from R&D investment be reliably measured and attributed, across financial and non-financial outcomes?

Question 2. What would have happened in the absence of a given R&D investment, and how can credible counterfactuals be constructed?

Question 3. What standards of evidence and accountability are appropriate for evaluating R&D, and how can they be applied consistently?

QUESTION DIFFICULTY	LOW	MODERATE	HARD	VERY HARD
---------------------	-----	----------	------	-----------

**Question 1. How can the returns from R&D investment be reliably measured and attributed, across financial and non-financial outcomes?**

Measuring the returns to R&D investment is the field’s most fundamental challenge and, to date, its least resolved one. Existing approaches tend to capture a narrow band of financial and bibliometric outcomes while leaving the broader range of social, environmental, and intrinsic value that public research generates largely unmeasured. Progress on this question is a prerequisite for credible answers to nearly every other question in this agenda.

**Key Audiences**



CONGRESS, INCLUDING THE CONGRESSIONAL RESEARCH SERVICE



OMB, STATE BUDGET OFFICES, SCIENCE AGENCY PROGRAM OFFICERS

**What specific questions do stakeholders have?**

- What methods best attribute downstream economic and social outcomes to specific upstream funding decisions, given long and indirect causal chains (for basic science research in particular)?
- How can non-pecuniary returns – such as scientific knowledge, health improvements, or cultural value – be assigned credible monetary or comparable values?
- How do ROI estimates vary systematically across fields, funding mechanisms, and research stages, and what explains those differences?
- What data infrastructure would need to exist to make attribution feasible at scale across federal R&D programs?

**Outcomes of Interest**

SOCIAL	RETURNS TO INVESTORY
MORTALITY REDUCTION, QUALITY-ADJUSTED LIFE YEARS, HEALTH IMPROVEMENT, LIFE EXPECTANCY, JOB CREATION, JOB QUALITY, DISTRIBUTIONAL EQUITY, PRODUCTIVITY GAINS	INTERNAL RATE OF RETURN, INVESTMENT MULTIPLE, CASH RETURNS, TAX REVENUE, TECHNOLOGY ASSET VALUE, OPTION VALUE, ENTERPRISE VALUE
SPIILLOVER	OTHER/CONTEXT
CITATIONS, SEMANTIC INFLUENCE, FIELD GROWTH, USER ADOPTION, SOCIAL STABILITY	NON-PECUNIARY VALUATION, EX POST ASSESSMENT, RISK-ADJUSTED RETURNS, MARGINAL EFFICIENCY, DATA AVAILABILITY

**How answerable is this question?**

We have made lots of progress on narrow sub-questions and specific outcomes listed above, but a credible, general/cross-sector measurement framework may not be within reach.

WHAT MAKES IT TRACTABLE?	WHAT MAKES IT HARD?
<ul style="list-style-type: none"> <li>• FINANCIAL RETURNS (IRR, TAX REVENUE) CAN BE ESTIMATED WITH REASONABLE CONFIDENCE IN SOME SECTORS</li> <li>• HEALTH ECONOMICS HAS WELL-DEVELOPED METHODS FOR VALUING NON-FINANCIAL OUTCOMES (QALYS, VSL)</li> <li>• LARGE LITERATURE TO BUILD ON – THIS IS THE MOST-STUDIED QUESTION IN THE FIELD</li> </ul>	<ul style="list-style-type: none"> <li>• ATTRIBUTION ACROSS LONG CAUSAL CHAINS IS NEARLY INTRACTABLE FOR BASIC RESEARCH</li> <li>• NON-PECUNIARY RETURNS ARE HARDER TO CREDIBLY MEASURE</li> <li>• RESULTS ARE HIGHLY SENSITIVE TO MODELING ASSUMPTIONS AND CAN BE EASILY WEAPONIZED SELECTIVELY</li> <li>• NO CONSENSUS ON HOW TO AGGREGATE ACROSS INCOMMENSURABLE OUTCOME TYPES</li> </ul>

QUESTION DIFFICULTY	LOW	MODERATE	HARD	VERY HARD
---------------------	-----	----------	------	-----------

**Question 2. What would have happened in the absence of a given R&D investment, and how can credible counterfactuals be constructed for public funding decisions?**

Even where returns can be measured, attributing them to a specific funding decision requires knowing what would have occurred without it – a question that the existing literature addresses inconsistently and often inadequately. The counterfactual problem is especially acute for basic research, where causal chains are long, indirect, and difficult to trace to a single funding source. Developing more rigorous and transparent approaches to counterfactual construction would strengthen the evidentiary foundation on which the entire field rests.

**Key Audiences**



CBO, CRS



PHILANTHROPY,  
PRIVATE INDUSTRY

**What specific questions do stakeholders have?**

- What do we know about the fate of unfunded grant applications?
- Under what conditions does public R&D crowd out private investment, and when does it crowd it in?
- How much value from publicly funded research would have been generated anyway?
- What natural experiments are available to estimate causal effects of R&D funding?

**Outcomes of Interest**

SOCIAL	RETURNS TO INVESTORY
NET SOCIAL BENEFIT, HARM AVOIDANCE, UNINTENDED BENEFITS, UNINTENDED HARMS	OPPORTUNITY COSTS, EXCESS RETURN, FIRM SURVIVAL
SPIILOVER	OTHER/CONTEXT
UNINTENDED CONSEQUENCES, NEARBY FIRM EFFECTS, LESSONS LEARNED, CO-INVESTMENT ATTRACTED	COUNTERFACTUAL BASELINE, CONTROL GROUP, FAILURE VALUE, RISK-ADJUSTED RETURNS

**How answerable is this question?**

This may be the most methodologically challenging question in this entire research agenda – while there are opportunities for rigorous, causal research designs that relate to this question, those designs would be narrow in scope and may lack the generalizability needed to inform decisions. And constructing a counterfactual for basic science may be out of reach entirely.

WHAT MAKES IT TRACTABLE?	WHAT MAKES IT HARD?
<p>OPPORTUNITY FOR CAUSAL DESIGNS:</p> <ul style="list-style-type: none"> <li>• REGRESSION DISCONTINUITY DESIGNS AROUND FUNDING THRESHOLDS OFFER SOME CAUSAL TRACTION</li> <li>• NATURAL EXPERIMENTS (BUDGET CUTS, PROGRAM CANCELLATIONS) CREATE USABLE VARIATION</li> <li>• SOME AGENCIES (NIH) HAVE DETAILED ENOUGH RECORDS TO SUPPORT QUASI-EXPERIMENTAL WORK</li> </ul> <p>DATA AVAILABILITY:</p> <ul style="list-style-type: none"> <li>• UNFUNDED GRANT APPLICANT TRACKING IS FEASIBLE AND UNDERUTILIZED</li> </ul>	<p>CHALLENGES WITH CAUSAL DESIGNS:</p> <ul style="list-style-type: none"> <li>• R&amp;D OUTCOMES ARE NON-FUNGIBLE – THERE IS NO TRUE CONTROL GROUP FOR A SPECIFIC DISCOVERY</li> <li>• THE COUNTERFACTUAL WORLD IS FUNDAMENTALLY UNOBSERVABLE OVER LONG TIME HORIZONS</li> <li>• EVEN VALID QUASI-EXPERIMENTS ANSWER NARROW LOCAL QUESTIONS, NOT THE GENERAL COUNTERFACTUAL</li> <li>• KNOWLEDGE SPILLOVERS MEAN “WHAT WOULD HAVE HAPPENED” IS A GLOBAL QUESTION, NOT A LOCAL ONE</li> </ul> <p>DATA AVAILABILITY:</p> <ul style="list-style-type: none"> <li>• AGENCIES RARELY TRACK UNFUNDED APPLICANTS SYSTEMATICALLY – DATA INFRASTRUCTURE IS MISSING</li> </ul>

QUESTION DIFFICULTY	LOW	MODERATE	HARD	VERY HARD
---------------------	-----	----------	------	-----------

**Question 3. What standards of evidence and accountability are appropriate for evaluating publicly funded R&D, and how can they be applied consistently across agencies and programs?**

Federal R&D programs are evaluated using a heterogeneous and often incompatible mix of methods, metrics, and standards, making it difficult to compare performance across agencies or accumulate knowledge over time. A more consistent approach to evaluation – one that sets appropriate expectations for rigor given the nature of the research being funded – would serve both the programs being evaluated and the oversight community that relies on evaluation findings. This question is as much about institutional design as it is about methodology.

**Key Audiences**



CONGRESS, GAO



OMB

**What specific questions do stakeholders have?**

- What level of rigor is feasible and appropriate for different types of R&D programs?
- How should program goals be specified ex ante to be meaningful and measurable without perverse incentives?
- What monitoring frameworks allow for real-time course correction without undermining researcher autonomy?
- How can evaluation findings be made credible and comparable across agencies using heterogeneous methods?

**Outcomes of Interest**

SOCIAL	RETURNS TO INVESTORY
EDUCATION OUTCOMES, TRAINING OUTCOMES, ENGAGEMENT, NET SOCIAL BENEFIT, HARM AVOIDANCE	EXPENDITURE ACCOUNTABILITY, PUBLICATIONS, RESEARCH QUESTIONS ADDRESSED, KNOWLEDGE GENERATED, RESEARCH RELEVANCE
SPILLOVER	OTHER/CONTEXT
UNINTENDED CONSEQUENCES, THIRD-PARTY VALIDATION, MISSION ADVANCEMENT, STAKEHOLDER LEGITIMACY, LESSONS LEARNED	EX POST ASSESSMENT, PROGRESS MONITORING, PROCESS MONITORING, DATA AVAILABILITY, CONTROL GROUP, PUBLICATION IMPACT

**How answerable is this question?**

The first part of this question, developing standards that can be used to evaluate R&D programs, is easier: we have a large literature base and evaluation community that can inform. The second part of this question, learning how to apply those standards, is much harder – it requires research funding agencies to allow evaluation of their practices and to act on what they learn.

WHAT MAKES IT TRACTABLE?	WHAT MAKES IT HARD?
<ul style="list-style-type: none"> <li>• EVALUATION SCIENCE LITERATURE ON STANDARDS OF EVIDENCE IS WELL-DEVELOPED</li> <li>• EXISTING FRAMEWORKS (E.G. WHAT WORKS CLEARINGHOUSE, TIERED EVIDENCE) OFFER TEMPLATES</li> <li>• SURVEY OF CURRENT AGENCY EVALUATION PRACTICES IS FEASIBLE</li> <li>• PARTLY A NORMATIVE/DESIGN QUESTION – DOESN'T REQUIRE NEW EMPIRICAL DATA</li> </ul>	<ul style="list-style-type: none"> <li>• R&amp;D RESISTS RCT-STYLE EVALUATION MORE THAN SOCIAL PROGRAMS DO</li> <li>• AGENCIES HAVE HETEROGENEOUS MISSIONS – ONE STANDARD MAY NOT FIT ALL</li> <li>• POLITICAL INCENTIVES ACTIVELY RESIST ACCOUNTABILITY FRAMEWORKS THAT COULD EMBARRASS PROGRAMS</li> <li>• CONSISTENCY ACROSS AGENCIES REQUIRES INTERAGENCY AGREEMENT THAT IS HARD TO ACHIEVE</li> </ul>

## Applied Questions

These are the practical, tractable questions; research on these topics will produce the most immediately usable findings. These get into the details of what science agencies, OMB, and Congress want to know now about funding instrument effectiveness, crowd-in multipliers, and researcher burden. And for the most part, we have the methodology to answer these questions, meaning researchers could turn around answers on a timeline more suited to policymaking.

### Questions

Question 4. How should R&D portfolios be constructed and evaluated as a whole, across varying risk, stage, and field?

Question 5. Under what conditions is public R&D investment justified relative to private alternatives, and how does the funding instrument affect returns?

Question 6. How do the benefits of R&D spill over beyond direct recipients, and how can indirect effects be measured?

Question 7. To what extent does public R&D catalyze additional private and philanthropic funding, and what designs are most effective at crowding in co-investment?

Question 8. How efficiently is R&D funding converted into research activity, and what reforms would reduce friction?

QUESTION DIFFICULTY	LOW	MODERATE	HARD	VERY HARD
---------------------	-----	----------	------	-----------

**Question 4. How should R&D portfolios be constructed and evaluated as a whole, and what allocation strategies maximize returns across projects with varying risk, stage, and field?**

Individual project evaluation tells only part of the story; the performance of a research portfolio depends on how its components are selected, balanced, and managed in relation to one another. Concepts from financial portfolio theory offer a useful starting point, but public R&D portfolios differ from financial ones in ways that limit direct application, including the difficulty of quantifying risk ex ante and the presence of social objectives that market returns cannot capture. Developing portfolio-level evaluation frameworks suited to the specific features of public research investment would address a significant gap in current practice.

**Key Audiences**



PHILANTHROPY



SCIENCE AGENCIES, OMB

**What specific questions do stakeholders have?**

- What is the optimal balance between high-risk/high-reward and incremental research within a public R&D portfolio?
- How should financial portfolio diversification principles be adapted for public research funding?
- At what stage of research development does public investment generate the highest marginal return?
- Can AI tools reliably improve portfolio selection decisions, and what are the risks of algorithmic prioritization?

**Outcomes of Interest**

SOCIAL	RETURNS TO INVESTORY
DISTRIBUTIONAL EQUITY, INCLUSION RATE, NON-INVESTOR VALUE, HUMAN CAPITAL	PORTFOLIO PERFORMANCE, RISK-ADJUSTED RETURNS, INVESTMENT MULTIPLE, GEOPOLITICAL ADVANTAGE, NEW PRODUCTS / SERVICES
SPILLOVER	OTHER/CONTEXT
COLLABORATIONS, FIELD GROWTH, MISSION DEFINITION	PORTFOLIO BALANCE, PORTFOLIO PERFORMANCE, SELECTION MECHANISM, COMPLEMENTARY INVESTMENTS, INCUBATOR ALIGNMENT, MARGINAL EFFICIENCY

**How answerable is this question?**

The challenge of this question is that there is likely no one-size-fits-all answer. However, we have the building blocks to understand the answer for individual agencies, which can still be informative to policymakers.

WHAT MAKES IT TRACTABLE?	WHAT MAKES IT HARD?
<ul style="list-style-type: none"> <li>• FINANCIAL PORTFOLIO THEORY PROVIDES A STRONG CONCEPTUAL STARTING POINT</li> <li>• AGENCY GRANT DATABASES OFFER DATA ON PORTFOLIO COMPOSITION OVER TIME</li> <li>• ARPA-E AND NIH PROVIDE CONTRASTING NATURAL EXPERIMENTS IN PORTFOLIO DESIGN</li> <li>• STAGE-OF-DEVELOPMENT QUESTIONS HAVE SOME EMPIRICAL LITERATURE TO BUILD ON</li> </ul>	<ul style="list-style-type: none"> <li>• DEFINING “RETURN” AT THE PORTFOLIO LEVEL REQUIRES RESOLVING OTHER FOUNDATIONAL QUESTIONS FIRST</li> <li>• CROSS-AGENCY PORTFOLIOS ARE POLITICALLY FRAGMENTED AND HARD TO ANALYZE AS A UNIT</li> <li>• RISK IN R&amp;D IS FUNDAMENTALLY DIFFERENT FROM FINANCIAL RISK – HARDER TO QUANTIFY EX ANTE</li> <li>• OPTIMAL PORTFOLIO COMPOSITION LIKELY VARIES BY MISSION IN WAYS THAT RESIST GENERALIZATION</li> </ul>

QUESTION DIFFICULTY	LOW	MODERATE	HARD	VERY HARD
---------------------	-----	----------	------	-----------

**Question 5. What is the distinct contribution of public R&D investment, and how does the choice of funding instrument affect the returns it generates?**

Public R&D investment can serve a number of different purposes: it funds research that markets would systematically underprovide; it pursues missions – national security, public health, environmental stewardship, the advancement of knowledge – that private investors are neither equipped nor incentivized to prioritize; it builds the scientific infrastructure and human capital on which private innovation ultimately depends. This question does not assume that market failure is the only or even the primary rationale for public investment. Rather, it asks which of those rationales hold up under scrutiny, and whether the instruments through which public investment is delivered – grants, prizes, contracts, tax incentives, procurement – are well matched to the goals being pursued.

**Key Audiences**



CONGRESS, INCLUDING THE CONGRESSIONAL RESEARCH SERVICE



PHILANTHROPY, PRIVATE INDUSTRY

**What specific questions do stakeholders have?**

- How do agencies define the problem (or “market failure”) their R&D investment is meant to solve, and how consistently is that definition used to evaluate whether the investment is working?
- What is the comparative effectiveness of funding instruments – grants, prizes, contracts, tax incentives, and procurement – for different kinds of research questions? When is it most appropriate to use each of these instruments?
- Is the marginal dollar at NIH or NSF generating positive returns, and how would we know if it weren't?
- How should policymakers weigh investment in basic research against more applied and translational priorities, and what does the evidence say about the potential long-run consequences of shifting that balance?
- How should the case for public R&D be made to skeptical audiences who question whether government funders can pick priorities as well as markets can?

**Outcomes of Interest**

SOCIAL	RETURNS TO INVESTORY
NON-INVESTOR VALUE, EQUITY OF ACCESS, SOCIAL COHESION, INDUCED INVESTMENT	OPPORTUNITY COST, EXCESS RETURN, GEOPOLITICAL ADVANTAGE, TAX REVENUE, NEW PRODUCTS / SERVICES
SPILLOVER	OTHER/CONTEXT
MISSION ADVANCEMENT, MISSION DEFINITION, SOCIAL STABILITY, R&D LEGITIMACY, R&D SUPPORT	COUNTERFACTUAL BASELINE, MARGINAL EFFICIENCY, SELECTION MECHANISM, LEGISLATION, IP RIGHTS

**How answerable is this question?**

Real progress can be made toward answering this question and especially specific sub-questions (i.e. comparing various funding instruments). But the broader question of what public investment in science is ultimately trying to achieve – and how competing rationales should be weighed against one another – is as much a question of values as of evidence, and consensus on it should not be expected or required.

WHAT MAKES IT TRACTABLE?	WHAT MAKES IT HARD?
<ul style="list-style-type: none"> <li>• MARKET FAILURE THEORY PROVIDES A CLEAR CONCEPTUAL FRAMEWORK</li> <li>• INSTRUMENT COMPARISON (GRANTS VS. PRIZES VS. CONTRACTS) HAS GROWING LITERATURE</li> <li>• SOME NATURAL EXPERIMENTS EXIST WHERE INSTRUMENTS CHANGED OVER TIME</li> <li>• INTERNATIONAL COMPARISONS OF POLICY INSTRUMENTS OFFER USEFUL VARIATION</li> </ul>	<ul style="list-style-type: none"> <li>• MARKET FAILURE IS HARD TO DEFINE EMPIRICALLY – AGENCIES OFTEN USE IT LOOSELY</li> <li>• INSTRUMENT EFFECTS DEPEND HEAVILY ON IMPLEMENTATION CONTEXT</li> <li>• THE “MARGINAL DOLLAR AT NIH” QUESTION REQUIRES A CREDIBLE COUNTERFACTUAL THAT IS VERY HARD TO CONSTRUCT</li> <li>• POLITICALLY CONTENTIOUS – FINDINGS WILL BE USED SELECTIVELY REGARDLESS OF NUANCE</li> </ul>

QUESTION DIFFICULTY	LOW	MODERATE	<b>HARD</b>	VERY HARD
---------------------	-----	----------	-------------	-----------

**Question 6. How do the benefits of R&D investment spill over beyond the direct recipients, and how can these indirect effects be measured?**

Some of the most significant value generated by public R&D investment accrues not to the funded researchers or their institutions but to firms, academic fields, and communities that were never directly involved in the work. These spillover effects are widely acknowledged but inconsistently measured, and the methods available to capture them – from citation analysis to semantic similarity to local economic modeling – vary considerably in their reliability and scope. Improving the measurement of spillovers is essential to any honest accounting of what public research investment produces.

**Key Audiences**



CONGRESS



OMB



RESEARCH ENTERPRISE

**What specific questions do stakeholders have?**

- How large are cross-sector spillovers from basic research, and how do they propagate?
- To what extent do other countries free-ride on US public R&D investment?
- Can computational methods like semantic similarity provide better estimates of knowledge diffusion than citation counts?
- How do R&D spillovers interact with complementary investments, and should these be evaluated jointly?

**Outcomes of Interest**

SOCIAL	RETURNS TO INVESTORY
CROSS-SECTOR GAINS, HEALTH IMPROVEMENT, MENTAL HEALTH, ADDICTION REDUCTION, NUTRITIONAL OUTCOMES, ENVIRONMENTAL IMPACT, INTRINSIC VALUE	NEW PRODUCTS / SERVICES, KNOWLEDGE GENERATED, PARTNERSHIPS, TECHNOLOGY ASSET VALUE
SPILLOVER	OTHER/CONTEXT
CITATIONS, IDEA DIFFUSION, GENERATIONAL INFLUENCE, PARADIGM SHIFTS, PUBLIC UNDERSTANDING, FIELD GROWTH, NEARBY FIRM EFFECTS, TRAINEE OUTCOMES, USER ADOPTION	COUNTERFACTUAL BASELINE, MARGINAL EFFICIENCY, SELECTION MECHANISM, LEGISLATION, IP RIGHTS

**How answerable is this question?**

The existing literature tackles measuring spillovers in a number of ways (citations, patents, nearby firm effects), so there has been meaningful progress toward answering this question. However, the types of spillovers that might matter most, such as paradigm shifts, are not easy to quantify.

WHAT MAKES IT TRACTABLE?	WHAT MAKES IT HARD?
<ul style="list-style-type: none"> <li>• CITATION AND PATENT DATA CAPTURE SOME SPILLOVERS RELIABLY</li> <li>• SEMANTIC SIMILARITY METHODS ARE PROMISING AND INCREASINGLY AVAILABLE</li> <li>• CROSS-SECTOR SPILLOVERS (E.G. EARTH SCIENCE → AGRICULTURE) HAVE TRACEABLE PATHWAYS</li> <li>• LOCAL ECONOMIC SPILLOVER LITERATURE (NEARBY FIRMS) HAS ESTABLISHED METHODS</li> </ul>	<ul style="list-style-type: none"> <li>• SOME IMPORTANT SPILLOVERS (INFLUENCE ON WAYS OF THINKING, FUTURE GENERATIONS) ARE THE LEAST MEASURABLE</li> <li>• SPILLOVERS BY DEFINITION DON'T SHOW UP IN THE FUNDED PROJECT'S OWN RECORDS</li> <li>• INTERNATIONAL FREE-RIDING EFFECTS REQUIRE DATA THAT US AGENCIES DON'T COLLECT</li> <li>• VALIDATING SEMANTIC SIMILARITY METHODS AS SPILLOVER PROXIES IS STILL AN OPEN RESEARCH PROBLEM</li> </ul>

QUESTION DIFFICULTY	LOW	MODERATE	HARD	VERY HARD
---------------------	-----	----------	------	-----------

**Question 7. To what extent does public R&D investment catalyze additional private and philanthropic funding, and what program designs are most effective at crowding in co-investment?**

One of the strongest arguments for public R&D investment is that it does not simply add to the total stock of research activity but multiplies it, by attracting private and philanthropic capital that would not otherwise have been deployed. The evidence on this crowd-in effect is promising but uneven, and relatively little is known about which program designs are most effective at generating it – or about the conditions under which public investment may instead deter private participation. This question has particular relevance for funders operating at the boundary between public and private investment.

**Key Audiences**



CONGRESS



OMB

**What specific questions do stakeholders have?**

- What is the empirical multiplier effect of public R&D on private capital flows?
- Which structural features of funding programs are most effective at attracting co-investment?
- How reliable are industry commitment signals as evaluation metrics?
- Under what conditions does public funding deter rather than attract private investment?

**Outcomes of Interest**

SOCIAL	RETURNS TO INVESTORY
INDUCED INVESTMENT, JOB CREATION, WORKFORCE DEVELOPMENT	PARTNERSHIPS, ENTERPRISE VALUE, GROWTH RATE, NEW PRODUCTS / SERVICES
SPILLOVER	OTHER/CONTEXT
CO-INVESTMENT ATTRACTED, COLLABORATIONS, FIELD GROWTH, JOBS PROGRAM EFFECT	CO-INVESTMENT ATTRACTED, COMPLEMENTARY INVESTMENTS, PORTFOLIO PERFORMANCE, TOPIC SALIENCE

**How answerable is this question?**

There is existing literature on this question that can help point the way, particularly on crowd-in investment. But it's clear that the idea of determining when public R&D funding is a complement vs. a substitute for private and philanthropic funding is important to decision-makers and underexplored, likely due to limitations in currently available data.

WHAT MAKES IT TRACTABLE?	WHAT MAKES IT HARD?
<ul style="list-style-type: none"> <li>• PRIVATE INVESTMENT DATA (VC, CORPORATE R&amp;D) IS RELATIVELY AVAILABLE</li> <li>• SOME EXISTING ECONOMETRIC LITERATURE ON CROWD-IN/CROWD-OUT EFFECTS</li> <li>• PROGRAM-LEVEL ANALYSIS (E.G. SBIR, ATP) HAS NATURAL BEFORE/AFTER DESIGNS</li> <li>• PHILANTHROPIC DATABASES (CANDID, FOUNDATION 990S) PROVIDE CO-INVESTMENT DATA</li> </ul>	<ul style="list-style-type: none"> <li>• SEPARATING CROWD-IN EFFECTS FROM SELECTION EFFECTS (GOOD PROJECTS ATTRACT BOTH) IS HARD</li> <li>• TIME LAGS BETWEEN PUBLIC FUNDING AND PRIVATE FOLLOW-ON ARE LONG AND VARIABLE</li> <li>• A KEY QUESTION – DOES PHILANTHROPY SUBSTITUTE FOR PUBLIC FUNDING? – HAS ALMOST NO LITERATURE</li> <li>• INDUSTRY COMMITMENT SIGNALS ARE OFTEN SOFT AND NOT RELIABLY TRACKED</li> </ul>

QUESTION DIFFICULTY	LOW	MODERATE	HARD	VERY HARD
---------------------	-----	----------	------	-----------

**Question 8. How efficiently is R&D funding converted into actual research activity, and what reforms would reduce friction and increase the share of resources reaching researchers?**

Not all of the resources allocated to public R&D reach researchers in a form they can use; a significant share is absorbed by administrative overhead, compliance requirements, and institutional transaction costs that consume time and money without contributing to research outcomes. This is among the most tractable questions in the agenda, in that much of the relevant data exists and several candidate reforms have been identified, though rigorous evaluation of their effects remains limited. Addressing it would increase the effective value of existing R&D investments without requiring additional appropriations.

**Key Audiences**



SCIENCE AGENCIES



UNIVERSITIES, RESEARCH ENTERPRISE

**What specific questions do stakeholders have?**

- What share of federal R&D funding is absorbed by administrative overhead, and how does this compare internationally?
- How much researcher time is consumed by compliance requirements, and what is the opportunity cost?
- Is the research university the most efficient performer of publicly funded R&D for certain types of work?
- What is the depreciation rate of R&D capital, and how should this inform funding levels and continuity decisions?

**Outcomes of Interest**

SOCIAL	RETURNS TO INVESTORY
TRAINEES SUPPORTED, SCIENTIFIC CAREERS, HUMAN CAPITAL	EXPENDITURE ACCOUNTABILITY, OPPORTUNITY COST, TAX REVENUE, NEW PRODUCTS / SERVICES
SPILLOVER	OTHER/CONTEXT
TRAINEE OUTCOMES, COLLABORATIONS, JOBS PROGRAM EFFECT	EXPERT TIME INVESTED, FACILITIES INVESTMENT, MARGINAL EFFICIENCY, INCUBATOR ALIGNMENT, COUNTERFACTUAL BASELINE

**How answerable is this question?**

There is existing literature on this topic, so the primary need is synthesis for decision-makers. The most challenging part of answering this question is convincing policymakers to experiment with and evaluate reforms that reduce administrative burden, though there seems to be openness to this in the current moment.

WHAT MAKES IT TRACTABLE?	WHAT MAKES IT HARD?
<ul style="list-style-type: none"> <li>• ADMINISTRATIVE DATA ON INDIRECT COST RATES AND GRANT OVERHEAD IS LARGELY AVAILABLE</li> <li>• SURVEY-BASED TIME-USE STUDIES OF RESEARCHERS ARE WELL-ESTABLISHED</li> <li>• INTERNATIONAL COMPARISONS OFFER NATURAL BENCHMARKS</li> <li>• REFORMS HAVE NATURAL BEFORE/AFTER EVALUATION DESIGNS</li> </ul>	<ul style="list-style-type: none"> <li>• AGENCIES RESIST SHARING DETAILED COST DATA</li> <li>• HARD TO DEFINE THE RIGHT EFFICIENCY BENCHMARK</li> <li>• CAUSAL LINK FROM REDUCED OVERHEAD TO BETTER RESEARCH OUTPUT IS CONTESTED</li> </ul>

## **Structural Questions**

Answering these questions would be both diagnostic and prescriptive: there is a need to clearly document the challenges, identify where the leverage points are, and give reformers the evidence they need to make the case for change. They would inform Congress and OMB about where institutional change is needed.

### **Questions**

Question 9. How do returns unfold over time, and what frameworks allow funders with different time horizons to assess impact?

Question 10. How do political and institutional incentives shape R&D decisions, and what governance structures protect long-term goals from short cycles?

Question 11. What can be learned from past R&D programs – including failures – and how should that knowledge inform future investment?

QUESTION DIFFICULTY	LOW	MODERATE	HARD	VERY HARD
---------------------	-----	----------	------	-----------

**Question 9. How do the returns from R&D investment unfold over time, and what evaluation frameworks allow funders with different time horizons to assess impact on their own terms?**

The returns to public R&D investment characteristically accrue over timescales that are poorly matched to the planning and accountability cycles of the institutions that fund it, creating persistent pressure to prioritize near-term outputs over longer-run value. Developing evaluation frameworks that allow funders operating on different timescales – from annual appropriations cycles to decade-long strategic plans – to assess impact on terms meaningful to them would help ease this tension, though it cannot eliminate it entirely. This question sits at the intersection of empirical research and institutional design, and progress on it requires both.

**Key Audiences**



SCIENCE AGENCIES, OMB



UNIVERSITIES, RESEARCH ENTERPRISE



PHILANTHROPY

**What specific questions do stakeholders have?**

- What is the empirical distribution of lag times between R&D investment and measurable impact?
- What leading indicators can reliably predict long-run returns early enough to be useful?
- How should discount rates be applied given the nonlinear, uncertain nature of R&D returns?
- Can milestone-based or staged evaluation frameworks bridge political accountability and long research timelines?

**Outcomes of Interest**

SOCIAL	RETURNS TO INVESTORY
BENEFIT DURABILITY, BENEFIT SUSTAINABILITY, TRAINEE TRAJECTORIES, BASIC RESEARCH PUBLICATIONS, SCIENTIFIC CAREERS	RESEARCH RELEVANCE, GROWTH RATE, DISCOUNT RATE, TIMELINE, OPTION VALUE
SPILLOVER	OTHER/CONTEXT
GENERATIONAL INFLUENCE, PARADIGM SHIFTS, PUBLIC UNDERSTANDING, TOPIC SALIENCE, IDEA DIFFUSION	EX POST ASSESSMENT, DIFFUSION RATE, COMPLEMENTARY INVESTMENTS, COUNTERFACTUAL BASELINE

**How answerable is this question?**

This is the question where the gap between what researchers can learn and what practitioners need is widest. Descriptive work on lag distributions is feasible; validated frameworks that work across different funder time horizons are much harder, because the problem is partly normative and partly political rather than purely empirical.

WHAT MAKES IT TRACTABLE?	WHAT MAKES IT HARD?
<ul style="list-style-type: none"> <li>• HISTORICAL R&amp;D INVESTMENTS PROVIDE LONG TIME SERIES TO ANALYZE</li> <li>• BIBLIOMETRIC AND PATENT DATA CAN TRACE KNOWLEDGE DIFFUSION OVER DECADES</li> <li>• SOME WELL-STUDIED CASES (E.G. NIH → DRUG DEVELOPMENT) PROVIDE ANCHORS</li> <li>• LEADING INDICATOR RESEARCH IS A GROWING FIELD</li> </ul>	<ul style="list-style-type: none"> <li>• EXTREMELY LONG LAGS MAKE CAUSAL ATTRIBUTION NEARLY IMPOSSIBLE</li> <li>• THE RESEARCH THAT MATTERS MOST (BASIC) HAS THE LEAST TRACEABLE RETURNS</li> <li>• NO VALIDATED LEADING INDICATORS EXIST FOR MOST RESEARCH TYPES</li> <li>• DISCOUNT RATE CHOICE IS NORMATIVELY CONTESTED, NOT JUST EMPIRICALLY UNCERTAIN</li> <li>• BRIDGING SHORT POLITICAL CYCLES AND LONG RESEARCH TIMELINES MAY BE STRUCTURALLY UNSOLVABLE BY RESEARCH ALONE</li> </ul>

QUESTION DIFFICULTY	LOW	MODERATE	HARD	VERY HARD
---------------------	-----	----------	------	-----------

**Question 10. How do political and institutional incentives shape R&D investment decisions, and what governance structures allow long-term research goals to survive short electoral and budget cycles?**

The allocation of public R&D funding is not determined by evidence alone (or in some cases, at all); it is shaped by political incentives, institutional arrangements, and budgetary dynamics that can pull investment toward visible, near-term payoffs and away from the basic research that generates the highest long-run returns. Understanding how those incentives operate – and what governance structures have successfully insulated research programs from their most distorting effects – is essential for anyone seeking to improve the conditions under which public investment in science is made and sustained. Research can clarify the mechanisms and document what has worked; changing the incentives themselves requires political will that research cannot supply.

**Key Audiences**



CONGRESS



OMB



UNIVERSITIES,  
RESEARCH  
ENTERPRISE

**What specific questions do stakeholders have?**

- How do congressional appropriations processes distort R&D funding priorities, and what reforms have successfully insulated programs?
- What institutional designs best protect research programs from short-term political interference?
- Is there evidence that politically salient topics are over- or under-funded relative to expected returns?
- What communication strategies help legislators understand and value long-horizon research?

**Outcomes of Interest**

SOCIAL	RETURNS TO INVESTORY
CIVIC ENGAGEMENT, SOCIAL COHESION, BENEFIT DURABILITY	GEOPOLITICAL ADVANTAGE, RESEARCH RELEVANCE, TAX REVENUE
SPILLOVER	OTHER/CONTEXT
R&D SUPPORT, R&D LEGITIMACY, PUBLIC OPINION, MISSION DEFINITION, SOCIAL STABILITY	LEGISLATION, TOPIC SALIENCE, SELECTION MECHANISM, INCUBATOR ALIGNMENT

**How answerable is this question?**

Unlike many of the other questions in this research agenda there is no need to develop new methods to answer this question. The only limitation is that it would be challenging to elicit causal claims about what governance designs work – but qualitative/descriptive research building on the existing political science literature would give policymakers a clear direction.

WHAT MAKES IT TRACTABLE?	WHAT MAKES IT HARD?
<ul style="list-style-type: none"> <li>• RICH HISTORICAL RECORD OF CONGRESSIONAL BUDGET CYCLES AND AGENCY FUNDING PATTERNS</li> <li>• POLITICAL SCIENCE LITERATURE ON INSTITUTIONAL DESIGN IS WELL-DEVELOPED</li> <li>• CASE STUDIES OF ARPA-E, NIH, AND NSF PROVIDE COMPARATIVE MATERIAL</li> <li>• QUALITATIVE INTERVIEWS WITH AGENCY STAFF AND HILL STAFF ARE FEASIBLE</li> </ul>	<ul style="list-style-type: none"> <li>• HARD TO SEPARATE POLITICAL EFFECTS FROM OTHER CAUSES OF FUNDING INSTABILITY</li> <li>• GOVERNANCE REFORM FINDINGS ARE CONTEXT-SPECIFIC AND DON'T GENERALIZE EASILY</li> <li>• SENSITIVE TOPIC – AGENCY STAFF MAY BE RELUCTANT TO SPEAK CANDIDLY</li> </ul>

QUESTION DIFFICULTY	LOW	MODERATE	HARD	VERY HARD
---------------------	-----	----------	------	-----------

**Question 11. What can be learned from past R&D programs – including those that failed – and how should that knowledge systematically inform future investment decisions?**

The federal R&D system has accumulated decades of experience with programs that succeeded, stalled, or were quietly discontinued, but that experience is imperfectly documented and inconsistently transmitted to the people making current investment decisions. Failure in particular is systematically underrecorded, despite carrying information that is often more instructive than success. Building more deliberate mechanisms for capturing and applying lessons from past programs – including the hardest ones to document – would improve the quality of investment decisions across the ecosystem.

**Key Audiences**



PHILANTHROPY



SCIENCE AGENCIES, OMB

**What specific questions do stakeholders have?**

- What are the most common causes of failure in publicly funded R&D programs, and are they acted upon in subsequent cycles?
- How is institutional memory maintained within funding agencies, and what is lost when staff turn over?
- Can systematic case study databases of past R&D programs improve future investment decisions?
- How should the expected value of learning from failure be factored into ex ante assessment of high-risk proposals?

**Outcomes of Interest**

SOCIAL	RETURNS TO INVESTORY
UNINTENDED BENEFITS, UNINTENDED HARMS, BENEFIT DURABILIT, TRAINEE TRAJECTORIES	KNOWLEDGE GENERATED, RESEARCH QUESTIONS ADDRESSED, RESEARCH RELEVANCE, FIRM SURVIVAL
SPILLOVER	OTHER/CONTEXT
LESSONS LEARNED, UNINTENDED CONSEQUENCES, GENERATIONAL INFLUENCE, TRAINEE OUTCOMES	FAILURE VALUE, EX POST ASSESSMENT, PROCESS MONITORING, PROGRESS MONITORING

**How answerable is this question?**

The biggest obstacle to answering this question is that there is much less written and said about failures than successes. But in an era where many programs are or may soon be discontinued, a deliberate effort to conduct interviews and capture institutional memory would be exceptionally valuable.

WHAT MAKES IT TRACTABLE?	WHAT MAKES IT HARD?
<ul style="list-style-type: none"> <li>• PROGRAM DOCUMENTATION, GAO REPORTS, AND IG FINDINGS ARE PUBLICLY AVAILABLE</li> <li>• CASE STUDY METHODOLOGY IS WELL-SUITED TO THIS TYPE OF QUESTION</li> <li>• COMPARATIVE ANALYSIS ACROSS PROGRAMS IS FEASIBLE WITH EXISTING DATA</li> <li>• RETIRED AGENCY STAFF ARE OFTEN WILLING TO SPEAK CANDIDLY ABOUT PAST PROGRAMS</li> </ul>	<ul style="list-style-type: none"> <li>• FAILURE IS SYSTEMATICALLY UNDERDOCUMENTED – AGENCIES HAVE LITTLE INCENTIVE TO RECORD IT</li> <li>• SELECTION BIAS: WE KNOW MORE ABOUT PROGRAMS THAT SURVIVED THAN THOSE THAT DIDN'T</li> <li>• TRANSLATING CASE STUDY LESSONS INTO GENERALIZABLE GUIDANCE IS METHODOLOGICALLY TRICKY</li> </ul>

## Conclusion

---

The questions assembled in this agenda are not new. Researchers, policymakers, and funders have been asking versions of them – about what R&D investment produces, for whom, and at what cost – for as long as governments have invested in science at scale. What is new is the urgency with which credible answers are now needed, and the degree to which the absence of those answers has begun to constrain the field's ability to make and defend the case for sustained public investment in research. While [some progress](#) has been made toward better empirically understanding the returns to R&D, greater clarity is needed to support decision-making (and in fact, the [Pop-Up Journal](#) initiative aims to deliver that clarity in the next several years by summarizing existing and emerging research on this topic).

The agenda is deliberately ambitious. Several of its foundational questions – particularly those concerning measurement, attribution, and counterfactual construction – may never be fully resolved. The causal chains connecting a basic research investment to its ultimate social and economic consequences are long, indirect, and resistant to the clean identification that rigorous evaluation requires. Acknowledging those limits honestly is not a concession of defeat; it is a precondition for doing the work well. An evaluation field that overstates what it can demonstrate invites the kind of skepticism that is difficult to recover from, particularly in a political environment where research budgets are contested and the burden of proof falls increasingly on those who defend public R&D investment.

At the same time, the agenda reflects a conviction that significant progress is possible. The applied and structural questions – on instrument effectiveness, spillover measurement, crowd-in effects, portfolio construction, administrative efficiency, and the governance of long-term research programs – are tractable in ways that the foundational questions are not. Credible work on these questions could reach a wide range of practitioners within a reasonable timeframe. The persona mapping exercise that grounds this agenda makes clear that there is the type of demand that exists for exactly this kind of research across the ecosystem: from program officers who need better language for what their portfolios are trying to achieve, to staffers and advisors who need geographically specific evidence that can survive a budget hearing, to philanthropic funders who need to understand whether their dollars are building durable public commitment or quietly substituting for it.

Three priorities stand out as most deserving of early attention. First, the development of rigorous, transparent approaches to counterfactual construction – not because the problem is easily solved, but because the credibility of nearly every other finding in the field depends on how it is confronted. Second, the systematic measurement of social outcomes from public R&D investment, which remains the largest and least documented category of effects despite representing the primary justification for public funding in the eyes of many of the stakeholders this agenda is designed to serve. Third, the creation of data infrastructure that would make attribution feasible at scale; without it, the most important questions in this agenda will continue to be answered through inference and assumption rather than high-quality evidence.

This agenda was built through engagement with the people who would use this evidence in practice. That origin is both its strength and its intended animating principle. The goal is not to inspire research that is methodologically sophisticated in the abstract but practically irrelevant to the decisions it was designed to inform. It is to close the distance between what researchers can demonstrate and what the practitioners, policymakers, and funders who steward public investment in science actually need to know. That distance is real, and it matters. Closing it is the work this agenda sets out to do.

## **About the Federation of American Scientists**

The Federation of American Scientists is dedicated to democratizing the policymaking process by working with new and expert voices across the science and technology community, helping to develop actionable policies that can improve the lives of all Americans. For more about the Federation of American Scientists, visit **FAS.org**.