United States Department of Agriculture

USDA

Economic Research Service

Economic Research Report Number 95

May 2010

Assessing the Benefits of Public Research Within an Economic Framework

The Case of USDA's Agricultural Research Service

Paul W. Heisey, John L. King, Kelly Day Rubenstein, Dale A. Bucks, and Rick Welsh

Visit Our Website To Learn More!

For more ERS research and analysis
on Agricultural Research, see:

www.ers.usda.gov/Briefing/AgResearch

Cover photos credit: Agricultural Research Service, USDA.

Assessing the Benefits of Public Research Within an Economic Framework

The Case of USDA's Agricultural Research Service

Paul W. Heisey, pheisey@ers.usda.gov
John L. King, johnking@ers.usda.gov
Kelly Day Rubenstein, kday@ers.usda.gov
Dale A. Bucks
Rick Welsh

Abstract

Evaluation of publicly funded research can help provide accountability and prioritize programs. In addition, Federal intramural research planning generally involves an institutional assessment of the appropriate Federal role, if any, and whether the research should be left to others, such as universities or the private sector. Many methods of evaluation are available, peer review—used primarily for establishing scientific merit—being the most common. Economic analysis focuses on quantifying ultimate research outcomes, whether measured in goods with market prices or in nonmarket goods such as environmental quality or human health. However, standard economic techniques may not be amenable for evaluating some important public research priorities or for institutional assessments. This report reviews quantitative methods and applies qualitative economic reasoning and stakeholder interviewing methods to the evaluation of economic benefits of Federal intramural research using three case studies of research conducted by USDA's Agricultural Research Service (ARS). Differences among the case studies highlight the need to select suitable assessment techniques from available methodologies, the limited scope for comparing assessment results across programs, and the inherent difficulty in quantifying benefits in some research areas. When measurement and attribution issues make it difficult to quantify these benefits, the report discusses how qualitative insights based on economic concepts can help research prioritization.

Acknowledgments

Contents

Summary

Agricultural research managers continually seek to identify the benefits of research programs, both to demonstrate social and economic impact and to prioritize future research. Tight budgets for public agricultural research have increased the need to get the most out of these investments. Meanwhile, shifting policy goals have pushed public agricultural research in new directions. Although economic analysis can provide quantitative estimates of research benefits, it may not be amenable to all situations. However, economic reasoning may be useful even when formal economic methods are not used.

What Is the Issue?

USDA's Agricultural Research Service (ARS) accounts for the majority of Federal expenditures on agricultural research. The most common method of Federal research evaluation is peer review—widely used throughout the ARS prioritization, planning and evaluation cycle. Peer review's strength is assessing scientific merit. However, it is not well-suited for quantifying market impacts or ultimate social benefits of research programs. Economic analysis, by contrast, can address these benefits, but often at considerable cost.

What Did the Study Find?

The standard techniques of economic evaluation—especially econometric and economic surplus techniques – have usually been applied retrospectively, and at aggregate levels of analysis. Three case studies of ARS research—bovine quantitative genetics/genomics, water quality/watersheds, and nutritional composition of food—illustrate opportunities and limitations of applying existing methods of economic evaluation to individual programs of research, and demonstrate the value of economic reasoning even in the absence of quantitative analysis. These case studies were selected to capture research diverse in its nature (basic versus applied), program scope, coordination roles, and relevance to USDA missions. As such, the case studies are not meant to represent the full range of issues addressed by ARS; rather, they focus on some of the more challenging aspects of research evaluation.

Case Study Specifics

- *Bovine Quantitative Genetics and Genomics*. ARS researchers conserve animal genetic resources, identify genes related to economically important production traits in beef and dairy cattle, and estimate the heritability of desirable genes. The benefits of ARS research can be measured by the value of the increased productivity engendered by the research, so standard economic analysis is plausible. Nonetheless, allocating attribution among the various contributors to increased productivity adds to the complexity of the problem.

- *Water Quality and Watersheds*. ARS research on water quality and watersheds examines physical and chemical properties of water and agricultural pollutants, their impact on agricultural production and water

quality, and methods to improve agricultural and environmental outcomes. This research generates both private and social benefits. The benefits of research leading to the development of specific products or services, such as irrigation control equipment, are amenable to standard economic analysis. But the benefits of research on factors affecting water quality are harder to evaluate, requiring the valuation of nonmarket goods/services and linking research with public policies and regulations.

- *Nutrient Data Laboratory*. The Nutrient Data Laboratory (NDL) provides data on the nutrient composition of foods in the American diet. Its databases are the foundation of virtually all public and commercial nutrient databases used in the United States and a number of other countries. The primary social benefit of NDL's research is improved human health, but the link between nutrition information and health outcomes, while widely recognized, is not well understood. Coupled with difficulty in assigning a dollar value to health outcomes, these measurement and attribution problems make standard cost/benefit analysis problematic.

Economic reasoning can provide qualitative analysis even when quantitative estimates of benefits are intractable. For instance, market failures and the presence of significant cross-State application of research findings may lead the private sector and State-funded institutions to underinvest in economically justifiable research. All of the case studies show that ARS research provides scientific results with few substitutes. Because of their "public goods" nature, the benefits of these research programs would be difficult to replicate by non-Federal research efforts. Thus, qualitative analysis can provide a clear indication of the public goods characteristics of a research program, even if it cannot rank multiple projects or programs, all of which demonstrably address public goods.

Furthermore, interviews with stakeholders such as food processors, natural resource managers, universities, other federal agencies and international research institutions suggest that ARS research facilitates numerous linkages between peers, contributors, and users. While it is difficult to put a dollar estimate on these benefits, interest on the part of a diverse group of stakeholders may indicate a broad set of benefits.

Finally, public research often aims to enhance the operations of Federal regulatory agencies and strengthen the scientific basis of government policies. Assessing research programs aimed at improving government regulations generally exceeds the scope of economic analysis because they work through the political process. However, economic reasoning is useful for tracing the demand for and performance of mission-related research. In these cases, ARS research often contributed to the regulatory and policy functions of the Federal Government in ways that other public sector research institutions did not. For example, with passage of the Clean Water Act, ARS adapted its soil movement models to examine the effects of sedimentation on downstream water quality. The findings of the three case studies provide additional insights about important issues in assessing research benefits.

- Basic research and nonmarket goods. In Federal research evaluation, benefit-cost analysis has been used primarily for Federal research programs that produce specific, near-market technologies While some

ARS research produces near-market technologies, two of the case studies (and much of the third) involved research producing significant nonmarket benefits or basic research with no immediate market applicability.

• Attribution. Research builds on previous findings, and numerous related efforts are often being performed by other institutions. Attributing benefits to one group of researchers is imprecise at best. All three case studies featured significant numbers of different researchers, with numerous horizontal and vertical linkages.

How Was the Study Conducted?

This report focuses on the feasibility of using economic and other evaluation methods to value ARS research, rather than the estimation of quantitative values. This study combined two principal methods—detailed literature reviews, and in-depth qualitative interviews with ARS research administrators, scientists, and stakeholders. The interviews used an interview guide approach that featured structured but open-ended questions, with information-rich respondents queried until answers confirmed working conclusions rather than revealing new important topics.

Introduction

Agriculture is one of the oldest human endeavors, but it is constantly renewed through research and development. New technological opportunities, changing production practices, and evolving pests are just some of the reasons for research and development (R&D) spending. In 2000, worldwide agricultural research and development spending totaled approximately $37 billion. Of that sum, an estimated $23 billion (62 percent) was spent by the public sector and certain nonprofit research entities; the remainder was spent by private firms (Alston and Pardey, 2006). In the United States, the Federal Government performed or funded approximately $2.9 billion of agricultural research in 2008, and States funded an additional $1.5 billion (USDA/NIFA, 2010). U.S. private firms undertake a larger share (58 percent of the total in 1998) of agricultural research than the combined public sector investment (Caswell and Day-Rubenstein, 2006).

More challenging than tallying these research costs is estimating the benefits of scientific research, for a number of reasons. Scientific research is often technologically complex and reliant on earlier work, making attribution of benefits to any one project difficult. Also, research benefits typically occur with long lags, with adoption or commercialization often occurring many years after the primary scientific effort is over. The experimental nature of scientific inquiry introduces risk and uncertainty that requires additional analysis for evaluation of results. Furthermore, in the case of public research, which is often directed at intractable problems with few private sector alternatives, it can be difficult to distinguish the success of the research in terms of scientific merit from other obstacles that prevent its implementation.

Despite the challenges, assessing the economic and social benefits of publicly funded scientific research is essential for the purposes of accountability and planning. Federal research agencies are accountable for their use of public funds and need to justify their expenditures by demonstrating returns on research investments. A substantial body of literature concludes that the social benefits from public research in agriculture have been large, with social rates of return usually falling within the range of 20 to 60 percent annually (see Fuglie and Heisey, 2007, for a recent overview, and Alston et al. (2000) or Evenson (2001) for extensive surveys).[1] Returns to public research have also been high in other sectors such as health (Cockburn and Henderson, 2000). Although it might be reasonable to conclude that past investments in agricultural research have been socially beneficial, demonstrating accountability is an ongoing responsibility for publicly funded R&D.

Assessment of research benefits is also important for research planning. Research managers attempt to determine the best use of limited funds, and sometimes compare potential benefits with potential costs. They also need to assess the probability that the research will meet its stated objectives, and the possibility that other institutions, including private firms, might conduct similar research at lower social cost. As a research program unfolds, the planning process is also likely to include evaluation of performance and implementation of changes in direction. All these planning stages require an understanding of research benefits.

[1]Although a conversion from estimated social rates of return to dollar values should be treated cautiously, these rates of return suggest that, very roughly, every dollar invested in agricultural research returns $5 to $15 to the economy in present value terms.

This study examines approaches for assessing benefits of public research in agriculture. This includes a review of assessment activities at Federal research agencies, describing efforts to address the need for benefits estimation. This study also presents three case studies of research programs within USDA's Agricultural Research Service (ARS). We do not make quantitative estimates of the value of ARS research, but rather explore the feasibility of economic and other evaluation methods in these three cases. The report is organized as follows:

• Chapter 2 describes the agricultural R&D system, discusses why it differs from many other R&D sectors, and offers possible explanations for why the public sector plays such a large role in agricultural R&D. It ends by discussing how views of the appropriate public sector role feed into how public R&D investments should be evaluated.

• Chapter 3 reviews the standard economic approaches for evaluating public R&D—econometric analysis and economic surplus methods—before looking at what, if any, economic analysis might be applied to different levels of research aggregation and discussing alternative methods of evaluating research when suitable economic data are not readily available.

• Chapter 4 describes the research evaluation methods typically used by Federal agencies and reviews the available literature on their application. ARS's organization and mission are then summarized, along with its planning activities and evaluation procedures.

• Chapter 5 presents the three case studies—covering ARS research on bovine quantitative genetics/genomics, water quality/watersheds, and the Nutrient Data Laboratory—to determine how research evaluation, and particularly economic evaluation, might be applied in each case. While quantitative analysis is not attempted, each case study identifies the economic methods that hold promise, as well as the complications that would need to be addressed.

• The main body of the report closes with a summary in Chapter 6, followed by an Appendix that discusses the challenges of and potential extensions to standard economic methods for evaluating scientific research.

Assessing the Economic and Social Benefits of Public Agricultural Research

To discuss appropriate ways to assess the benefits of public research activities, it is first necessary to understand the role of public research within the agricultural R&D system. Public sector agricultural R&D investments are not made in a vacuum; rather, they occur in a system in which private sector firms, universities, and other nonprofit institutions participate. In many cases, researchers share educational and training experiences and might work on different aspects of the same technological problem. Researchers often find it expedient to form research partnerships with other institutions; for example, firms form research joint ventures and some Federal laboratories are located on university campuses. As a result, flows of research funding and research personnel among institutions are common. Given the complexity of these relationships and the cumulative nature of scientific investigation, the results of public sector R&D depend on the functioning of the entire system.

Aside from funding flows, working relationships, and employment mobility, other activities unify different sectors and disciplines of the research system. Norms of science, such as peer review and presentations at scientific conferences, are examples of such activities. Still, researchers perform differently according to incentives provided by their institutional settings. For example, the way firms and Federal research agencies acquire resources is very different. Firms raise funds in capital markets, perform research, and attempt to appropriate the resulting benefits. Federal funding of research that might be difficult to commercialize is determined through an entirely different appropriations process.

Special characteristics of agricultural research. The interdependence of public, private, and university scientists within a research system is common to most fields of scientific inquiry, but research in agriculture has some distinctive characteristics with implications for assessing benefits. Alston and Pardey (1996) note the biological nature of agricultural research and the atomistic, spatially diffuse structure of agricultural production.

The biological nature of agricultural production. Agriculture is based on the growth of living organisms, which implies greater unpredictability than dealing with inert matter. Some aspects of this feature, such as the importance of epidemics and disease, share elements with human health research; others, such as the influence of weather, are particular to agriculture.

The reproductive characteristics of different domesticated crop and livestock species are biological aspects of agriculture that create risk for commercial inventors. For instance, the genetic information in improved seed varieties that are developed through research investment can often be reproduced simply by replanting the seed. As a result, plant breeders may not be able to appropriate enough returns from seed sales to make seed development profitable, resulting in less than socially optimal private investment in crop improvement research. Depending on the crop, the ability to create economically viable hybrid varieties, or other technical factors, can create much greater incentives for private firms to invest in plant breeding (Fuglie et al.,

1996; Morris et al., 1998; Heisey et al., 2001). For livestock, differences in fecundity rates and gestation periods make private sector investment far more profitable in poultry than in cattle, with swine investment moderately profitable (Narrod and Fuglie, 2000).

Changes in science, such as shuttle breeding and biotechnology, have reduced agriculture's dependency on strictly biological factors over time. Similarly, other technical and policy changes, such as hybridization or increasing application of intellectual property rights to biological technologies, have created more incentives for private firms to invest in agricultural research. Though these developments have led to significant increases in private investment, they have not eliminated the uncertainties of biology from agricultural research.

Firm size in agriculture. Relative to other industries, the agricultural sector is characterized by numerous, decentralized producers. According to the 2007 Census of Agriculture, 2.2 million farms sold approximately $297 billion of agricultural products, or $134,807 per establishment (USDA/NASS, 2009). In contrast, the 2007 Economic Census for all U.S. establishments with payrolls, which altogether represent a majority of economic activity in the United States, found 7.2 million establishments with average sales of $4.0 million (U.S. Census Bureau, 2009). The atomistic or fragmented nature of production agriculture usually does not carry over to the firms in the sectors that supply inputs to agriculture, nor to the processing firms that buy agricultural products (which are included in the Economic Census payroll survey). For example, according to annual reports, Monsanto's seeds and genomics division had sales of around $4 billion in 2006; Archer Daniels Midland, a leading producer of agriculturally derived products, had annual sales of around $35 billion in 2005.

Alston and Pardey (1996) suggest several reasons why the industry structure of agriculture affects its research system. Because of its atomistic structure, individual farm operators are unlikely to realize the market share and industry dominance necessary to recoup R&D costs. As a result, "individual farmers in such settings are more inclined to confine their R&D investments to searching and screening activities, introducing and perhaps adapting technology that has been developed elsewhere, and developing ideas, methods, and information specifically relevant to their individual situations" (Alston and Pardey, 1996, p. 235).

Spatial diffusion of agriculture. The importance of land as a factor of production in agriculture is one reason for its atomistic industry structure. Land used for agricultural purposes totaled nearly 1.2 billion acres in 2002, over 50 percent of the total U.S. land area (Lubowski et al., 2006), while the Census Bureau estimated over 938 million acres as "land in farms" (USDA/NASS, 2004).[2] [3]Across such a large geographical area, different crops and production practices are best suited to specific locations because of differences in weather, soils, pest and disease pressure, and day length; this accounts for the location-specific nature of agricultural R&D. As such, applicability of new technologies "may vary significantly among firms, even across short distances, adding to the costs of decision making" (Alston and Pardey, 1996, p. 237). As a result, even successful agricultural research is fragmentary in its practical application.

[2]The difference between the two estimates is largely caused by grazing lands which the Census of Agriculture does not define as being in farms, and some underestimation of cropland in the Census of Agriculture (Lubowski et al., 2006)

[3]Agricultural land use in the 48 contiguous States is a higher percentage of total land area.

The degree of location specificity may vary by commodity; for example, livestock research may be less location specific than crop research (Evenson, 1989; McCunn and Huffman, 2000). Even for crops, biological factors can make research for some crops (e.g., wheat) less location specific than for other crops (e.g., corn or potatoes) (Byerlee and Traxler, 2001; Maredia and Byerlee, 2000; Griliches, 1957; Walker and Fuglie, 1999).

Extensions beyond production agriculture. Most of the discussion so far has focused on research directed at production agriculture. Increasingly, however, agricultural research institutions have focused effort on areas such as environmental improvement, human health and nutrition, and rural development. Thus, the public agricultural research sector, as a whole, has a broad portfolio of research activities with objectives that go well beyond production agriculture.

The Roles of Public Research

Prior to assessing the benefits of research, it is important to establish what goals public research is intended to accomplish. While some research projects and programs might have easily identifiable objectives, methods, outputs, and applications, the link between research intent and research outcomes is not always so clear. Some research might be more exploratory, and other research may not be directed at a final product, but toward development of a research tool, such as the RUSLE2 model for predicting annual soil loss. With so many confounding factors, it is helpful to draw some generalizations about the intended roles of public research and to create categories for different goals and rationales.

Three paradigms of public research policy. Bozeman (2000) argues that the history of U.S. public policy toward R&D can be conceptualized under three paradigms: market failure, mission orientation, and cooperative technology. In this view, the three paradigms are defined primarily by the extent of public sector involvement. The market failure paradigm requires the least active public role, with public policy directed toward filling gaps left by the predominant private sector. The cooperative technology paradigm has the most central public involvement, with the public sector choosing research areas, marshalling resources, providing incentives for complementary private sector involvement, and brokering the development of broad areas of research and commercial activity. The mission orientation paradigm is intermediate, entailing an active public role within narrowly defined agency missions.

Market failure. The United States has a primarily market-based economy, and a central tenet of economics is that incentives for the productive and efficient use of assets are strongest when people act on their own behalf. However, individual incentives often disregard the full impact—beneficial or detrimental—of personal decisions on the rest of society. The aggregation of individual decisionmaking by markets can lead to the inefficient use of economic resources, with significant effects on economic growth.

One type of market failure is the failure to produce public goods (Samuelson, 1954). Nelson (1959), Arrow (1962), and others provide economic analysis of the public good nature of scientific knowledge, which shares the two

defining characteristics of public goods. First, information produced from scientific R&D is nonrival in consumption, meaning that it can be used simultaneously by many people.[4] In fact, the value of information can increase if more people possess and act on it at the same time. Second, the cost of duplicating information is low compared to the costs of creating it. Although generating research results often involves significant costs in the form of scientific and technical labor, computing power, experimental facilities, and materials, communicating those research results through journal articles, research reports, or Internet publication is relatively inexpensive. The related difficulty of excluding others from accessing research results[5] makes information (and other public goods) less profitable to produce, even though information is valuable to society.

This mismatch of private incentives and public benefits results in a market failure wherein the private sector undersupplies scientific research. One type of market failure is in the quantity of overall R&D spending: private incentives fail to produce R&D commensurate with potential social impact. A second type is when the composition of R&D is directed away from areas of science where the goods produced are themselves prone to market failure, as with environmental goods such as clean air or water (Popp, 2006). The role for public sector research suggested by these twin market failures is both to increase the amount of R&D in society generally and to target research toward public goods.

Several public policies address market failure in the quantity of research. First, the Federal Government performs significant amounts of research, totaling $24.4 billion or 7.1 percent of the $342 billion spent on research in the United States in 2006 by all institutions—private firms, universities, and the Federal Government (NSF, 2007). Second, the public sector supports an even larger amount in the research system: public sector-sponsored research in industry, universities, and federally funded R&D centers accounted for $72.4 billion (21.1 percent) of research spending in 2006 (NSF, 2007). This support typically comes in the form of research grants. The public sector also increases private sector research through the establishment of intellectual property rights and occasionally through tax policy regarding research investment.

Public provision of R&D addresses both quantity and composition aspects of scientific research. Grants and other research funding address the total amount of research activity in society, but offer weaker control over composition because grant recipients exercise some discretion over carrying out research.[6] Intellectual property rights provide incentives for increasing private sector R&D, but their generality permits very little guidance of the composition of R&D activities.

Mission orientation. The mission paradigm assumes that government agencies have well-defined missions arising from identified public needs. The Government may perform R&D in service of these missions "in which there is a national interest not easily served by private R&D" (Bozeman, 2000). Examples of mission-oriented research include national defense— dating back to the sponsorship of the National Armory beginning in 1777 (Hounshell, 1984)—energy production, and space exploration. In agriculture, explicit missions of the Federal Government include customs inspections,

[4]Thomas Jefferson, writing on this theme nearly 200 years ago, noted that, "He who receives an idea from me, receives instruction himself without lessening mine; as he who lights his taper at mine, receives light without darkening mine."

[5]This property of public goods is called "nonexcludability" by economists.

[6]Research prizes are an alternate type of funding awarded to the first successful demonstration of a technological feat. Research prizes probably exert greater influence on the composition of resulting research, but they are uncommon in practice despite some contemporary and historical examples (Wright, 1983; Sobel, 1995).

regulating food safety, and preventing outbreaks of plant and animal disease. Historically, certain activities such as seed importation and classification were carried out by the Agricultural Division of the Patent Office even before the establishment of the U.S. Department of Agriculture in 1862 (Huffman and Evenson, 2006). Research oriented toward public agencies aims to provide program administrators and regulators with improved tools, data, and technology to address their missions, and to develop cheaper and more effective ways of fulfilling mandates of the Federal Government. In some areas such as environmental policy, direct regulation or the provision of economic incentives for conservation, rather than research, often plays the primary role, but research may inform regulatory and program design (Abler and Shortle, 1991, 1995).

Cooperative technology. The theoretical underpinning of the cooperative technology paradigm comes from industrial policy theory and regional economic development theory. In this view, governments may be more efficient than relying solely on markets to spur sustainable economic growth. Government policy tools such as tax breaks, competitive research grant programs, cooperative research agreements, and other combinations of public and private funds are used to steer private sector research and investment in particular directions. Federal laboratories take leadership roles within economic sectors to spur economic growth and development, essentially "jump-starting" a new technology area by assembling all the pieces necessary for multiple actors to make a mutual decision to move forward. Cooperation rather than competition is assumed to be particularly desirable in early stages of technology development (Larsen and Wigand, 1987; Wigand and Frankwick, 1989; Link and Tassey, 1987).

The cooperative technology paradigm has had greater influence in Europe and Japan than it has in the United States. However, the cooperative technology paradigm was a significant influence on policy changes that began in the 1980s, including changes in intellectual property policy. A number of major congressional measures for technology transfer, including the Bayh-Dole and Stevenson-Wydler Acts of 1980, were predicated on the assumption that federally sponsored or federally performed research was constrained from reaching industrial partners. By creating patent rights for that research, these policies intended to increase benefits from collaboration with the private sector. The creation of the Small Business Innovation Research program in 1982 and Cooperative Research and Development Agreements (CRADA) in 1986 were other U.S. policies aimed at increasing coordination between public and private sectors.

Any research organization, particularly if it is large, is likely to fit more than one of the three paradigms of public research as set out by Bozeman. Research is often intended to serve more than one goal, and the three paradigms described above overlap considerably. In practice, research in Federal agencies fits into multiple paradigms at once. Crow and Bozeman (1998) admit that these three paradigms promote "stereotypes," and they provide examples of hybrid organizations with mixed goals that fall into multiple paradigms.

Beyond Bozeman's three paradigms of public research, another way to categorize the roles of public research begins with properties intrinsic to the

pursuit of scientific research. The scientific method and the process by which scientific research results are transmitted throughout society suggest categories for public sector research roles.

Basic and applied research. Scientific investigation builds cumulatively, proceeding from previously established results. In some instances, the value of earlier research is not fully appreciated until subsequent applications are developed. This aspect of scientific research suggests a linear model that distinguishes between the "two types of work commonly referred to as basic and applied research" (Bush, 1945). In this model, basic research that "involves long-term investigation of a fundamental nature" has few direct applications to societal problems. But basic research can lead—directly or indirectly—to future discoveries that do have useful applications. Many of these applications will have prospects for commercial success and will attract private sector investment. But the inability to predict direct commercial benefits from basic research--even when it is reinforced with intellectual property rights, which are limited in duration--suggests a public role in providing basic research as a necessary precursor for socially useful applications.

Stokes (1997) refined this linear model of basic and applied research by distinguishing the motivation with which each research endeavor is undertaken. In particular, Stokes asks first whether research is motivated by greater general understanding of phenomena and second whether research is motivated by potential uses of the research. In the resulting four-quadrant model, there is no contradiction between "basic" research that is oriented toward greater general understanding of the natural world and simultaneously "applied" to potential uses. He references the work of Louis Pasteur as an example of research that is simultaneously fundamental (e.g., disproving spontaneous generation in favor of the germ theory of disease) and applied (e.g., improving food safety and decreasing spoilage). Thus, the Stokes model suggests that the public sector can have an important role in applied research without abandoning its role in providing basic research, especially if the private sector does not fully pursue a particular area of research or its more general aspects.

Risk and uncertainty. Inherent in the experimental nature of scientific research is the presence of risk and uncertainty. When allocating resources for investing in research, the prospective benefits need to be discounted to reflect the fact that research costs and benefits occur over time. Fuglie et al. (1996) and Tassey (1997) discuss proper accounting of discount rates for R&D. Furthermore, risk preferences in the face of uncertainty may be incorporated into the choice of an appropriate discount rate. Where risk is a systematic feature, as in R&D, Mendelsohn (1981) argues that the public discount rate—rather than the normally higher private (or market) discount rate—is the correct way to discount public investments.

Areas with a marked divergence in the public and private discount rate—and the biological, spatial, and atomistic industry characteristics of agriculture might be reasons for such divergence—create additional justification for a prominent public role in funding or performing research. The fragmentary nature of agricultural practices emphasizes another risk-reducing aspect of public research. Because of the location-specific nature of agricultural research, its results may be applicable only in certain regions or organisms.

The Federal Government, with more expansive research programs than those undertaken by any other single entity, is more likely to find a useful application of a particular research result.

The Federal Government has a lower cost of bearing risk, reflecting its lower borrowing costs relative to private firms. This lower cost derives from its taxation authority, but also from its ability to pool risk across all citizens. In turn, one role of public research in agriculture is to develop technologies that decrease risks of agricultural production and consumption. Because of the wide variety of agricultural production and the universality of food and fiber consumption, public sector research is likely to result in a wide range of benefits.

Coordinating role for public research. Because the public sector has such a large role in performing and funding research, it can often be uniquely positioned to coordinate research effort. In agriculture, the public sector funds or performs a plurality of R&D.[7] As a result, public science agencies can take advantage of their predominant role in the overall research system to prevent duplication of research effort and to promote the dissemination of new research results.

There are two primary benefits from coordination of research. The first is to reduce duplication of effort. "Winner-take-all" incentives may occur across all levels and institutions of the research system. Incentives provided by intellectual property policies, and norms of science that award large benefits for scientists who are first to arrive at a scientific result but few awards to other scientists, can be wasteful. Loury (1979) and Dasgupta and Stiglitz (1980) note that duplication of research effort in patent races can dissipate the entire value of the potential benefits if the total research costs of competing firms outweigh the benefits provided by the patent winner (see also Shapiro, 2007). By allocating researchers among different research projects and limiting the number of grants awarded toward a specific research problem, public sector research managers can limit the duplication of research effort.

In industries where the public sector does not perform a plurality of research, it still may serve a coordinating role. In the U.S. semiconductor industry, standard-setting organizations help to reduce duplication of effort, although problems can arise from voluntary participation, hold up,[8] and antitrust concerns. In industries such as telecommunications and transportation where infrastructure requirements create an incentive to reduce industry fixed costs, the public sector usually plays some role in regulating or coordinating industry effort.

Another coordinating role for public research is the dissemination of research results. There can be institutional barriers that prevent the timely spread of information. Research across different scientific fields, with different applications or requiring multidisciplinary perspectives, can produce useful results that go unnoticed. In areas of common public mission, different agencies can coordinate efforts and form long-term working relationships to facilitate information sharing. For example, solving water quality problems related to agricultural activity may require other disciplines in addition to environmental science, such as crop or livestock science, soil physics, soil chem-

[7]Collectively, firms in the private sector perform more R&D, but each individual firm performs less than the public sector.

[8]"Hold up" in this instance refers to a situation in which the owner of a patent begins to charge high royalties when a technology using that patent is chosen as an industry standard.

istry, engineering, and computer science, as well as behavioral sciences such as economics. Furthermore, different Federal agencies have water quality mandates. Public coordination can help bring the relevant agencies and scientific disciplines together.

Microeconomic models of strategic information sharing also suggest rationales for coordinating research and data collection. Whereas most physical goods increase in value with scarcity, information goods can be more valuable when they are shared by many people. Competing firms will tend to share information (including data and research) depending on how information sharing affects the correlation in strategic decisions, and whether correlation of strategic decisions is good or bad for individual actors (Vives, 1984; Gal-Or, 1986). The public sector role can be to facilitate information sharing when it is voluntary, providing a clearinghouse and credible platform for verifying information.[9] When private firms find it individually profitable to withhold information, the public sector can step in to increase competition by creating and sharing information. For example, agricultural extension can promote information sharing and correlation of production techniques through the sharing of research results.

[9]Trade associations are another common venue.

Critical Perspectives on the Roles of Public Research

The roles of public research do not suggest the practical difficulties in managing it for lowest cost and prioritizing its most beneficial research targets. Each of the twin decisions on the quantity and composition of public sector research spending poses difficulties. It is difficult to establish the proper extent to which the public sector should perform or fund research above the amounts that the private sector would sponsor in the absence of government intervention. Marglin (1963) notes the tendency to stop well short of funding all possible investments with an adjusted positive rate of return after discounting for risk and time. Firms presumably face the same difficulties, but competition provides clear price signals and strong informational incentives.

With regard to the composition of R&D funding, research in areas of science that themselves are prone to market failure, such as environmental or nutritional science, often lacks market prices that communicate an economically efficient amount of investment. Difficulties include determining the research priorities to emphasize, setting an appropriate incentive that is large enough to be effective but small enough not to be wasteful, and coming up with the exact specifications that need to be met to indicate that the research has been successful.

Public research and the political process. Political economy and public choice theory emphasize that government policies are the result of a political process that is subject to the influence of interested parties. While the public research system interacts with other research organizations, including private firms, its inherent public roles dictate that it adhere to potentially more socially beneficial arrangements. The danger is that a single interest group could "capture" public research and direct it to the interests of that group instead of toward societywide benefits (Huffman and Evenson, 2006).

Research in support of broad public missions may be specified as the outcome of a political process. Since taxpayer funding is required for most government-sponsored research, accountability is an issue. Ideally, decisionmaking for resource allocation and research direction at different levels might be vested in fully informed decisionmakers, with scientific considerations afforded greater weight the more disaggregated the research program or project (Alston and Pardey, 1996). In many cases, there is limited public input and government missions may be assessed mostly through the technical capacities of experts. On the other hand, political economy perspectives suggest well-organized interest groups might play a large role in determining stated government missions. In agriculture, as in many other research areas, the relative influence of interest groups, the research bureaucracy, and scientific experts is contested, and "many cross-cutting, complementary, and contradictory forces help shape priorities and resource allocations" (NRC, 2003).

Research outcomes and social evaluation. From the perspective of public policy, evaluating large-scale, long-term societal outcomes might be important. Current research assessment efforts tend to focus on *outputs* rather than *outcomes*, in large part because the former are more easily and more quickly available.

Sarewitz et al. (2004) exemplify an attempt to develop an analytical framework that can shed light on interrelationships between public science policy decisions, scientific discoveries, technological change, and societal outcomes. They assert that most attempts to assess science and technology policy only consider benefits and not costs or downsides. For example, if science and technology policy can be linked to economic growth within a nation or region, then the same policy can also be associated with any negative social outcomes from that growth. However, such accounting is rarely done. Sarewitz et al. (2004) make the point that technological innovation and the resulting economic growth in the United States have also been blamed for reduced employment and increasing global inequality. "Science and technology policy are obviously not themselves directly responsible for rising inequality or unemployment," but "neither are they directly responsible for economic growth."[10]

Ultimately, Sarewitz and colleagues (2004) wish to see science policy broadened to encompass knowledge use as well as its creation. Taking into account the perspectives and needs of the potential users of knowledge recognizes that scientific research and technology development trajectories are not independent from social forces but rather are the product of social, economic, and political influences. Furthermore, the development of physical technologies toward certain social outcomes is contingent on the concurrent development of "social technologies," or the capacities of potential users of that physical technology. For example, the eradication of smallpox was made possible by the development of the appropriate type and form of vaccine (physical technology) in conjunction with monitoring and assessment activities (social technology). Social engagement is therefore a key feature of reforming science policy and harnessing it more directly for positive social outcomes.

Bozeman (2003) proposes a concise way to interpret the social outcomes of public policies. Beginning with a "public value" that is a desirable

[10]These considerations are related to the broad question of attribution in research evaluation, to which we return below. In general, it is more advisable to address multiple social objectives through multiple public policies, and not through research policy alone (Alston et al., 1995). Determining the appropriate policy mix and determining the extent to which research policy should take the credit for positive social outcomes, or the blame for negative social outcomes, are not easy tasks.

outcome based on social consensus, the Public Value Mapping (PVM) framework investigates the reasons for success or failure in realizing that public value. When private sector market forces are chiefly responsible for creating a public value--for example, the efficient creation and distribution of commodity goods--it is a "market success." Citing Bator (1958), Bozeman (2002) argues that efficient markets may not always be the most desirable vehicles for delivering public value, and that public failure is a possibility even when markets are functioning efficiently. For example, if the production of a commodity results in excess pollution beyond the efficient amount justified by public value, it is a market failure. Absence of a policy to address this market failure, or a policy that permits too little or too much of the good in question, is a public failure. Public policies that align outcomes with public value can increase public success.

Bozeman cites the tobacco industry in 1950 as an example of an industry with high market success and high public failure. That is, tobacco product sales were high and tobacco firms were profitable, but public health was compromised. Assessing the same situation in the late 1990s, he concludes that government regulation in the interest of public health values had created a "moderate public and market outcome" (Bozeman, 2002, p. 156). In this interpretation, government intervention through regulation, investment, or other action does not merely fill gaps the market has left, but instead adds positive value to social outcomes. Bozeman's example of a public and market success was the commercialization of the Internet. After significant public investment from the Department of Defense and the National Science Foundation to meet national security and scientific research needs, the Internet was extended to commercial use. Examples of public value success and market failure include environmental regulations, polio inoculations, and higher education. Examples of public value failure and market failure include elementary/secondary education and municipal public works. Bozeman contends that in these final cases, public institutions failed to provide all characteristics society might value, while markets lacked incentives to provide them. He argues that the market failure approach for guiding public policy development and management can be complemented by a public value failure approach, though the latter cannot wholly substitute for the market failure approach.

Methods for Assessing Economic Benefits of Public Agricultural Research

The economic value of research is the benefit resulting from research over what would have occurred in its absence, net of the costs of doing the research. Research does not occur in a vacuum, but should improve upon prevailing conditions. Analysts of research benefits must construct a plausible scenario of what would have happened to the benefits in question had the research being evaluated not been performed, and use this scenario as the basis for comparison.

A standard scenario for the assessment of public agricultural research is that the resulting knowledge leads to new or improved technology that farmers adopt. This adoption raises the productivity of inputs employed in agricultural production. The net result is lower costs, higher production, exit of resources such as labor from agriculture, or some combination of these changes (Fuglie and Heisey, 2007). Or new data or models might help policymakers better oversee agricultural activities.

When markets have observable prices that reflect costs of production and benefits to consumers, standard economic methods can be used to measure net benefits of research and to estimate the improvement they represent. This chapter reviews these standard approaches and looks at what, if any, economic analysis might be used to evaluate research when suitable economic data are not readily available for a full-scale analysis.

Standard Economic Approaches To Measuring Research Benefits

Economists traditionally have applied one of two major methods to evaluating agricultural research. The first approach is through econometric or statistical analysis and the second through economic surplus, or project evaluation, methods.[11] This distinction is maintained both in the standard text on economic evaluation of agricultural research (Alston et al., 1995), which thoroughly reviews both approaches, and in major literature surveys (Evenson, 2001; Alston et al., 2000).

Both methods have to answer the same kinds of questions regarding measuring the benefits and costs of research, constructing appropriate counterfactual scenarios, and addressing problems in linking benefits and costs, such as correctly specifying research lags and accounting for research spillovers and other questions of attribution. Whereas econometric analysis relates some measure of costs to some measure of benefits via statistical estimation, economic surplus analysis relates costs to benefits synthetically. Both methods, through defining and measuring economic benefits, use some welfare concepts, but economic surplus tends to rely more explicitly on ideas from welfare economics, as through the use of consumer and producer surplus.[12] Likewise, even though economic surplus models do not relate research benefits to research costs statistically, some of the parameters they use, such as supply and demand elasticities, may be obtained through statistical estimation.

[11]Occasionally *non-parametric* methods have been used to analyze research impacts (e.g., Chavas and Cox, 1992). These methods use economic properties but do not specify functional forms as do most econometric models.

[12]Economic surplus methods are fairly closely related to cost-benefit analysis.

Econometric/statistical analysis. Econometric models such as multivariate regression estimate statistical relationships that include economic behavior. The ability of econometric models to account for a variety of changes simultaneously makes them very useful for research assessment. Not only can they estimate a baseline or counterfactual scenario to model how variables would look without research, but deviations from this estimate represent one way to assess the benefits of research.

Econometric approaches can be for direct estimation of production functions, which model agricultural output quantities as a function of inputs; response functions, which model the factors that influence an important variable such as crop yield; or productivity functions, which estimate output per unit of input, especially at the aggregate level (see Fuglie et al., 2007, for an overview of agricultural productivity measures). Econometric approaches can also be indirect, using models of firm and producer behavior to incorporate changes in economic data. An indirect approach might involve treating the sector being analyzed as a firm and estimating an aggregate profit or cost function. In either type of approach, econometric models include data to measure research activity, and the estimated models are used to infer the impact of research on output or productivity. This isolates the effect of research from other influences, even as it models how research interacts with other aspects of the agricultural economy (see box, "Using Econometric Analysis To Analyze Research Spillovers and Returns to Agricultural R&D").

Economic surplus methods. Economic surplus methods, based on some of the tools of welfare economics, look at how the interaction of supply and demand determines the value of market transactions. Applied to research assessment, economic surplus methods show how research affects supply, demand, and their resulting market outcomes. By augmenting product qualities that increase demand, or lowering costs of production that increase supply, research can expand markets and change the economic rewards for market participants. Surplus measures are based on the differences between actual prices and prices that consumers would be willing to pay, or producers would be willing to accept, for a given good.

Economic surplus methods begin with supply and demand curves in the market in which research-based technological change takes place. Technical change that reduces supplier costs reduces the lowest price at which quantities can be supplied. This increases the volume of transactions and the economic surplus available to divide among suppliers and demanders. Consumer and producer surplus measures are well-defined areas (fig. 1) showing market supply and demand, and benefits from research are measured by changes in these areas.[13] Analysts can then compare the gain in surplus against the costs of the research necessary to create it, and assess the benefits of research to market participants. A similar analysis is possible for research that increases product quality, which increases demand instead of supply (see

[13]The figure gives one special case, that of a parallel shift in the supply curve. Other shifts, such as pivotal shifts, are sometimes used, and different formulas apply for calculating producer and consumer surplus in these models. The chosen supply shift is an important assumption of the model; it is difficult to ascertain what kind of supply shift is most appropriate either theoretically or empirically (Alston et al., 1995).

Figure 1
Change in economic surplus resulting from a research-induced supply shift
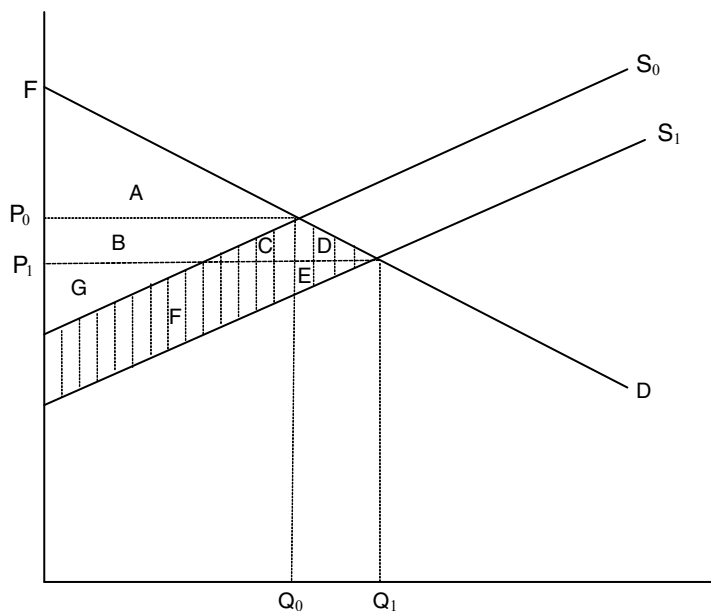
Change in consumer surplus: B + C + D
Change in producer surplus: (G + F + E) − (B + G) = F + E - B
Change in total surplus = Change in consumer surplus + change in producer surplus
$$= (B + C + D) + (F + E − B)$$
$$= C + D + E + F \text{ (shaded area)}$$

Using Econometric Analysis To Analyze Research Spillovers and Returns to Agricultural R&D

Gopinath and Roe (2000) used econometric analysis to analyze research spillovers and both private and social rates of return to research in three aggregate, vertically linked U.S. sectors: food processing, primary agriculture, and farm machinery and equipment. They focused on private sector research in food processing and farm machinery and public sector research in primary agriculture between 1960 and 1991. In addition to outputs, factor use, and factor prices, the stocks of research capital for each of the three sectors were also used as variables in the estimated equations.

Their theoretical model led to a system of 12 nonlinear equations, estimated using sets of pairwise equations for two of the three sectors at a time. The dependent variables for each of the unit factor demand functions were labor/output, capital/output, and inputs/output, with output for each sector measured in dollar terms. Simple calculations, or economic theory, were used to derive measures of labor, physical capital, and material inputs, used in the dependent variables, and factor prices, used as explanatory variables, from data on total expenditures on each factor as well as information on variables such as number of agricultural laborers. Some of the most important explanatory variables were measures of research capital, derived from data on annual research expenditures in each of the three sectors following a procedure outlined by Hall (1993).

Gopinath and Roe's procedure allowed them to estimate private and social rates of return to research in each sector, with and without taking research spillovers from sector to sector into account. Their estimate of the direct rate of return to public R&D was about 37 percent without spillovers, and about 46 percent when research spillovers into the other two sectors were taken into account.

box, "Using Economic Surplus Analysis To Estimate the Rate of Return to Kansas Wheat Breeding").[14]

Cost-benefit analysis is closely related to economic surplus analysis. It can be applied to many different types of investment projects, not simply research investments, which is why some authors use the terminology "project investment methods." In many applications, cost-benefit analysis, either explicitly or implicitly, has amounted to the assumption either that extra production in a market is measured by an outward shift of a vertical supply curve against a horizontal demand curve, or that the value of inputs saved is measured by the downward shift of a horizontal supply curve against a vertical demand curve (Alston et al., 1995). However, there is no reason why cost-benefit analysis cannot use more sophisticated specifications based on theoretical welfare economics (Just et al., 1982; Boadway and Bruce, 1984). Cost-benefit analysis is usually directed at a single investment project; the "project" in an economic surplus analysis of research investment is likely to be an entire program of research, at a larger scale than any individual research project.

Using Economic Surplus Analysis To Estimate the Rate of Return to Kansas Wheat Breeding

Barkley (1997) used economic surplus analysis to analyze the economic impact of the wheat breeding program at the Kansas State Agricultural Experiment Station. He constructed a two-sector model featuring supply and demand and producer and consumer surplus for Kansas and the rest of the world (ROW), modifying a model from Alston et al. (1995).

Barkley's model required data on market variables and market parameters, e.g., wheat production and demand for both Kansas and ROW, wheat prices, and supply and demand elasticities (a supply (demand) elasticity measures the percentage change in supply (demand) given a 1-percent change in price); the yield gains resulting from the Kansas wheat breeding program; and wheat breeding research expenditures. Supply and demand elasticities for Kansas and ROW were obtained from previous studies. Barkley used experimental wheat yield trial data from 1978 to 1996 to construct an index of varietal improvement following the methodology of Feyerherm et al. (1984), Brennan (1984), and Zentner and Peterson (1984). The ultimate objective was to measure the percentage shift in cost savings in each year. It was calculated by dividing the annual increase in Kansas wheat production by the elasticity of supply for 1979 to 1994.

Finally, Barkley aggregated data in real terms for wheat genetics research expenditures from the State of Kansas, the Kansas Wheat Commission, USDA/CSREES, and USDA/ARS through cooperative arrangements. He used an estimate of research lags for wheat breeding obtained from Kansas Agricultural Experiment Station scientists, requiring research expenditures over the previous 17 years.

The model found that the major beneficiaries of Kansas wheat improvement research were Kansas wheat producers who adopted the new varieties. Kansas consumers (wheat millers) were also made better off, but only by an extremely small percentage (0.04 percent) of the value of wheat purchased. There was also an annual transfer of economic surplus of approximately $41 million from non-Kansas producers to ROW consumers, caused by the small decrease in the world price of wheat. Overall, Barkley estimated a benefit-cost ratio of nearly 12 for this research, or an internal rate of return to the program of 39 percent.

Concepts such as net present value, benefit-cost ratio, and internal rate of return, which often appear in the cost-benefit literature, can also be applied to economic surplus analysis.

Advantages and disadvantages of each method. Econometric methods have a number of advantages (table 1). First, they are usually applied at a fairly high degree of aggregation--for example "all agricultural research," "all crops research," or "all wheat research." As such, they may be more likely to measure impacts taking into account a broad sweep of research, including both research successes and failures. This is an improvement over anecdotal evidence, which might rely too much on the circumstances surrounding a narrowly defined research project. Also, because of the degree of aggregation, econometric models can capture the effects of minor improvements in farming practices resulting from research that might be overlooked in analyses that look at major technological changes. The cumulative effect of minor productivity improvements can be significant over time, and econometric models are well suited to capturing gradual changes in farm organization, scale, and specialization that derive from research and technological change (Huffman and Evenson, 2006).

Second, statistical approaches can also account for the concurrent effects of other public investments on agricultural productivity, and therefore may be less likely to assume all changes in agricultural output or productivity result from research alone (Alston et al., 1995). These other investments may include investments in farmer education, in infrastructure (for example, roads or communications networks), or in farm programs generally. Likewise, using econometric models to control for other influences prevents analysts from relating outcomes to research costs when success or failure may be due to unrelated events.

Table 1

Advantages of econometric versus economic surplus methods

Econometric methods	Economic surplus methods
Reflect economic impacts, taking into account research "failures" as well as research "successes."	More likely to focus on "success stories.
	More likely to double count research benefits.
Can capture the effects of minor improvements in practices.	More likely to focus on specific technologies developed by research.
Can capture the effects of gradual changes in farm organization, scale, and specialization that take place as the result of technological change.	
Can account for the concurrent effects of other investments on agricultural productivity, including education, infrastructure, and farm programs.	Focus on impacts of agricultural research only.

Sources: Compiled by authors, based on Alston et al. (1995); Huffman and Evenson (2006); Fuglie and Heisey (2007).

A straightforward application of an economic surplus approach is a useful framework to analyze and assess the effects of research on agricultural markets. But a powerful feature of economic surplus methods is the ability to explore more complicated questions. For instance, if research on product quality enhancement was funded by a commodity group, an important comparison might be between research costs and producer surplus, rather than between research costs and total surplus gained. If research affects prices and quantities in one market, the effects can be linked to their impact in related markets. Economic surplus methods can also be extended to analyze more than one product, or to consider market distortions caused by agricultural and other economic policies or by externalities (Alston et al., 1995). Also, market supply and demand is usually composed of different groups of suppliers and buyers; economic surplus analysis can show the different effects of technology on producers or consumers with different characteristics (table 2).[15]

In short, economic surplus methods are useful in analyzing particular rather than aggregate effects of research. Because they can be applied to less aggregated data, they can be used more effectively to analyze particular technologies or research programs, thus allowing greater emphasis on cause-effect relationships. An early example of such a study focused on the introduction and spread of hybrid corn (Griliches, 1957; 1958). In addition to finding the total value of research benefits, surplus methods are useful for determining how research benefits are distributed among buyers, sellers, and market intermediaries. These models can disaggregate benefits both "vertically" (over producers, consumers, and market intermediaries) and "horizontally" (over agro-ecological zones, regions, or countries; over types of producers; or over income classes of consumers). If the analyst is willing to define and forecast research benefits carefully, economic surplus methods can also be used for *ex ante* as well as *ex post* research evaluation.

[15]In some cases, if reliable estimates for supply and demand parameters (e.g. elasticities) are available for the market or markets being analyzed, economic surplus methods could require somewhat less data collection than econometric methods.

Table 2

Advantages of economic surplus versus econometric methods

Economic surplus methods	Econometric methods
Focus on particular technologies or research programs, with greater emphasis on cause-effect relationships.	Confined to more aggregated data.
Can allow vertical disaggregation of benefits over producers, consumers, and market intermediaries.	More likely to measure aggregate economic impact.
Can allow horizontal disaggregation of benefits over countries, regions, and agroecological zones; types of producers; and income classes of consumers.	More likely to measure aggregate economic impact.
Can allow measurement of effects in markets related to the principal market of interest.	Primarily measures effects for sector or commodity being analyzed.
Can be used for *ex ante* analysis as well as *ex post* analysis.	Used primarily for *ex post* analysis.

Sources: Compiled by authors, based on Alston et al. (1995); Huffman and Evenson (2006); Fuglie and Heisey (2007).

The relative strengths and weaknesses of econometric and economic surplus methods sometimes appear to be complementary. It is not always possible to see how these strengths and weaknesses play out in practice, but in one area, some information is available. Our evaluation of the two methods suggests that research benefits and returns are more likely to be overstated by economic surplus methods. Evenson's (2001) survey of the literature on returns to agricultural research, however, shows that the median internal rate of return for studies using econometric methods was actually *slightly higher* than the median internal rate of return for economic surplus studies; furthermore, he notes that econometric studies tend to have more outliers with extremely high rates of return. On the other hand, Alston et al. (2000) apply a regression meta-analysis in a survey of much the same literature, and find that econometric studies do tend to produce lower internal rates of return. However, the regression coefficient comparing the two types of studies was statistically insignificant.

A whole host of factors may contribute to the reliability of either approach in practice, and potential economic extensions exist for each approach. In addition, methods such as public value mapping (PVM) exist that attempt to broaden program assessment by focusing on social rather than economic benefits. These issues are explored in the Appendix ("Challenges of and Potential Extensions to Standard Economic Methods").

Intermediate Economic Analysis

Economic analysis is data and time intensive. The greater the "public goods" aspects of the research in question, the more difficult it may be to measure benefits in observable markets. Standard economic analysis has primarily been applied at a relatively high degree of aggregation, particularly in the case of econometric analysis. Even economic surplus methods have often been applied to agriculture in the aggregate or to entire commodities to make use of available aggregate-level data. On the other hand, cost-benefit analysis, when it has been applied to Federal research evaluation, has often been directed to a single research project or cluster of projects. How can economic analysis be modified or used to inform research evaluation at an intermediate level? The following four options are aimed at the evaluation of research *programs*—which lie somewhere between research projects and wide-ranging research subjects such as commodity research or natural resources research. Table 3 gives an example of different degrees of aggregation in research based on USDA's classification of water-related research.

There are four general approaches to intermediate-level economic analysis (table 4) that can provide information on the economic benefits of public R&D when data are not readily available, or when more targeted analysis is desired. These approaches are not, in all cases, full-fledged economic analyses, but they may be very useful depending on the objectives of the evaluation.

First, one of the major justifications for public sector research, from an economic perspective, is market failure because of the "public goods" character of knowledge generation. As a result, a simple but powerful approach is to *identify areas where governments should be involved in R&D by evaluating the public goods aspects of the research*. There are important advan-

Table 3

Examples of different levels of research investment

The level of aggregation at which research is assessed depends on the reason for the analysis. The project level provides detailed information about specific research outcomes, but a broader level is usually necessary to capture spillover effects into other areas of scientific or economic activity. As a result, information about research is collected at several different levels. The following tables show how two research categorization systems in use by USDA might represent different levels of water research.

The narrowest level in the top table corresponds very roughly to the broadest level in the bottom table. However, the $65.6 million in CRIS-reported expenditures for ARS water-related research probably includes several water-related projects from outside NP-211, and probably excludes some NP-211 projects that simultaneously pursue other research objectives.

Current Research Information System (CRIS) database			
Level	**CRIS unit**	**Examples**	**Reported expenditures (FY 2006)**
Broadest	Research program group	Natural resources and environment research	$964.9 million
Intermediate	Research program	Water resources/ watersheds/wetlands/ riparian systems	$207.6 million
Narrowest	Research program by type of institution	Water-related research at ARS	$65.6 million

From ARS National Program summary		
Level	**National Program summary description**	**Examples**
Broadest	National Program	NP-211, Water Availability and Watershed Management
Intermediate	Program subcomponents	Effectiveness of Conservation Practices Irrigation Water Management Drainage Water Management Systems Integrated Erosion and Sedimentation Technologies Watershed Management, Water Availability, and Ecosystem Restoration Water Quality Protection Systems
Narrowest	Research project	Project 3625-13000-009-00: Water Quality Improvement from Management Practices in Agricultural Watersheds (South Fork Iowa River and Walnut Creek, Story County watersheds)

The research project listed probably falls under more than one "problem area" within NP 211.

Sources: Current Research Information System (CRIS); USDA-Agricultural Research Service at http://www.ars.usda.gov/research/programs.htm.

Table 4

Advantages and drawbacks for research program evaluation methods

Method	Advantages	Drawbacks
Identifying public goods in public research	+ Focuses attention on impact + Can lead to development of appropriate measurement	– Lacks quantitative metrics – Does not allow easy comparisons across projects or resources
Narrative/anecdotal assessments	+ Communicates to non-technical audiences + Focused anecdotes are quicker to complete	– Selection of examples might not represent the entire program
Economic surplus/cost-benefit analysis for related projects	+ Provides detailed information and analysis + Able to capture distinctive aspects of research	– High time and resource costs of implementation – Research impacts more difficult to measure than costs (still forthcoming, hard to value, etc.)
Economic surplus/cost-benefit analysis for an entire research program	+ Most detailed and comprehensive + Best basis for making allocation decisions	– Most expensive method – Research impacts might be difficult to measure

Source: Compiled by authors.

tages to this approach. Careful analysis can identify R&D that can be left to the private sector, freeing up resources for better public use. Furthermore, areas identified should at least have the possibility of providing positive economic returns to public research. On the other hand, merely identifying public goods aspects does not estimate quantitative economic benefits to the research, nor does it provide a basis for linking research benefits to research costs, or for choosing among alternative research investments. Like all economic approaches, it is based primarily on the market failure paradigm and does not address questions of broader social impact. In addition, analysts who take this approach must be able to define the economic concept of "public goods" within the context of R&D, discriminate among alternative investments, and be able to communicate carefully with research policy-makers and scientists concerning the meaning of public goods. This method can be used in conjunction with other methods for assessing projects, components of a program, or entire programs.

A second approach to intermediate level analysis that might have some economic component is the *narrative or anecdotal approach*. In this approach, the analyst would present carefully gleaned examples of successful research projects or program components within the program being evaluated. This approach is quicker and less data intensive than full-fledged economic analysis. It is also useful for meeting formal requirements of Federal research evaluation, such as the Program Assessment Rating Tool (PART) or Government Performance and Results Act (GPRA). Politically, such an approach may at times prove helpful in justifying research budgets. At times, examples of successful research can provide some quantitative estimates of economic benefits. On the other hand, quantitative estimates of benefits in this approach are usually fairly rough. Again, the potential for linking research benefits to research costs is limited, and the approach

gives little basis for allocating resources across different research activities. Any approach of this type is also subject to the charge that analysts "cherry pick" successful projects and do not provide a broader research evaluation that includes management options to make unsuccessful research more productive.

A third way of applying standard economic methods to research program evaluation would be to *apply economic surplus methods to individual projects or groups of projects within the given research program.* A simplified version of this approach would be the use of price and quantity data for the markets affected by the results of the projects studied, rather than full-fledged economic surplus analysis. In fact, the more disaggregated the analysis, the more likely the analyst will resort to this kind of shortcut. This kind of approach could, however, permit more careful measurement of economic benefits, and unlike the simpler approaches outlined above, provide a means of linking benefits to costs. *Ex ante* economic surplus analysis could be helpful in research planning exercises; *ex post* analysis could be useful in demonstrating the effectiveness of a given research program. Particular care must be taken with economic surplus analysis of projects or clusters of projects within a research program, as the generation of quantitative estimates of benefits could give a spurious appearance of precision. As with the narrative or anecdotal approach, economic surplus or cost-benefit analysis is subject to the criticism that it focuses on "successful" research projects.

A final method of economic analysis would be the *application of economic surplus analysis to the entire research program.* Such an approach would share many of the same strengths and weaknesses as the application of economic surplus analysis in a more narrowly defined context. In addition, the application of economic surplus techniques to an entire research program could permit particularly careful measurement of economic benefits, reduce the likelihood of double counting, and increase the likelihood of capturing smaller scale benefits, spillovers among related projects, or benefits from individual research projects that might not appear to be unqualified successes when standing alone. On the other hand, whole-program analysis would be even more data and time intensive than analysis focusing on particular projects or project clusters, and it would require particularly close attention to the multiple sources of research results that influence measured benefits.

Assessing Federal and ARS Research

Assessment of public sector investments is an important basic function of government. Ultimately, the conduct of public research must be accountable to the citizens who finance it. At the Federal level, R&D spending is authorized by Congressional legislation and administered by the Executive Branch. Determining the soundness of public investments and ranking competing projects to select the best alternatives is critical for accountability of Federal agencies conducting R&D.

As we have seen, R&D generally presents difficult assessment challenges. Research seeks to discover novel results, which are intrinsically difficult to compare with prior examples. The time horizon of a typical R&D project can require years, and the adoption and impact of R&D results might not be known for decades. (A hypothetical model of the flow of research costs and benefits over time is shown in figure 2.) Therefore, early assessments of R&D will generally involve forecasting error, or they might focus on progress toward intermediate goals. This contrasts with many other Federal activities that can be assessed on an annual or even more frequent basis.

Despite these difficulties, the need for assessment of R&D has generated significant effort at the Federal level, although much of this assessment is reported outside of the peer-reviewed literature. Recent governmentwide planning and assessment activities include the Government Performance and Results Act (GPRA) of 1993 and the Program Assessment Rating Tool (PART) of 2002. In particular, PART provides special guidelines for evaluating research agencies, noting that "assessing the outcomes of basic research in particular is never easy" (OMB, 2007, p. 72).

The way in which the Federal Government supports research— direct performance versus research sponsorship—may influence the types of research outputs and affect the relative importance given to different methods of research assessment.[16] Over half of total Federal research expenditures in any given year are defense related. A little over 20 percent of Federal expenditures in recent years were intramural--i.e., <u>performed</u> by Federal agencies.[17]

[16]See chapter 2 for a review of different public approaches to funding research.

[17]This percentage is a decrease from the figure in the mid-1990s, when Federal intramural research comprised roughly a third of the Federal total (Cohen and Noll, 1996).

Figure 2

Flows of resources, research benefits, and research costs

Source: Alston et al., 1995.

Recent research obligations spent intramurally for 2005-07 by Federal agencies, including USDA's Agricultural Research Service, on non-defense-related research are shown in table 5. This table does not include non-defense Federal research obligations made through grants.[18] It also excludes federally funded research and development centers (FFRDC) that are administered by universities and corporations. Even when grants administered by NIH are excluded, NIH is the largest non-defense Federal intramural program.[19]

Literature on Federal Assessment

Although limited in coverage, a brief but focused literature describes past efforts toward assessing Federal R&D (table 6). Surveys by Kostoff (1995) and the National Academy of Sciences (1999) suggest five rough (and overlapping) approaches to this assessment: peer review, narrative assessments, quantitative analysis (including bibliometric and cost-benefit methods), benchmarking, and retrospective analysis.

Peer review decisions are reached by a panel of scientists with expertise in the research programs to be assessed. For example, the NSF makes competitive, merit-based reviews central to their selection of externally funded

[18]Two of the principal Federal institutions that solicit research proposals and administer research grants are the National Institutes of Health (NIH) and National Science Foundation (NSF).

[19]NASA, which has a large intramural research budget, like NIH grants much of its budget to research partners.

Table 5
Federal non-defense intramural research obligations by agency

Agency	Research obligations FY 2005-2007 average, $1,000
NIH	5,203.3
NASA	1,212.6
ARS	990.1*
USGS	481.2
EPA	466.3
VA	408.4
NOAA	339.5
NIST	314.7
FS	286.6
Energy Efficiency/Renewable Energy	204.8
CDC	191.2
Smithsonian	121.3
Fossil Energy	116.3

NIH: National Institutes of Health, U.S. Department of Health and Human Services; NASA: National Aeronautics and Space Administration; ARS: Agricultural Research Service, U.S. Department of Agriculture; USGS: United States Geological Survey, Department of the Interior; EPA: Environmental Protection Agency; VA: Department of Veterans Affairs; NOAA: National Oceanic and Atmospheric Administration, Department of Commerce; NIST: National Institute of Standards and Technology, Department of Commerce; FS: Forest Service, U.S. Department of Agriculture; Energy Efficiency/Renewable Energy: U.S. Department of Energy; CDC: Centers for Disease Control and Prevention, U.S. Department of Health and Human Services; Smithsonian Institution; Fossil Energy: U.S. Department of Energy.

*As is often the case with other Federal intramural research agencies, some funds in the ARS budget are used for cooperative agreements and contracts with research partners, particularly in universities, whose work complements ARS' research mission. This is the main reason for the difference between this estimate and the figure for the ARS budget in the text.

Source: NSF, 2008. Obligations rather than outlays are reported in this table because NSF lists agency obligations, but not outlays, by research performer.

Table 6
Literature on assessments of benefits from Federal research

Authors	Year	Subject	Method 1	Method 2	Method 3	Method 4
Survey articles						
Nielson and Brazzel	1980	Ag science	Survey of methods			
OTA	1986	Federal Government	Survey of methods			
Logsdon and Rubin	1988	10 Federal agencies	Survey of methods			
Kostoff	1995	Federal Government	Survey of methods			
Hicks et al.	2002	Federal Government	Survey of methods			
Assessments						
Salasin et al.	1980					
Martin and Irvine	1983	Radio astronomy	"Converging partial indicators" method	Bibliometric	Benchmarking	
Link	1995		Bibliometric			
Brown	1996	Army research lab	Bibliometric			
Cozzens	1995	GPRA requirements				
NAS/COSEPUP	1999	GPRA requirements	Bibliometric	Peer review	Peer evaluation of programs	
Cahill	2000	DoD/SBIR*	Bibliometric	Peer assessment		
Fitzsimmons	2001	Automotive industry	Bibliometric			
Audretsch et al.	2002	DoD/SBIR	Case study/ interviews	Tobit model (of commercialization)		
Tassey	2003	DoC/NIST**	Cost-benefit	Case studies		
Link and Scott	2004	ATP/Optical Fiber Communications	Cost-benefit	Case study	Prospective & retrospective "social benefits" estimation	Peer assessment
Peretz et al.	2005	Car projects at DOE	Case study	Bibliometric	Peer review	Benchmarking
Macauley	2006	NASA/Spectro Radiometer	Retrospective			

SBIR = Small Business Innovation Research; GPRA = Government Performance and Results Act ;DoD = U.S. Department of Defense; DoC = U.S. Department of Commerce; NIST = National Institute of Standards and Technology
Source: Compiled by authors.

projects (OMB, 2007).[20] Survey pieces by the Office of Technology Assessment (1986), Kostoff (1995), NAS (1999), and Hicks et al. (2002) find that Federal agencies prefer using peer review to assess their research projects.

Narrative assessments use case studies and anecdotal assessments to provide a description of a research program. Although they are well adapted to the unique characteristics of R&D programs, they are difficult to use when comparing projects or programs. The NSF makes extensive use of narrative assessments in their PART reviews, and ARS requested and received an OMB waiver to use a narrative approach in its GPRA planning.[21]

Quantitative analysis has several perceived advantages. GPRA, PART, and other policies that stipulate quantitative measures are based on the assumption that such measures allow comparisons across agencies. Properly designed quantitative measures can be useful decisionmaking tools, but poorly constructed measurements promote facile understanding of R&D activities. OTA (1986) and Hicks et al. (2002) note that quantitative assessment is rare in the review of Federal research programs.

Cost-benefit approaches, rates of research return, and related quantitative approaches appear especially desirable in that they reduce projects to a common measure, often a present-dollar value of program benefits and program costs or a ratio of costs and benefits. These measures, reviewed by Tassey (2003), are closely related to the economic surplus approach. Tassey indicated that these approaches should be conducted externally to minimize reporting bias, although the Australian examples cited by Mullen (2004) demonstrated the advantages of including internal analysts who understand the markets being addressed by particular research programs and the scientific approaches undertaken.[22] Tassey (2003) and Mullen (2004) found that detailed quantitative program assessments are expensive.

Bibliometric techniques take advantage of the fact that new knowledge created by R&D programs is generally published in scientific journals and (less often) as patents. Although the knowledge generated by R&D is not confined to these outlets, they represent a way to quantify and compare R&D results. Counting research publications and analyzing their citation patterns is known as bibliometric analysis (see Hicks et al. for an extensive discussion of bibliometric use in the Federal Government). Problems with bibliometric analysis include a lack of objective standards other than comparisons among similar projects in similar areas of science. Martin and Irvine (1983) found that, if the analysis is carefully constructed, basic research can be assessed with bibliometric methods, in combination with peer reviews. They note that past performance is one of the best indicators of future performance, and therefore is likely to be useful for policymakers. NAS (1999) indicated that bibliometric methods still require expert evaluation of the content, quality, and relevance of citations and patents to be effective. Bibliometric methods are not well suited to assessing current research projects, since citations lag publication and publication lags discovery and achievement.

Benchmarking is the measurement of progress against internally determined performance measures. Annual reports produced to meet GPRA requirements make extensive use of benchmarks. This is one way to overcome the fact that

[20]The Committee of Visitors is further used to examine NSF management.

[21]Special Report on the Government Performance and Results Act of 1993, Senate Appropriations Committee, 106-347.

[22]Alston et al. (2000) compare rate-of-return studies done by analysts within the agricultural research institution being evaluated and those done by external analysts. They find that "self-evaluators" actually report significantly *lower* estimated rates of return, perhaps because they are better informed, and perhaps because they bias their estimates downward to be considered plausible. Alston et al. also find that government-based evaluators tend to report *higher* rates of return than analysts employed by universities, the private sector, or unspecified employers.

science results are only evident in the long term, whereas budgeting, planning, and accountability exercises occur more frequently. These measures, based on internal processes rather than external impacts, can be acceptable for Federal R&D programs (OMB, 2007).[23]

Surveys, Guides, and Additional Examples of the Federal Evaluation Literature

In addition to the surveys mentioned above, Logsdon and Rubin (1988) provided an early survey of evaluation methods focused on 10 Federal agencies. The authors detailed the assessment techniques and culture of each agency/department. While the findings are dated, it is one of the few reviews that considered USDA's assessment methods. Logsdon and Rubin described the R&D evaluation process at ARS as highly structured and formal. The Current Research Information System (CRIS) provided transparency through detailed funding summaries that describe the subjects and goals of ARS research. As of 1988, ARS was using patent, publication and citation/bibliometric analysis, as well as peer review. CSRS (was CSREES and now NIFA) used cost-benefit analysis, though most of this research was carried out at the State, not Federal, level. OTA (1986) provided a case study of agriculture focused on the many studies finding high rates of return to agricultural research (see Fuglie et al., 1996; and Fuglie and Heisey, 2007, for a summary of studies). OTA did not perform its own review of USDA. The study did note the unique structure of the agricultural sector, specifically a high degree of cooperation between the public and private sectors, and the strong history of support for public research. An even earlier study by Nielson and Brazzel (1980) described the base of R&D evaluation in agriculture as narrow.

A 1999 study (NAS, 1999) by the National Academy of Science's Committee of Science, Engineering and Public Policy (COSEPUP) lists the pros and cons of different assessment methodologies within the context of the Government Performance and Results Act (GPRA). The study explored bibliometric methods, economic-impact studies, and expert reviews; the latter were deemed the most effective means of evaluation. "Expert review" included peer review, relevance review (that is, review by potential users of the research), and benchmarking (evaluation of the program relative to international standing).[24] The COSEPUP report stated that quantitative measures can be used to confirm expert findings. However, measures such as economic impact studies are inappropriate for basic research, unless used to assess the practical benefits of basic research retrospectively.

Tassey (2003) reviewed analysis of more than 30 programs using retrospective microeconomic impact studies, mostly variants on the benefit-cost approach.[25] Although focused on one agency, the National Institute of Standards and Technology (NIST), this study provided a detailed discussion on the selection of metrics, analysis level and timing, and impact measures. The technologies assessed by NIST were in such areas as communications, energy, chemicals, and semi-conductors. In addition to noting the cost of these studies, Tassey argued that most Federal agencies lack the internal capability to assess research impacts appropriately.

Several other studies have looked at particular Federal research programs or program components. Cahill (2000) combined bibliometric and peer assess-

[23]Recall that Bozeman (2003) advocated the use of benchmarking as a tool for public value mapping, in evaluating research contributions toward meeting broad social goals.

[24]This "external" benchmarking differs from benchmarking against internally defined targets.

[25]A useful review of the application of benefit-cost analysis to public sector agricultural research in Australia is provided by Mullen (2004).

ment techniques to review the Department of Defense's Small Business Innovation Research (SBIR) grants, primarily focusing on commercialization. Audretsch et al. (2002) used a case study approach, including a tobit model, to evaluate commercialization of technologies from the same SBIR program. Link and Scott (2004) also applied case study methodology to the NIST Advanced Technology Program (ATP). They used the "Griliches/ Mansfield" method for estimating social rates of return to research arising from an improved standard reference material used in optical fiber networks.

Ruegg and Feller (2003) assessed the overall performance of the ATP using 13 studies.[26] Prospective case studies indicated that the benefits of the program far exceed their costs. An analysis of 45 ATP evaluation studies found that modeling underlying program theory was the most frequently used evaluation method, followed by case studies and surveys. Program theory is a concept from public administration. Program theory models outline logical relationships between program activities and desired outcomes (Bickman, 1987). In the case of the ATP, these models often focused on the determinants of collaborative research success. Use of program theory analysis, case studies, and econometric/statistical methods has increased over time.

Finally, Peretz et al. (2005) reviewed six projects at the Department of Energy. Within a case-study approach, the authors used bibliometric, peer review, and benchmarking techniques. Among their findings was that firms involved in the DOE projects would not have engaged in that R&D without DOE funds. On a broader level, the authors observed that the Federal Government serves as an organizer of R&D efforts, which enhanced synergies among the players.

Ruegg and Jordan (2007) have written a quick reference guide to evaluation for managers of R&D programs.[27] They defined 14 evaluation methods, including their uses and limitations, and provided at least 1 example of each. Ruegg and Jordan stated that evaluation studies should take into account the level of effort to be employed; study design (descriptive, normative, or impact); the level of analysis (project, program, portfolio, system, or organization); whether the evaluation is to be retrospective or prospective; and the intended audience. They also noted that different evaluation methods might be used depending on the phase of research performance being analyzed. These phases are:

1. Designing/revising, planning, selecting and budgeting;
2. R&D progress, review of process mechanisms, outputs;
3. Output dissemination and achievement of interim outcomes; and
4. Commercialization, market acceptance, outcomes, impacts.

The manual gave a roadmap (fig. 3), based on an "R&D logic model" for DOE.[28]

Planning, Directing, and Monitoring the ARS Research Portfolio

Each Federal research institution has developed its own methods of research evaluation, influenced by government directives such as GPRA, PART, and their predecessors; agency mission; available evaluation techniques

[26] Ruegg and Feller (2003) conducted a detailed analysis of 45 evaluation studies commissioned by the ATP from which they developed a "toolkit" to guide evaluators. Drawing on diverse government programs for illustrations, the Ruegg-Jordan (2007) report is a quick-reference guide to evaluation for managers of R&D programs who are not evaluators.

[27] Ruegg and Jordan (2007) illustrate the application of each method from a variety of programs, among them DOE programs..

[28] A logic model is a representation of inputs and processes that influence an organizational goal or other desired outcome. Ruegg and Feller (2003) provide a generic evaluation logic model.

as described in the preceding sections; and precedent. The Agricultural Research Service (ARS), USDA's principal inhouse research agency, is no exception.

ARS is charged with conducting publicly funded research for the benefit of the Nation. Research conducted by ARS is supported by an annual appropriation of over $1.1 billion (FY 2006/7).[29] ARS scientists participate in agricultural research and development that addresses a broad range of production and management programs, high- and low-technology techniques for livestock and crop production, natural resources and the environment, high-quality and nutritious food, and alternative energy production. Stakeholders of ARS research include agricultural producers, food processing industries, natural resource managers, universities, and other Federal agencies and non-profit research institutes. Considered more broadly, beneficiaries of ARS research include U.S. consumers, other Federal agencies, and international markets.

Research at ARS is divided among four broad program areas, each headed by a Deputy Administrator. Within the program areas, more than 1,000 active research projects are grouped into National Programs (NPs), each headed by a National Program Leader(s) (table 7). Research projects take place at over 100 ARS laboratories and locations around the United States, Puerto Rico, and the U.S. Virgin Islands, and four laboratories overseas (fig. 4). Laboratories are led by research leaders who oversee a number of related research projects and who are responsible for ensuring the quality and perfor-

[29]A little under 10 percent of the total ARS budget is passed through to research partners in other institutions. This accounts for most of the difference between this estimate and the estimate in table 5.

Figure 3
Basic logic of R&D programs and evaluation questions
Performance assessment questions span the performance spectrum.
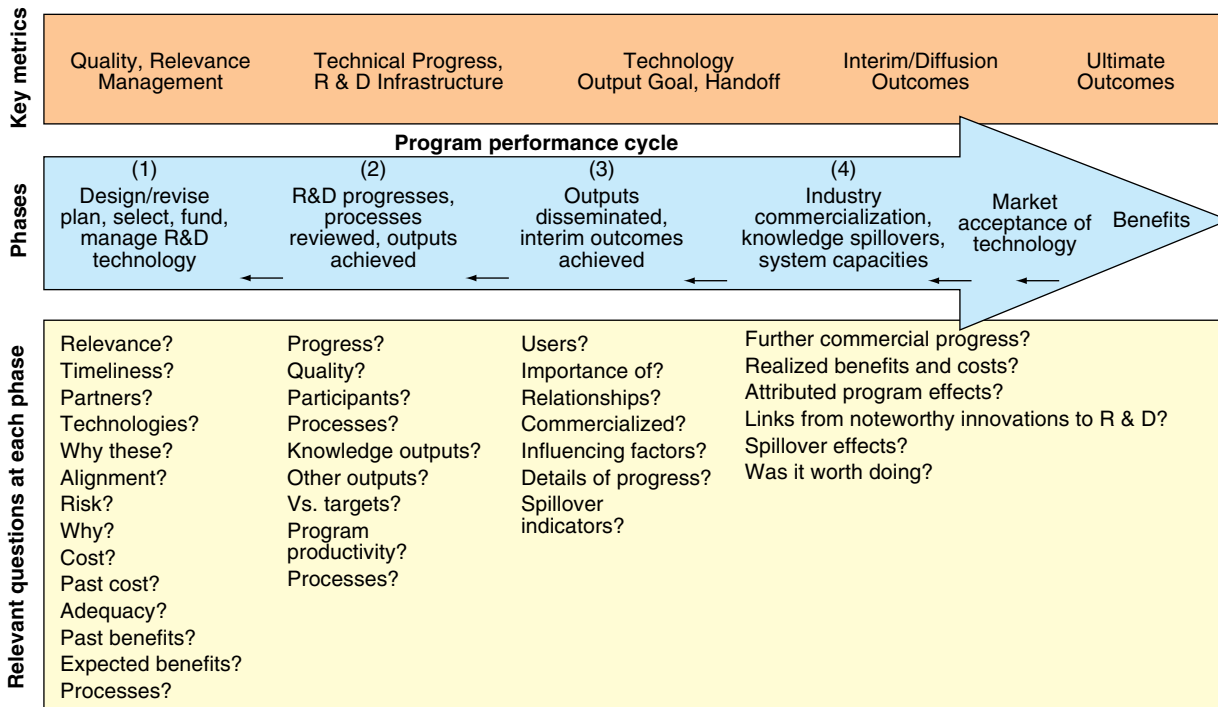
Source: Gretchen Jordan, Sandia National Laboratories, as reported by Ruegg and Jordan (2007).

mance of each project. In turn, individual projects are headed by lead scientists or research teams.

The organizational structure of ARS lends itself to assessment at different scales, which allows the use of different approaches. Assessment at the project level allows detailed observation adapted to the scientific effort and project impacts. Assessment at the laboratory or NP level permits analysis of impacts at the problem, commodity, subject, or regional level, and allows research managers to coordinate projects to minimize overlap, identify program gaps, and serve as liaison to related research at other agencies or

Table 7
ARS National Programs

Nutrition, Food Safety and Quality	Animal Production and Protection
• Human Nutrition • Food Safety (animal and plant products) • Quality and Utilization of Agricultural Products	• Food Animal Production • Animal Health Veterinary, Medical, and Urban Entomology • Aquaculture
Natural Resources and Sustainable Agricultural Systems	**Crop Production and Protection**
• Water Resource Management • Soil Resource Management • Air Quality • Global Change • Rangeland, Pasture, and Forages • Manure and Byproduct Utilization • Integrated Agricultural Systems • Bioenergy and Energy Alternatives	• Plant Genetic Resources, Genomics, and Genetic Improvement • Plant Biological and Molecular Processes • Plant Diseases • Crop Protection and Quarantine • Crop Production • Methyl Bromide Alternatives

Source: USDA's Agricultural Research Service..

Figure 4
Agricultural Research Service areas and locations
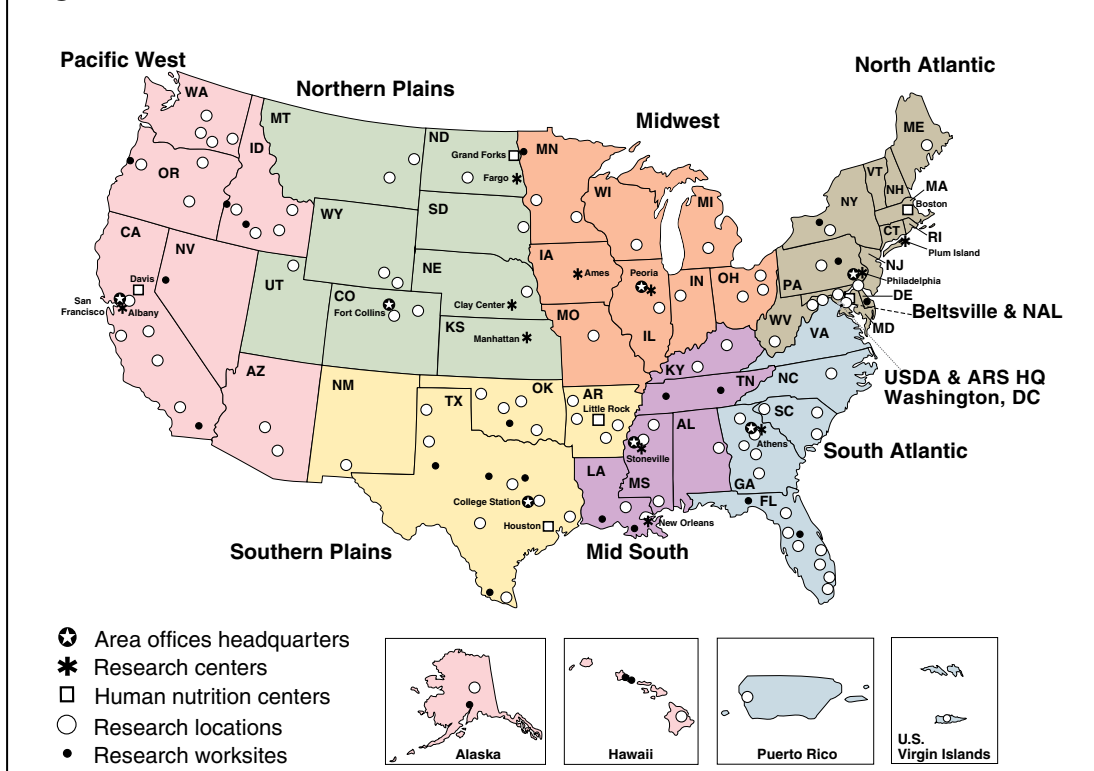
institutions. Full-scale assessment at the program area or Agency level is suited to aggregate measures of research impact, especially over longer periods of time.

The ARS Planning Cycle and Tools of Research Assessment. Many of the primary assessment tools currently used at ARS are tightly bound to a 5-year planning cycle (fig. 5). Each NP has a 5-year cycle that allows initial setting of expectations and identification of activities to be tracked. At the outset of the planning cycle, a cross-section of ARS stakeholders and customers identify and prioritize research needs for each NP. Along with input from stakeholders and consideration of administration goals and priorities, NP leaders set research priorities and goals in 5-year Action Plans that guide the overall research for that NP. NP leaders allocate funds and assign specific research objectives to individual research projects, which are the fundamental units of ARS research. Lead scientists or research teams in charge of projects develop Project Plans, which serve as working agreements between NP staff, middle managers such as the lead scientists, and the science teams.

After review by research managers, including leaders of national-level laboratories and area offices, project plans are submitted by the lead scientists for peer review managed by the ARS Office of Scientific Quality Review (OSQR). Peer review by OSQR, unlike peer review at Federal grantmaking agencies such as NSF, does not affect project selection or funding levels. Rather, OSQR peer review focuses on the scientific methods employed in ARS research, to ensure that they are acceptable and transparent. Peer review panels are made up almost entirely of non-ARS scientific professionals headed by an external panel chairperson with expert knowledge pertinent to the planned research projects. Each panel evaluates proposed research methodology, probability of success, and scientific merit. When major revi-

Figure 5
ARS 5-year research program cycle
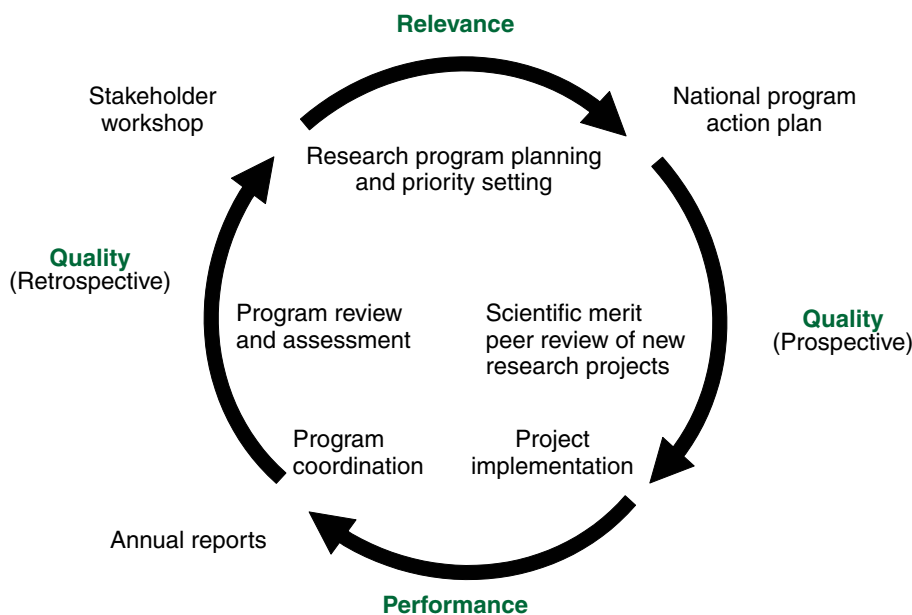
Source: ARS.

sions are required, project plans are sent back to lead scientists with suggestions for improvements needed to win approval. If revised project plans are rejected, management may reallocate human and fiscal resources and/or redirect the focus of the research so that it will result in an acceptable plan.

Assessment activities begin as the research projects get underway. Lead scientists submit *Annual Progress Reports* for their projects, which help NP leaders to compile annual reports assessing the yearly progress that each NP has made in meeting the goals set in their 5-year Action Plans.[30] Near the end of the 5-year program planning and evaluation cycle, NPs undergo another level of review. NP teams provide external *Retrospective Review* peer panels with accomplishment summaries for each NP, using the aggregate information contained in the annual reports and projects aligned with that NP. After evaluating the aggregated accomplishment summaries, the *Retrospective Review* panel assesses the value of actual research accomplishments compared with project goals in NP Action Plans. The panel also makes recommendations for future research priorities, which are then used by NP teams as they launch new 5-year planning and assessment cycles. Retrospective reviews began in 2005, and procedures and guidelines continue to be refined.

Long-Term Agency Accountability. ARS uses its *5-Year Strategic Plan* as a key benchmark to assess Agency performance and progress. The plans outline several broad goals that clearly define what the Agency hopes to accomplish in the next 5-year cycle. Using a strategic plan in this manner was facilitated in part by the passage of the GPRA. After an interim period of pilot studies, ARS released strategic plans that satisfied GPRA in 1997 (covering fiscal years 1998-2002), 2002 (covering fiscal years 2003-2007), and 2007 (covering fiscal years 2006-2011).

In response to directives developed to promote governmentwide accountability, ARS compiles *Annual Performance Plans and Annual Performance Reports* that review progress made in meeting goals established in the ARS 5-Year Strategic Plan. The first, Annual Performance Plans, identifies specific performance indicators and anticipated outcomes that will, if accomplished, indicate progress toward achieving set objectives and goals in the year to come. The second, a single Annual Performance Report, reviews the mission and progress of the Agency in meeting the performance measures and indicators of progress in the previous year's *Annual Performance Plan.*

The Office of Management and Budget compiles some of the GPRA reporting requirements, but ARS is also accountable for other external reviews by OMB. In response to the President's Management Agenda, instituted in 2001, ARS began assessing agency activity using OMB's *Program Assessment Rating Tool (PART)* to evaluate four program components: purpose and design, strategic planning, management, and results/accountability. PART analysis provides ARS with an ongoing yardstick for measuring the overall, long-term effectiveness of its research programs. The information gained from PART and its Research and Development Investment Criteria review helps ARS identify low-performing and/or low-priority research. It is used in shaping future budget requests, steering program management decisions, and adjusting the project's base funding.

[29]These reports are available at www. ars.usda.gov.

PART emphasizes the metrics of outputs/process control rather than the evaluation of research investments (NRC, 2008). The measures that ARS has used in fulfilling PART requirements have included peer-reviewed publications and technologies licensed. Individual research programs may have different quantitative targets. For example, water quality research (see chapter 5) may record the number of models completed and delivered or the acres of highly erodible land set aside in response to conservation policy. Measurement of highly erodible land uses the universal soil loss equation (USLE), a metric developed by ARS and USDA's Natural Resources Conservation Service (NRCS).

Research Evaluation Tools Used by ARS. ARS makes extensive use of at least three of the research evaluation tools outlined earlier. External *peer review* that addresses the scientific quality of the research is a major part of the prospective evaluation of research projects. External peer review is also the major component of retrospective reviews at the National Program (NP) level. *Benchmarking* of achievements against targets—both through the internal processes of progress reports and through agency compliance with GPRA and PART—helps to assess research progress as well as indicate potential redirection of research. Progress reports at the project level and performance reports at the agency level also make considerable use of *narrative assessments* in addition to reporting on quantitative process targets.

Both peer review panels and progress/performance reports often indicate broad economic goals of ARS research. However, measuring the economic benefits of scientific research is difficult for research managers and policymakers, especially in the short term. Several aspects of the research process reflect this difficulty:

- Outcomes/impacts of research are difficult to identify and measure in advance.

- The value of knowledge and information gained is not always immediately recognized.

- Results are not always predictable.

- Negative results—discovering the absence of an effect or treatment—is useful information that is difficult to quantify.

Because of these factors, ARS continues to rely on a narrative approach in describing accomplishments in *Annual Performance Reports*, rather than attempting to quantify research outcomes using numeric metrics alone.

In summary, ARS research assessments are similar to those of other Federal R&D in that they are often non-quantitative and use peer review extensively. The body of literature using quantitative methods to evaluate Federal research in general is small, and much of it is not actually economic in content, focusing instead on process metrics. Quantitative analysis of economic benefits from Federal research programs is found most often outside the life sciences, in fields like energy with near-term applications for industry with outcomes that are more readily available for economic analysis. Attempts to quantify basic research have been largely theoretical, or have involved highly specific case studies.

Case Studies of ARS Research

To examine specific approaches to better assess the economic benefits of research, a team of social scientists and research managers examined components of three distinct National Programs (NPs) at ARS. The resulting case studies do not attempt to estimate economic benefits of the research programs, but instead summarize ARS research programs and compare ARS evaluation activities with suggestions from the economics literature and practices used throughout the Federal Government. The case studies illustrate potential advantages and difficulties for assessing economic benefits, and which strategies are most likely to strengthen such assessment.

Case Study Selection and Methodology

Multiple case studies were chosen to represent a range of ARS research and assessment and to explore as many economic hypotheses as possible. Cases were selected based on three criteria: how the case study would help in understanding the roles of ARS research; what potential obstacles might interfere with the case study; and logistical concerns about conducting the case study.[31]

For understanding the **roles** of ARS research, we asked to what extent the science involved was fundamental in nature; how the research addressed USDA's mission, and if it served some regulatory function outside ARS or USDA; whether ARS played a major role in coordinating research across agencies, universities, or the private sector; to what extent the research addressed some externality or market failure; and to what extent the market structure of potential target industries might affect the use of research results. Regarding possible **obstacles** to the case studies, we examined whether the potential case study would attract the attention of agricultural interest groups, and whether there might be potential confidentiality issues for that particular case study. For **logistics**, we asked whether there are well-defined limits to the research program.

The research team met with each of the four ARS Deputy Administrators and developed a list of 12 potential case studies. With guidance from ARS Deputy Administrators, the team decided that ARS research on (1) bovine quantitative genetics and genomics, (2) water quality and watersheds, and (3) the Nutrient Data Laboratory would best characterize ARS research.[32]

Research programs generally comprise research activity over a longer term and at a larger scale than an individual research project. This definition of research "programs" is not identical to a National Program within the ARS management structure, but case studies at this scale allow the assessment of a portfolio of several integrated projects while still permitting the use of detailed assessment techniques. The scope of the selected case studies varied widely, allowing analysis of issues relating to program scale and research coordination. For instance, some ARS research on water quality may be conducted outside of the main ARS National Program (NP 211) in which water research resides. A portion of ARS research relating to food nutrient content resides in other laboratories within NP 107, Human Nutrition, but outside the Nutrient Data Laboratory.

[31] Another initial consideration, whether key participants could readily be identified for the case study, was eliminated after it became evident that this step would pose no difficulties.

[32] The first alternative case study was area-wide pest management programs.

A range of case studies also permitted the analysis of different rationales and motivations for ARS research on a topic. Initially, we expected that response to market failure would significantly characterize research about water quality/watershed resources and research by the Nutrient Data Laboratory. We hypothesized that an emphasis on basic research would guide ARS research on bovine quantitative genetics and genomics, and that ARS research in this area would also illustrate close research coordination with stakeholders. In addition, we supposed that water quality and watershed research would illustrate some of the ways that ARS research might support Federal regulatory functions. Having examples of different primary rationales for public research influenced our selection of cases.

Case studies were conducted through two principal and interacting modes. Initially, we assembled relevant literature from management documents (e.g., National Program Action Plans and Annual Progress Reports) and scientific publications within case study areas from both inside and outside ARS. Then, the major component for developing the case studies was to obtain further information about ARS research programs through indepth qualitative interviews with ARS research administrators, scientists, and stakeholders.

Investigators conducted taped interviews for this report, following an interview guide approach (Patton, 1990). A semi-structured interview guide was used, meaning that a set of core questions and the sequence were pre-determined. However, many interview questions were open-ended, and researchers were free to pursue unplanned topics as they arose. This methodology achieved the goal of addressing specific topics while allowing investigators to let interview subjects influence the direction of the interviews (Patton, 1990).

Interviews in this study drew on a "grounded theory" approach, in which the investigators probe emerging questions with interview subjects and test out tentative conclusions with further research. In this approach, investigators stop collecting data when the findings reach a point of saturation, whereby new data tend to confirm working conclusions rather than reveal new critical dimensions (Strauss and Corbin, 1998). The rationale for this method is the concept of the information-rich case. Subjects are not chosen at random; rather, the sample is selected purposefully to identify information-rich cases. In support of this approach, Harper (2001, p. 27) argues that "a small number of well-informed informants are, in fact, a better sample than much larger samples of minimally involved subjects."

To select our sample of interview subjects, we first identified key informants for each case study: someone in a position to know the program under consideration as well as the major actors and their roles within that program. The key informant was interviewed in order to learn more about the case, to develop an initial list of people to interview to complete the study objectives, and to prepare a list of publications relevant to the case. Additional interview guides were developed based on information collected from the key informant interviews and relevant documents provided or recommended by the key informant.

The investigators then interviewed the initial list of ARS personnel and research stakeholders using semi-structured interviews and the grounded

theory approach. At the end of each interview the subject was asked about other people to whom the investigators should speak. Interview guides were adjusted as more information was collected: some questions were added or deleted or reworded to provide clarity. Interviews continued until the investigators discerned that the research was nearing a saturation point. Generally, two investigators conducted each interview—one person asked questions and the other took notes and operated a tape recorder. Most interview subjects were interviewed individually, and all answers have been kept confidential.

Case Study: Bovine Quantitative Genetics and Genomics Research

Defining the Program

ARS and predecessor agencies within the U.S. Department of Agriculture have made important contributions to bovine quantitative genetics research since the early part of the 20th century. Today, research on bovine quantitative genetics and genomics falls under ARS' National Program 101, Food Animal Production.[33] This national program currently has an appropriated budget of approximately $46.5 million annually. NP 101 is divided into three components: (1) understanding, improving, and effectively using animal genetic resources; (2) enhancing animal adaptation, well-being and efficiency in diverse production systems; and (3) measuring and enhancing product quality. Component 1, which includes bovine quantitative genetics and genomic research, currently consists of 18 projects. Component 2 consists of 20 projects, and component 3 covers 5 projects.

Using the CRIS, expenditures on bovine quantitative genomics and genetics research can be found at the intersection of research on the use of animal genetic resources with research on beef or dairy commodities. We approximated investments in beef or dairy quantitative genetics and genomics by relying primarily on two key assumptions:[34]

- All research classified as "genetic improvement of animals" or "animal genome" under a Knowledge Area applies only to *live* farm animals, i.e., it does not apply to research on meat or dairy products.

- Allocation of research for a Knowledge Area within a species is proportional to the total amount of live animal research allocated to that Knowledge Area.

We converted all research expenditures to 2006 dollars by applying a research deflator maintained by ERS. Public sector expenditures on beef and dairy quantitative genetics and genomics research by State agricultural experiment stations (SAES) and other cooperating institutions are considerably higher than ARS expenditures (figs. 6 and 7).

Over the past decade, ARS has maintained a more even balance between genetic improvement and genomics research, while animal research at the SAES has shifted from genetic improvement to genomics. ARS and SAES expenditures on both beef and dairy genome research have risen 15 to 18 percent annually (in real terms) over the past decade. Over the same period (1998-2006), ARS investment in beef genetic improvement rose 9 percent per year, as did investment in dairy genetic improvement. In contrast, real

[33]The two categories are closely related. As currently defined, "genetic improvement" is comprised of a number of different research topics, including estimation of genetic parameters such as heritability and breeding values, selection studies, breed evaluation studies, and mating systems. "Genomics" includes gene mapping, gene identification, and genetic engineering. The former category, genetic improvement, also includes the incorporation of molecular and genomic information from the latter category, genomics, into applied genetic improvement programs (CRIS Classification Manual VII, http://cris.csrees.usda.gov/manualvii.pdf).

[34]More accurate estimates could be made by obtaining and analyzing CRIS data at the individual project level.

expenditures in beef genetic improvement by the SAES fell, with no significant trend in SAES dairy genetic improvement expenditures (figs. 6 and 7).

Another indicator of ARS investment in beef and dairy quantitative genetics and genomics research is the number of scientists working in this area. In the current Action Plan for NP 101, 42 of the 102 scientists do some research that falls under Component 1, with a focus on beef, dairy, or general bovine research. As of February 2008, ARS directly counts 22 full-time-equivalent (FTE) scientist years devoted to bovine quantitative genetics and genomics (R. Green, former National Program Leader, NP 101, personal communication).

Figure 6
Public sector expenditures on beef genetics and genomics research, 1993-2006

$ million (2006)
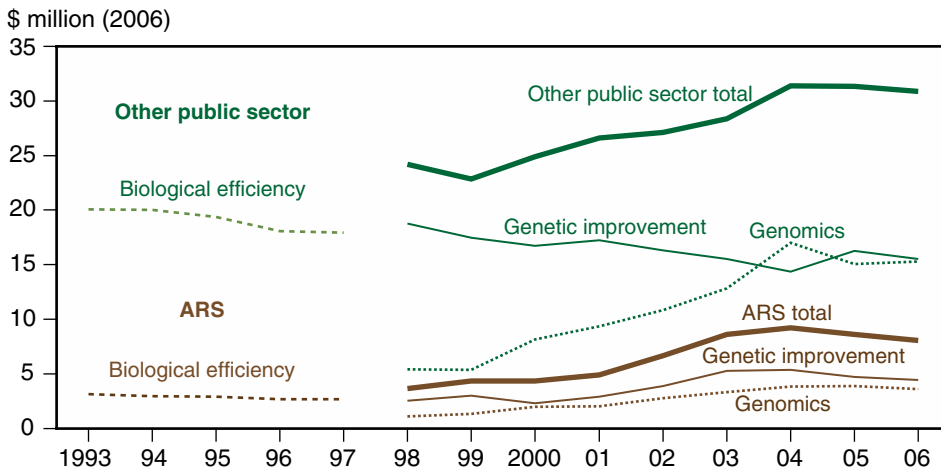
Source: Current Research Information System (CRIS); ERS.

Figure 7
Public sector expenditures on dairy genetics and genomics research, 1993-2006

$ million (2006)
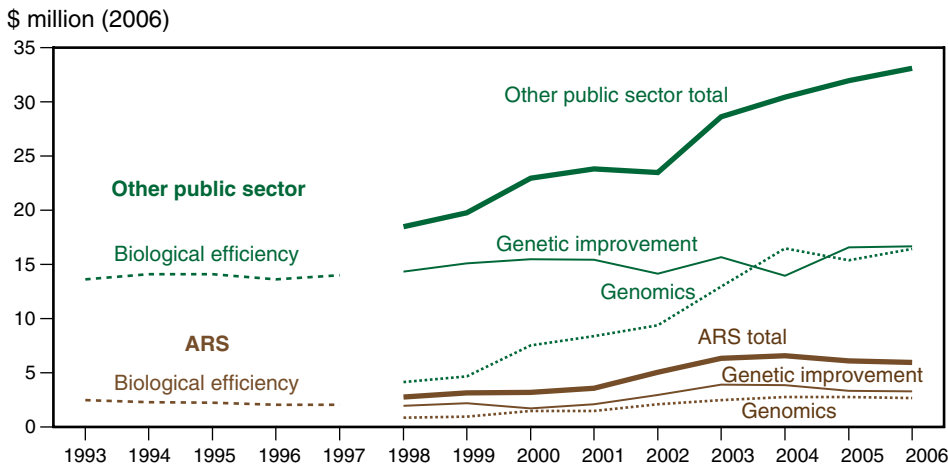
Source: Current Research Information System (CRIS); ERS.

Research Outputs

Across all public and private actors conducting bovine quantitative genetics or genomics research, research outputs can be categorized broadly as information or genetic material. The former might be codified in research publications or research protocols, or embodied in scientists' or animal breeders' tacit knowledge. Genetic material might consist of animals, or of preserved genetic resources such as frozen semen or embryos. Immediate outputs of ARS' research programs in bovine quantitative genetics and genomics include:

- Scientific publications, including:
  - Peer-reviewed journal publications,
  - ARS reports,
  - Germplasm reports,
  - Conference papers, and
  - Internet material, which may include some of the above.

- Information exchanged through personal contact, including:
  - Direct contact,
  - Interaction at workshops or conferences, and
  - Research collaboration.

- Directly transferred biological material.

ARS and USDA contributions to bovine quantitative genetics date back to the early years of the 20th century, when USDA inbreeding experiments in dairy cattle led to definitions of the genetic relationships among animals, inbreeding coefficients, and the concept of heritability of a trait (AIPL, 2007; Weller, 1994). Many of the current institutional arrangements for ARS beef and dairy genetic improvement research originated with the founding of the Meat Animal Research Center (MARC) in Clay Center, Nebraska, in 1966, and the Animal Improvement Programs Laboratory (AIPL) in Beltsville, Maryland, in 1972. From the late 1960s through the 1990s, Germplasm Evaluation and Utilization projects at MARC took the lead in defining the levels of genetic variation for beef cattle traits that contribute to lifecycle efficiency. In other words, the data and analysis from these MARC projects were crucial to understanding how to use both within-breed and between-breed genetic variation in industry breeding programs (R. Green, personal communication). MARC research on sire breed evaluation for topcross performance, across-breed expected progeny differences (EPDs, which estimate the genetic value of an animal as a parent not only within a breed but across breeds), and heterosis all contributed to much wider use of cross-breeding. Today, some 75-80 percent of beef cows are cross-bred; only the Angus herd remains primarily purebred (L. Cundiff, USDA/ARS/MARC-retired, personal communication, May 2007).

After the institution of the AIPL, new evaluations were added to evaluations developed prior to World War II for milk yield and milkfat composition. Protein and solid-not fat evaluations were undertaken in the 1970s; productive life, somatic cell score (related to evaluation for mastitis resistance), and net merit evaluations in the 1990s; and calving ease and daughter preg-

nancy rate criteria since 2000. The AIPL has played a significant role in the application of statistical animal models to dairy evaluation since the 1980s, allowing unbiased estimates of breeding values across herds and across environments.[35]

Over the last 20 years, one of the most significant advances in livestock genetic improvement research has been research at the molecular, or DNA, level. ARS has spearheaded many significant steps in bovine genomics research. In the 1990s, MARC researchers were instrumental in producing one of the first, low-marker density genetic linkage maps of the bovine genome, and then a second-generation linkage map (Kappes et al., 1997). Several important statistical designs for determining the relationships between marker loci and quantitative trait loci (QTL) in animals had been developed by 1990 (Weller et al., 1990). The widespread use of artificial insemination (AI) in dairy cattle results in natural populations that fulfill the requirements of these designs, and so ARS dairy researchers could develop QTL data for dairy cattle using Dairy Herd Improvement Association (DHIA) records (Van Tassell et al., 2001). However much of the work for the first bovine linkage maps was done by MARC, which developed large half-sib beef families specifically to create the linkage maps and identify QTL.

In the first decade of the 21st century, ARS has collaborated with a variety of State, national, and international institutions—including Baylor College of Medicine, CSREES (now NIFA)/land grant universities, the Commonwealth Scientific and Industrial Research Organization (CSIRO, Australia) and the National Institute of Health's Human Genome Research Institute—in planning and completing a sequence of the bovine genome (Bovine Genome Sequencing and Analysis Consortium, 2009; Bovine HapMap Consortium, 2009). ARS also led the effort to plan further USDA efforts in functional and structural genomics (Green et al., 2007).

Other Actors in Bovine Quantitative Genetics and Genomics Research

Public sector or nonprofit institutions doing research on bovine quantitative genetics or genomics also include universities with beef or dairy improvement programs. ARS and universities with such research programs stand at the interface between fundamental sciences such as genetics, statistics, or molecular biology and institutions such as breed associations, artificial insemination companies, seedstock companies, and DNA technology companies. These private sector institutions sell improved animals, artificial insemination and other reproductive services, DNA technology, and recordkeeping/classification services to beef and dairy producers. Thus, users of ARS research outputs are wide ranging, and the relationships among these institutions can be complex.

Users of ARS research outputs in bovine quantitative genetics and genomics stated they would find it difficult to obtain similar results from alternative sources, indicating the potential public good characteristics of ARS research. Depending on the type of user or the type of technology, alternative suppliers might be universities, private companies, or foreign research institutions. Reasons cited for the uniqueness of ARS research outputs included:

- The complexity and completeness of ARS data;

- ARS' ability to maintain large, pedigreed animal populations;

- ARS' ability to apply systems approaches to data analysis;

- ARS' ability to integrate new research with older findings;

- The high quality of DNA data from ARS; and

- The availability of ARS technology compared to technology from other sources, in particular because of intellectual property issues.

A national focus and a long research time horizon underpin ARS' comparative advantage in bovine quantitative genetics and genomics research. ARS has been able to analyze data from large animal populations over long periods of time. Its research has also contributed to the practical validation of theoretical findings. The singularity of some ARS research outputs does not, however, imply that ARS covers all scientific areas where it has a research advantage, nor that ARS avoids all research in areas where other institutions may have an advantage.

Outcomes of Research and Economic Evaluation

Productivity measures for important outputs in both beef and dairy production show strong rates of increase over time. Average dressed weight for U.S. commercial cattle increased from around 600 pounds in the 1970s to nearly 750 pounds in this decade, for an average annual growth rate of 0.7 percent. These increases probably reflect both higher weaning weights and higher rates of post-weaning weight gain. From 1946 to 2006, the average rate of gain in annual milk production per cow was around 2.4 percent annually. After a period of decline, milkfat percentage has leveled off since the 1970s. In Dairy Herd Improvement Association (DHIA) cows, protein percentages have also been relatively level since the late 1970s. When milk production per cow and component percentages are taken together, the amounts of milkfat and protein per cow per year have also increased significantly over time (ERS calculations from NASS, various years).

Undoubtedly, much of this productivity increase is associated with investments in bovine quantitative genetics. Determining the contribution of genetic improvement more exactly, however, requires further information. It requires understanding the interaction between genetic improvement and other research like animal nutrition or general management. Genetic and management research can improve performance directly in production traits such as carcass meat yield or quality, milk yield per cow, or milkfat or protein content. But genetic and management research can also increase output by improving reproductive or lifecycle efficiency in cattle, particularly in cows. And other factors, such as price policy or farmers' response to changes in both prices and technology, also affect productivity measures. Finally, producer costs in obtaining improved beef or dairy output would have to be netted out from economic benefits to determine dividends from research conducted by ARS, universities, and private sector firms.

Although there have been many economic studies of beef or dairy markets, most have focused on supply and demand parameters, and sometimes the

effects of price policy. Few economic studies of research have focused on beef or dairy research, separate from all livestock or all agriculture, and even fewer have focused on particular aspects of bovine research, such as quantitative genetics or genomics.

Attribution questions are likely to influence the type of economic evaluation conducted for bovine quantitative genetics and genomics research. **Institutional attribution** of economic benefits to particular research investments or research actors is an important issue. Many actors produce and use bovine quantitative genetics and genomics research, as well as the more theoretical sciences of genetics, statistics, and molecular biology that underlie quantitative genetics and genomics. **Technical attribution** of economic benefits refers to sorting out the relative contributions of quantitative genetics and genomics and other research, such as management research. Determining the effects of research on production traits or lifecycle reproduction efficiency could also be difficult.

Both technical and institutional attribution issues suggest that economic studies of the benefits of bovine quantitative genetics and genomics research would be more likely to: (1) be retrospective; (2) use economic surplus as opposed to econometric analysis; and (3) combine the benefits of research performed by different institutions within a single analysis. On the other hand, given that products are sold in two different markets, beef and dairy research analyses would probably be performed separately.

In some cases, providing attribution for research benefits might be easier if a study focused on a particular technology. Economic benefits from a particular technological development would be compared with benefits without that technology. If the analyst's interest is in a particular research performer, such as ARS, a study could concentrate on a technology or technologies for which ARS' role has been particularly prominent. Such studies would nonetheless be stronger and more credible if they also indicated the parts other institutions might have played in the development of these technologies. They could be either retrospective or prospective (or some combination), depending on the technology. Examples of such studies would be an analysis of sire evaluations using DNA data instead of conventional progeny tests, or a study of the relative economic weights to be given to production traits and reproductive/ lifetime efficiency traits in animal breeding.

Case Study: ARS Research on Water Quality and Watersheds

ARS research on water quality and watersheds examines the physical and chemical properties of water and agricultural pollutants, analyzes the use of water in U.S. agriculture and its impact on agricultural production and off-farm water quality, and explores methods to improve agricultural and environmental outcomes. ARS research on water quality and watersheds also integrates water management decisions, such as irrigation, drainage, and flood control. Projects often combine water quality objectives with other goals, such as maximizing crop yields or livestock production while managing runoff. Research results typically have specific applications to agricultural research problems or provide platforms of models, data, and discovery that are accessed by the broader scientific community. ARS scien-

tists conduct research on water quality and watersheds primarily within the National Program on Water Availability and Watershed Management (NP 211).[36]

History

Early USDA water research focused on the role of water in soil erosion and soil salinity, with an emphasis on conserving soil to maintain farm productivity. Subsequent research included water quantity management, especially irrigation and drainage to expand and improve arable land. Following the creation of ARS in 1953, research on water quality adapted to address new concerns. For example, prior research on soil erosion and subsurface soil drains modeled soil movement with the intention to conserve productive soil; ARS adapted these models to examine the effects of sedimentation on downstream water quality. In another example, an experimental watershed system built to examine the impact of flood control structures on large-scale hydrology also included funding for water quality research.

Over time, ARS research began to address water quality more directly. One impetus was the signing of the Clean Water Act in 1977, which expanded the Federal role in regulating surface-water quality. Amendments to the Clean Water Act in 1987 addressed nonpoint sources of water pollution, including agriculture.

The President's Water Quality Initiative, beginning in 1989, emphasized water quality and watershed research. The Water Quality Initiative also emphasized the importance of coordination with other agencies. ARS research informed the creation of the first Hydrologic Unit Areas (HUA) in 1990, a program administered by USDA's NRCS and CSREES to provide assistance where agricultural activities contributed significantly to water quality problems. Also in 1990, the first Management Systems Evaluation Areas (MSEA) were established to coordinate research and management throughout USDA and with other Federal agencies such as the Environmental Protection Agency (EPA) and U.S. Geological Survey (USGS). The Conservation Effects Assessment Project (CEAP) in 2003 continued the trend of increasing the scale of analysis to include larger areas and more complex systems.

Research Scope

ARS water quality and watersheds research is broader than the other two case studies in this report. Research on water quality within NP 211 occurs at very different levels of spatial scale: large-scale research at the basin or watershed level (e.g., remote sensing that provides satellite data of sediment transport phenomena); field-scale research (e.g., experimental plots that test crop yield and runoff under different water management treatments); and laboratory research (e.g., chemical properties of agricultural pollutants). Although field-scale experiments are necessary to understand the impact of agronomic practices, larger scale studies that include numerous producers within an agriculturally productive watershed are necessary to show the cumulative impact of nonpoint-source pollution. Unlike a point source of water pollution, which is traceable to discharge at a particular place and time, nonpoint

sources are more diffuse and difficult to trace, and therefore more difficult to monitor or regulate.

ARS research on water and water quality engages a greater number of related research topics than most other USDA-supported research. Table 8 summarizes 30 years of USDA-supported water research as reported to the Current Research Information System (CRIS).[37] Compared with research performed or funded by other research institutions or agencies, ARS water research spans the widest range of research topics (140). Despite the wide breadth of topics on which ARS conducts research, ARS research projects exhibit greater focus than other public sector water research projects, as demonstrated by the low average number of research problem areas (1.71) per project. The average ARS project size, in dollar terms, is larger by far than research projects from other funded sources.[38] Together, these facts are consistent with a science agency that can weigh in authoritatively on narrowly defined research on a broad range of subjects.

This case study focused on a subset of ARS water quality and watersheds research, attempting to narrow the topic while still examining the broad scientific and economic relevance of ARS research. Interviews emphasized sedimentation and sediment transportation, soil drainage, the development of physical modeling tools, and the interaction of ARS scientists with counterparts at other agencies. Other topics—such as soil salinity, hypoxia, and irrigation—were touched on but not explored in as much depth.

Outputs

The chief output of ARS research is oriented toward publication of scientific results in journal articles, monographs, books, and other outlets. Scientists in NP 211 publish hundreds of scientific results every year. Results of ARS research are in widespread use, in four distinct areas: (1) specific technolo-

[37]The table describes all research projects with a commodity or subject of inquiry (SOI) of "water," or with one of the following five research problem areas : (1) conservation and efficient use of water; (2) pollution prevention and mitigation; (3) drainage and irrigation systems and facilities; (4) soil, plant, water, and nutrient relationships; (5) watershed protection and management. This is not a perfect intersection with NP 211: it includes research projects with a primary classification in other national programs, it excludes certain water quality and watershed research projects, and its long time period spans different programmatic emphases.

[38]CRIS funding reports for ARS include full-time equivalent salaries for researchers, but might only include part-time salaries for researchers at other institutions through different funding mechanisms. Because salaries generally comprise a majority of agricultural research expenses (Klotz et al., 1995), this might overstate ARS project size relative to other categories to some extent.

Table 8
USDA-supported water research, 1973-2003

Funding type	Number of projects	Average funding per research project (2006 dollars)	Average number of research problem areas (RPAs) per research project	Total number of RPAs addressed by all water research projects (out of 158 possible)
Formula funds (Hatch, McIntyre-Stennis, etc.)	17,769	67,910	2.05	118
USDA in house (ARS)	8,245	573,254	1.71	140
Cooperative agreement	1,232	126,934	1.64	74
Competitive grant	1,078	50,253	2.30	51
Other	2,098	98,917	2.19	80

Source: Economic Research Service; Current Research Information System.

gies developed by ARS scientists; (2) applied science, often from physical models of erosion, pollutant fate and transport, water quality, and the like; (3) tools for policy guidance; and (4) further scientific advances like understanding biochemical processes in soils.

Specific technologies arising from ARS research are sometimes licensed or offered for sale. Examples include irrigation and subsurface drainage control equipment, improved products to limit nutrient leaching into waterways, and other inventions that are either based on or relevant to NP 211 research. In addition to licensing these technologies, ARS sometimes releases them into the public domain to promote widespread commercial use (Heisey et al., 2006).

ARS applied science includes biophysical models that relate climate, geography, and agronomic practices to the physical, biological, and chemical processes of water moving through agricultural areas to estimate their impacts. The Universal Soil Loss Equation (USLE), developed by scientists in the ARS National Soil Erosion Laboratory in 1958, is an example. It predicts annual soil loss per acre based on factors including rainfall, field slope, and management practices. Updates and improvements include the Revised Universal Soil Loss Equation (RUSLE), RUSLE2, and the Water Erosion Prediction Project (WEPP). These models are used by other USDA agencies to assist agricultural management, with appropriate modifications for users in specific regions or with specific needs. Models and applications are often distributed electronically at no charge via CD-ROM or over the Internet. Users include:

• Extension specialists and NRCS agents helping agricultural producers to analyze and manage environmental impacts of their operations;

• Developers of erosion management plans to meet USDA conservation compliance requirements on highly erodible land;

• State, Federal, and international regulatory authorities who promulgate rules and regulations for agriculture and other environmental and land-use applications; and

• Outside researchers who use the models to support and validate their own models.

NP 211 research also influences water quality policies and programs through the development of technologies that can create new policy options. The Conservation Reserve Program (CRP), the Conservation Stewardship Program (CSP), and the Environmental Quality Incentives Program (EQIP) are examples of USDA programs that utilize ARS science to direct conservation program funding to agricultural producers. CRP and CSP are voluntary programs that provide payments to farmers who undertake resource conservation with environmental benefits. Similarly, EQIP provides financial and technical assistance to producers who address natural resource concerns through installation and maintenance of subsurface drains, irrigation controls, vegetative buffers, and other investments.

Eligibility and payments for CRP are based in part on an Environmental Benefits Index (EBI) that includes predicted water quality and soil erosion

impacts from ARS modeling efforts. ARS research on erosion, soil quality, and organic material was also incorporated into the Soil Conditioning Index (SCI), which is used to calculate some CSP payments. Similarly, many of the abatement techniques and benefit estimations for EQIP incorporate ARS research, either in the way States award EQIP contracts or in the development of eligible EQIP practices (for example, research on vegetative buffers or subsurface drains that curb erosion or improve yields).

Other Research Actors

ARS conducts joint research with other USDA agencies and other research-performing institutions funded by USDA (notably land-grant universities funded by CSREES, now NIFA). In some circumstances, ARS contributes research personnel and other resources. In others, ARS assumes a coordinating role among other science and program agencies. Other agencies have complementary capabilities, such as field offices for data collection and applied orientations. NP 211 also brings scientific and agricultural expertise to State, regional, and national water quality issues. NP 211 scientists interact with researchers at EPA, USGS, the U.S. Army Corps of Engineers, and numerous other Federal agencies, especially through intergovernmental and regional committees.

One ongoing example of the way ARS interacts with other agencies is the Conservation Effects Assessment Project (CEAP) to determine the impacts of conservation activities. In addition to five lead agencies (NRCS, ARS, NIFA, FSA, NASS), CEAP includes input from several other USDA and Federal agencies. The goal of CEAP is to (1) identify conservation practices and techniques currently in use by private agricultural landowners, (2) identify the USDA role in developing them, and (3) determine their impacts on water quality and other environmental concerns.

Scientists at other Federal research institutions also perform water quality and watershed research at different scales. ARS research is especially focused on agricultural sources of water quality concerns. Since agriculture is both a major water user and a large contributor to water quality problems, ARS research with detailed information on agricultural practices has few substitutes. Thus, soil and water models linked to agricultural production practices constitute a particular area of comparative advantage for ARS water quality research.

Evaluation of Economic Benefits

Each of the four areas of ARS research use identified—technologies, applied science, policy tools, and further scientific advances—presents different opportunities, challenges, and roles for evaluation of economic benefits. Economic evaluation often tries to estimate dissimilar benefits in terms of their monetary value, which can then be compared with costs and alternatives for providing those benefits. Evaluating the benefits of technologies for sale or license is perhaps most straightforward because they can be quantified with market price signals to determine value. Even then, the value of ARS technology in commercial use is not always easy to establish if it is released to the public domain. Other situations also require analysis to estimate

physical quantities utilized, effects on prices, or other observable results of economic behavior.

Economic evaluation in water quality involves the additional challenge of assessing economic values for goods and commodities that are not traded extensively in markets. In most cases of NP 211 research on water quality and watersheds, some primary users of ARS technologies are other public sector actors, and therefore benefits are difficult to assess. Other effects of water quality have indisputable economic value, but the precise value of these effects is difficult to measure. One example is human health, which can be estimated by the cost of health interventions necessary to treat health conditions resulting from water quality impairments, or the cost to a water treatment plant of attaining EPA drinking water standards. For instance, a 2006 benefit-cost analysis (CVision Corporation, 2006) estimates that widespread installation of subsurface drainage control units in the Upper Midwest could reduce agricultural nitrate runoff by 128 million pounds. In another study, Des Moines Water Works estimated that denitrification to meet drinking water standards requires operating costs of $6 a pound. Thus, benefits of $768 million are plausible from the drainage technology.

The concept of quality-adjusted life years also attempts to estimate the value of reduced life expectancy or reduced life enjoyment because of poorer health. Costs of treatment and reductions of life expectancy can be estimated actuarially or epidemiologically, but the value of good health is subjective and likely varies among people. The aesthetic value of pristine water and natural wildlife are also goods that have economic, but subjective, value that must be estimated carefully. Contingent valuation, which relies on responses to survey questions about valuation, and travel cost methods, which rely on people's observed willingness to incur expenses while traveling to natural amenities, are economic techniques to infer economic valuation of nonmarket goods (Feather et al.,1999).

Even when water quality affects a commercial activity such as electricity generation or barge shipping, estimation of the economic value of water quality can be technically challenging. Hansen and Ribaudo (2008) provide regional estimates of soil conservation benefits ranging from $0 to $8.81 per ton of runoff for each of several types of downstream activities.[39] But they caution that not all activities are relevant within regions, and that location-specific effects within regions can significantly increase or reduce soil conservation benefits.

[39] The National Resource Inventory (USDA NRCS, 2007) estimates that annual water erosion of U.S. cropland soil was 2.6 tons/acre in 2003, for a total of 970 million tons.

An important benefit of the modeling tools developed from applied ARS science is the ability to estimate alternative scenarios. To estimate the benefit of a policy, or the environmental consequences of no change, one must estimate what would have occurred, even though only one set of conditions can actually be observed. For instance, many of the examples of economic evaluation in Hansen and Ribaudo (2008) are based on ARS research; the CVision (2006) study was based in part on DRAINMOD, a biophysical model developed with support from the CSREES National Research Initiative. The use of models to develop counterfactual scenarios operationalizes the value of research; the estimated results of decisions made with and without the new information provided by science-based models are observable and comparable.

Accordingly, one contribution of ARS research is to improve estimates of cost-benefit analyses, and to make new analyses possible. ARS models are used extensively by other agencies in determining the effects of changes in policy. Regulatory agencies such as the EPA perform numerous cost-benefit analyses of investments in water quality and water pollution abatement. Other examples of cost-benefit analysis can be found outside of the public sector (CVision Corporation, 2006). These studies show extensive use of ARS science, and illustrate some of the complexity and detail required for cost-benefit analyses.

Another example of this type of research benefit is the use of targeting (Claassen, 2003; Walter et al., 2007). Targeting in conservation programs is intended to channel more abatement toward environmentally sensitive areas, large sources of pollution, or sources of abatement that can be obtained with minimal costs. ARS research identifies the characteristics and locations of areas that make targeting worthwhile. However, targeting benefits in this way might require program changes, possibly including new legislation.

Case Study: Nutrient Data Laboratory

The Nutrient Data Laboratory (NDL) is one of six units in the Beltsville Human Nutrition Research Center within ARS' Human Nutrition National Program (NP 107). USDA has provided data on the nutrient composition of foods in the American diet for 115 years through a series of institutions that have evolved to be the NDL. The NDL assesses more than 7,000 food items for up to 140 nutrients. Sampling one food, including sampling cost and nutrient analysis, but not including salaries or other fixed expenses, costs approximately $17,000-$20,000.[40] Sources of USDA's food composition data include original analytical data generated through the NDL's National Food and Nutrient Analysis Program, as well as data for selected nutrients and foods contributed by the food industry, collected from the scientific literature, or from other agencies (e.g., FDA's Total Diet Study).

[40]The chemical analyses are performed by universities, private contractors, and, for special nutrient analyses, the Food Composition and Methods Development Laboratory, also part of the Beltsville Human Nutrition Research Center in Beltsville, MD.

The food composition data developed by NDL are reported in the National Nutrient Database for Standard Reference. The Standard Reference (SR) databases are the foundation of virtually all public and commercial nutrient databases used in the United States and a number of other countries (NP 107 review panel, 2006). Supplementary databases include fluoride, isoflavone, and choline content of selected foods. In addition to continually updating values for new foods, the NDL databases in recent years have added foods specifically addressing items commonly consumed by minority populations, such as Hispanics and Native Americans. These databases are the NDL's primary output. Additional outputs include publications in the professional literature regarding food composition research, data analysis reports, proceedings of the National Nutrient Databank Conference, presentations at scientific meetings, and direct communications with food processing companies and other professional users of the NDL's output and with the general public. Finally, NDL scientists share food composition research methods for food sampling, sample handling, analytical techniques, data quality assessment, and data compilation with scientists in other countries.

Funding

The NDL's budget has varied over 2002-07. In 2007, total funds were $4.9 million when converted to 2006 dollars by using a research deflator: $3.1 million in gross appropriated base funds and $1.8 million[41] from other Federal agencies, chiefly NIH, usually through specific agreements for specialized purposes (fig. 8).

NDL funds have declined in both real and nominal terms since 2002, when they totaled more than $7 million (in 2006 dollars). Base appropriated funds from USDA have remained roughly constant in real terms; declines in other Federal funds and personnel cuts have reduced research expenditures.

At the time the case study was conducted, the NDL had six permanent scientists on staff, having lost staff members during 2007. Six support scientists worked on the database, and five additional support scientists were funded by industry or with "soft funds" (i.e., NIH funds and funds that are not direct Congressional appropriations to ARS). The contract employees were hired to work on specific databases, which, along with diminishing funds and retirements, account for some of the variation in staffing levels.

Use of NDL's Output

NDL users fall into several categories: medical and nutrition researchers, public policymakers, regulators, the food and dietary supplement industries, educators (public and private), and secondary users. Nearly everyone in the field of nutrition, whether doing research or applied work, draws upon NDL information (M. Kretsch, ARS, personal communication, Oct. 12, 2007). NDL information underlies any effort to determine, understand, or comment on the connection between nutrition and health. The NDL's outputs serve most frequently as tools for other research and information creation.

[41] The $1.8 million figure is based on "net to location" data, to keep it consistent with the other data in figure 8.

Figure 8
Real Nutrient Data Lab budget summary, 2002-07

$ million (2006)
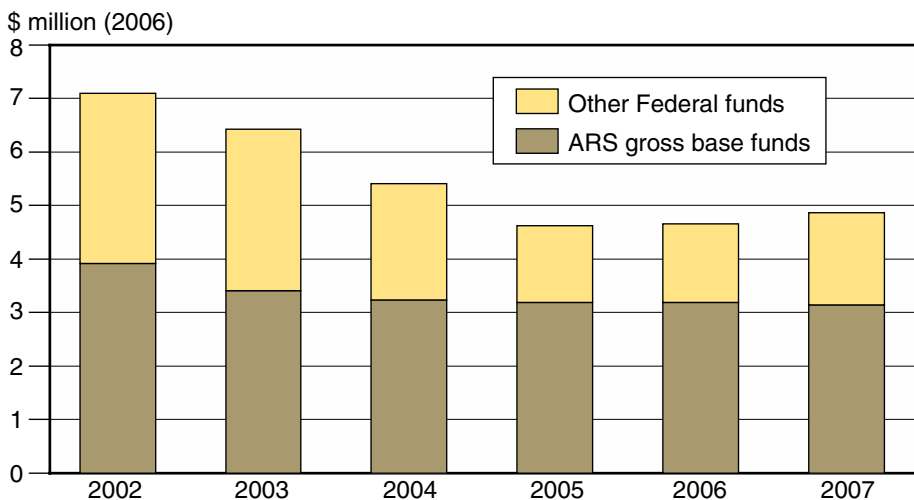
Source: Nutrient Data Laboratory, ARS.

Research

Medical research. The Standard Reference database (SR) is used in medical research to determine the role of nutrition in the causation of disease and modifying risk factors, and the efficacy of different nutrition-related treatments. The SR data are combined with dietary intake data for epidemiological and population-based nutrient/disease studies (A.G. Ershow; National Heart, Lung, and Blood Institute; NIH; personal communication July 10, 2007).

Nutrition research. The SR is used in research to better understand the biological functions of different nutrients. Research on normal metabolism functions (both clinical and basic) provides a baseline for metabolism-related studies. The NIH's Office of Dietary Supplements supports the NDL's development of the Dietary Supplement Ingredient Database (DSID) to provide information on the composition of dietary supplements commonly used by the U.S. population.[42] The Office of Dietary Supplements also uses the DSID and SR for information about the nutritional content of dietary supplements and food additives.

[42] www.usda.nlm.gov/DSID

Technological research. The NDL conducts research to determine the most effective and accurate methodology to analyze what is in foods. Food sampling is a significant methodological challenge: the NDL assessed various methods to determine key foods that are the primary contributors to nutrient intakes in the United States and now collects these foods to be tested periodically (Pehrsson et al., 2000). Because the size and complexity of the NDL database necessitate a comprehensive quality control plan (Bodner-Montville et al., 2006), the NDL continually evaluates methods for data processing and dissemination.

Public Policy

The NDL's work is used in almost every aspect of nutrition policy The SR underlies the *Dietary Guidelines for Americans*, which form the cornerstone of Federal nutrition policy and nutrition education activities (Dietary Guidelines Advisory Committee, 2004). Federal nutrition assistance programs such as USDA's Child Nutrition (school meals) and Supplemental Nutrition Assistance (food stamp) Programs, and the WIC program (Special Supplemental Nutrition Food Program for Women, Infants, and Children) use the *Dietary Guidelines* to help design benefit structures and nutrition education. The Department of Defense also uses NDL data in the formulation of its feeding program (OMB, 2006).

Information in the SR also helps international organizations plan for interventions in countries with malnutrition by providing a means to measure the nutritional content of the food supply. Countries throughout the world, especially developing countries without their own food nutrition composition databases, use the USDA food composition databases to develop their dietary guidelines and national nutrition plans (B. Burlingame, UN FAO, personal communication Nov. 14, 2007). NDL scientists also share research methods on determining food composition (sampling, sample handling, data compilation and analytical techniques, and data quality assessment, for example) with scientists in other countries.

In addition, consumers use the NDL's research in formal and self-directed nutrition education programs. Nutrition professionals use the SR databases to educate clients and to construct dietary plans for individuals, hospitals, schools, and other institutions. Education programs use the SR in textbooks and student software.

Regulatory Requirements

Research conducted by the NDL also supports several regulatory efforts. The Child Nutrition Database of the Food and Nutrition Service, USDA, is based on SR and is used to create plans for meals at schools and day care centers. SR data underlie both the regulations for school meals and the assessment of compliance with government rules.

FDA uses NDL data in the regulation of nutrition labels. For example, the FDA uses the SR in reviewing health claims and determining the nutrient content of foods that may be eligible for proposed claims. The FDA also regulates nutrition labeling claims and values.[43] USDA's Food Safety and Inspection Service also supports the use of SR data in the labeling of meat and poultry products.

Use of the NDL databases to inform regulation is not limited to the United States. The Codex Alimentarius Commission was created in 1963 by the Food and Agriculture Organization of the United Nations (FAO) and the World Health Organization (WHO) to develop food standards and guidelines. The FAO provides scientific advice to Codex committees using USDA's food composition database and other sources (B. Burlingame, personal communication).

Assistance for Government Food Surveys

The SR is used to support the Food and Nutrient Database for Dietary Surveys (FNDDS) and search tools from the Food Surveys Research Group, also at the Beltsville Human Nutrition Research Center. Food intake data from the National Health and Nutrition Examination Survey—conducted by the National Center for Health Statistics of the Centers for Disease Control and Prevention, U.S. Department of Health and Human Services—are combined by the Food Survey Research Group with the NDL's databases to create What We Eat in America (WWEIA), the only nationally representative and major continuing dietary and nutrient intake survey of the U.S. population.

Customized Databases

The SR supports many customized databases used in public and private research. One such database is the University of Minnesota's "Nutrition Data System for Research," a proprietary software system geared toward clinical research and epidemiological studies. The Nutrition Data System for Research combines data from the SR with highly specific information on dietary intake to create specialized databases that allow for the comparison of dietary intake over time. Other examples of customized databases are those geared toward the generation of nutrient labels, as seen in the Nutrition Facts Panels on U.S. retail products. While the food industry is responsible for the

[43]For example, in 2006 the FDA used the SR in updating guidelines for voluntary in-store labeling of the 20 most frequently consumed raw fruits, vegetables, and fish to enable consumers to make more informed purchasing decisions.

content of the Nutrition Facts Panels, much of this information is based on data obtained from the SR.

Evaluation of Economic Benefits

The primary benefit of the NDL, common to all its uses, is improved public health. Public health can be measured and, to a degree, valued, through estimates of health care costs, loss of productivity, and measures such as quality-adjusted life years. However, connecting public policy interventions (e.g., dietary recommendations, nutrition requirements, and food labeling) with improvements in public health is difficult. The role of nutrition is recognized, but not fully understood. The relationship between health, disease, environment, behavior, and genetics is unclear. Moreover, measurement of dietary intake is incomplete. While WWEIA gives a picture of what Americans eat, it is not conclusive. A better understanding of nutrient intake helps to fashion effective policy interventions and nutrition education. Still, it is unlikely that the results of most nutritional interventions could be measured accurately.[44] And there is no reliable method of apportioning the contribution of the NDL relative to that of other scientists and the policy process.

On the other hand, the health benefits of improved nutrition information clearly can have a huge value. Variyam and Cawley (2006) estimated the effects of the Nutrition Labeling and Education Act (NLEA) of 1990 by comparing body weight changes of individuals who used labels for food selection before and after NLEA's implementation with changes in body weight over the same period for individuals who did not use labels.[45] For non-Hispanic White women, the total monetary benefit of decreased weight over a 20-year period was an estimated $63 billion to $166 billion.

Focusing specifically on the NDL presents some unique challenges to conventional research assessment techniques. The NDL's SR and DSID are the major outputs. The NDL database products are available free of cost and are accessed primarily through the Internet. Secondary uses of the databases are significant, so it is difficult to measure use of the NDL solely through website hits alone. The SR is used so widely that it often goes uncited, even among research scientists. Also, the SR is frequently used in regulatory and policy planning functions and results of regulatory science may not be published because the work is classified, confidential, or "politically volatile" (Ershow, 2003). Methods of delivering or utilizing SR data may be appropriable, but the underlying information has public goods characteristics and cannot be protected with intellectual property.[46] Therefore, methods such as citation or patent counts would not fully capture the benefits of the NDL.

Alternatives to the NDL

One user characterized the SR as the "numerical bedrock of all human nutrition research" (A.G. Ershow, personal communication). Since no equivalent research product exists, there is nothing to which it can be easily compared. Theoretically, nutrient values for specific items could be measured by others, though at substantially higher cost and reduced reliability (J. Dwyer, Office of Dietary Supplements, NIH, personal communication July 23, 2007). However, determining nutrient values of individual items could not replace

[44]For example, vitamin D affects absorption of other nutrients, secretion of certain hormones, skeletal health, and immune system functions. Different components of vitamin D appear to differ in their benefits. The vitamin D content of a given food can vary significantly from sample to sample, presenting methodological challenges for food composition analysis. Moreover, there is uncertainty about the estimation of the intakes of food that contains vitamin D. Finally, even if vitamin D levels could be better measured, the exact relationship between vitamin D levels and public health is not clear. Therefore, while ensuring adequate vitamin D levels would enhance public health, measuring such benefits from an economic perspective would be extremely difficult.

[45]Previously, labeling occurred under voluntary rules established by the FDA.

[46]Even in cases where proprietary databases or software are based in part on the SR, observed license fees may not reflect a "market value" or "willingness to pay." For example, the University of Minnesota's NDSR software/database is licensed to roughly 200 institutions, with about 700 copies in use. The fees for these licenses are called "support fees." Many of the copies are actually used by NIH, and further copies may be used by academic researchers, who could also be partially funded by NIH grants. The NDSR is not currently raising support fees because so many users are currently constrained by lower grant awards. Furthermore, recall that the NIH has also funded part of ARS' work on the Standard Reference. In summary, NIH is active on both sides of the purported "market"—data supply and data use—so it is unlikely that support fees could be used as the basis for market valuation.

a nationally based, census-driven, food composition measurement system, backed by the credibility of ARS and the rigor needed for statistically significant medical studies (Pehrsson et al., 2003), and this is one major area in which ARS research has a particular comparative advantage.

Peer Review and Efforts To Maximize Research Returns

The NDL program has been evaluated by two standard mechanisms used to review ARS research—peer review through the ARS Office of Scientific Quality Review (OSQR), and an OMB Program Assessment Ratings Tool (PART) review. The OSQR external panel found the quality of NDL research accomplishments to be high, and noted that the SR is the most comprehensive and highest quality information of its type in the world (NP 107 Panel, 2006). The PART review of the Human Nutrition National Program found the NDL component was meeting its targets in terms of reports. The assessment also said that the NDL's research is used extensively by other Federal agencies, other countries (e.g., Korea) and international organizations, (e.g., FAO), and provided specific examples (OMB, 2006).

Past peer review has led to significant improvements in the analytical quality of nutrient data provided by the NDL. In 1997, the National Food and Nutrient Analysis Program (NFNAP), a multi-agency effort (led by USDA and NIH) distinct from the ongoing operations of the NDL, was initiated (NDL, n.d.). The NFNAP proposal was subject to peer review before the work began, and the NDL was required to address certain weaknesses in advance. According to Ershow (2003), the prospective analysis and the lab's response significantly raised the proposal's credibility. NFNAP has focused on five goals: (1) improving information on key foods (i.e., regularly consumed foods), (2) evaluating the quality of existing data, (3) developing nationally based sampling plans, (4) analyzing samples using state-of-the art methods, and (5) compiling newly generated data to achieve a current and representative food composition database. NIH, CDC, FDA, and several industry associations have contributed to the support of NFNAP.

NFNAP has expanded the NDL's work into more basic forms of research. The NDL work on sampling technologies has increased the quality of the SR databases and has contributed to understanding sampling methodology. The lab is producing more peer-reviewed, professional publications. The burnishing of NDL's science credentials has enhanced its reputation and likely improved its standing in peer-review metrics. However, NFNAP's greatest value is in improving the accuracy, reliability, and breadth of food composition data, whether it be used for science, regulation, policy, or nutrition education.

Analyzing the Research Environment of the Case Studies

Despite the breadth of ARS research and the wide differences among the three case study research programs, it is still possible to systematically compare the case studies. Table 9 specifies intermediate and end users of ARS research outputs for all three cases and identifies the areas where economic impacts of the research might also be found. As hypothesized, both the water quality/watersheds case and the Nutrient Data Laboratory

Table 9

Case study comparison—research environment

	Bovine Quantitative Genetics and Genomics	Water Quality and Watersheds	Nutrient Data Laboratory
Other public sector research providers?	Land grant and other universities with specific beef or dairy improvement research	• Numerous Federal agencies • Universities • Nonprofit institutions	University food composition data, which generally are either for specific components or synergistic data that build on NDL's unique, nationally based, census-driven effort
Private sector research providers?	• Private research focused on germplasm (AI, seedstock companies) and related technologies • Private DNA technology companies	Very limited	Providers of food industry composition data for specific products and proprietary databases based in part on Standard Reference
Other sources of supply for ARS research outputs? (Particular areas of comparative advantage)	Limited—ARS results are often based on data from large animal populations over a long period of time, or ARS results provide practical validation of theoretical findings	Limited—particularly with respect to soil and water models linked to agricultural production practices	Very limited—no other source of information is nationally based and census driven
End users and other stakeholders	• Breed associations • Seedstock companies • AI companies • DNA technology companies • University research programs • Beef and dairy producers	• Farm conservation programs • Federal and State water quality agencies • Producers (e.g., subsurface drainage research, irrigation/water re-use research)	• Medical and nutritional researchers • Public policy programs (e.g. nutrition programs) • Regulatory agencies (e.g. FSIS/FDA regulation of labeling) • Consumers
Research output serves other public research, regulatory, or policy function?	University beef and dairy improvement research may use ARS research results (or vice versa)	• Analytical tools for conservation program implementation and evaluation • Widespread use in State, EPA regulation	Standard Reference used for • Medical and nutritional research • Regulation of school and daycare meals, nutrition label claims • Creation of *Dietary Guidelines for Americans*, which guide nutrition programs and policy
Coordinating role?	• Evaluation of national dairy data • Leadership in bovine functional genomics	Development of technologies that improve water quality and methodologies for measuring physical impacts on water quality	NDL is guided by needs of stakeholders both public (e.g., NIH, FDA, and other USDA agencies and consumers) and private (e.g., food industry)
Observable prices in ultimate impact market	• Direct for beef and dairy products • Indirect for characteristics or components of meat and milk • Not observable for environmental effects of beef and dairy production	Water quality is essentially nonmarket; techniques for valuing water quality are indirect	Ultimate benefits are human health; techniques for measuring value of human health are indirect

case illustrated situations in which ultimate benefits—higher water quality, better human health—were in goods for which market prices are generally not observable. Although indirect valuation techniques exist for both environmental and health goods, they are complex, and reinforce the difficulty of economic evaluation using standard tools. Bovine quantitative and genetics research exists within the more traditional commodity research areas. Ultimate outputs of this research can be measured directly by market prices for the primary products (beef and milk) and indirectly from market prices for characteristics or components of the primary products (e.g. beef quality, milkfat or milk protein percentages). Environmental effects of beef or dairy production would, however, generally not be directly measurable using market prices.

All three research programs exist in a complex research environment, with multiple actors, both in terms of other research providers and other users or stakeholders of the ARS research. Nonetheless, in all three cases ARS provides research that users found valuable, and that would be incomplete, costlier, or nonexistent were stakeholders to seek the research outputs from alternative suppliers. This suggests that in these three cases, ARS research has public goods aspects. Research linkages, however, differ among the three cases. The Nutrient Data Laboratory may have the most "vertical" set of linkages, with numerous pathways between the output of the laboratory and the ultimate impact area of human health. Water quality and watersheds research may have the most "horizontal" set of linkages, with numerous Federal agencies, as well as universities, performing some kind of water quality-related research. Bovine quantitative genetics and genomics is an intermediate case, with both vertical linkages to underlying sciences and to private companies that focus on germplasm or DNA technology, and horizontal linkages to quantitative genetics and genomics research in universities.

In all three areas, there is relatively little private sector provision for the types of research that ARS performs. Private sector research in the water quality area is probably the most limited. Some private sector database developers or companies create proprietary databases of nutrient data, but all are built to some extent on the Standard Reference maintained by the Nutrient Data Laboratory. Private research focused on animal germplasm tends to generate either genetic material such as semen, embryos, or live animals, or related reproductive technologies, rather than the statistics-based outputs of quantitative genetics research. Private research from DNA technology companies and ARS is often complementary—the private companies may provide sequencing and DNA platforms, while ARS contributes information and software tools.[47]

Similarly, for all three case studies, ARS research often fills a niche that other public sector institutions do not. Numerous Federal agencies concern themselves with water quality and watershed research, but ARS is the leader in linking soil and water models to agricultural production practices. Also, as the case study shows, ARS performs more water quality research than universities, and ARS water quality projects tend to be considerably larger than projects from other public sector sources. In the nutrient case study, many databases are built in part on the Standard Reference maintained by ARS' Nutrient Data Laboratory. Only the Standard Reference is both nationally based and census driven. In contrast to water quality research, more

[47]See J. Karow, "USDA-Led Team Uses Illumina Sequencer to Discover, Describe New SNPs in Cattle," In Sequence, March 4, 2008. http://www.in-sequence.com/issues/2_10/webreprints/145538-1.html

research money is spent by universities than by ARS on bovine quantitative genetics and genomics. ARS' NP 101, however, tends to concentrate its research on areas with a national focus and a long time horizon. ARS research is particularly valuable when data analysis for large animal populations over a long period of time is required.

In all three case study areas, ARS both uses research results from other sources and provides research that other public sector research is based on. Other research building on ARS output is most evident in the NDL case, where medical and nutritional research studies often begin with Standard Reference information. Other instances can be found in the other case studies; for example, ARS maintains DNA data from bull semen provided by artificial insemination companies, and university researchers use this DNA data to relate the genetic information to superior traits. The most straightforward regulatory and policy functions associated with ARS research are in the water quality/watersheds area, where ARS provides analytical tools for conservation program implementation and evaluation. ARS tools are also used widely in the development of EPA and State regulations. The NDL Standard Reference also serves important regulatory and policy roles with respect to regulation of school and daycare meals, nutrition label claims, and development of the Dietary Guidelines for Americans.

Feasibility of Economic Evaluation Methods for the Three Case Study Programs

The preceding discussion has indicated that certain types of economic analysis can be applied to all three case studies. For example, in all three cases it is possible to identify the areas of ultimate economic impact. It is also possible to identify related research, and to explore the areas in which public sector research at the national level, such as that conducted by ARS, might concentrate its efforts with the greatest comparative advantage. On the other hand, in all three case studies, the uncertain attribution of research benefits to particular research actors or particular research cost streams poses particular difficulties for economic analysis. Furthermore, in both the NDL and water quality/watersheds cases, the nonmarket nature of the ultimate goods associated with the research makes benefit measurement difficult (table 10).

Measuring benefits in general appears most difficult for the NDL case study. A number of different economic methods are available for valuing human health. However, human health benefits would need to be defined appropriately, with the understanding that nutrition is only one of several factors affecting health. Furthermore, there are different types of nutrition-related pathways to human health outcomes[48] and there are several different means by which one factor, the nutrition information provided by the Standard Reference, will influence the well-being of individuals.

Water quality benefits are also nonmarket, and difficult to measure. Economists have developed methods for measuring environmental benefits, but applying these measures requires more data than measuring benefits in a commodity market with observable prices, and thus data collection and analysis is costlier. Attribution of benefits could well be incomplete, or "superattribution" of benefits could occur if multiple actors are indispensable

[48]For example, nutrients may come from diet, supplements, or even be induced by the effect of sunshine on skin. Different forms may have different effects. Nutrients may need to be consumed in combination to maximize their effectiveness. Medical conditions vary in their response to nutrients.

to ultimate water quality impacts and each has a claim on the same estimated benefit. If benefits of water quality and watershed research could be given a spatial focus, more complete accounting might be possible. Therefore, quantitative economic studies might focus on a particular watershed rather than on a complete research program.

For the NDL case study, we conclude that both standard economic evaluation methods—econometric and economic surplus analysis—are likely to be infeasible. For the other two case study areas, the fact that the research effort is fairly disaggregated is likely to make econometric analysis particularly difficult to apply; furthermore, the nonmarket aspects of water quality suggest that economic surplus analysis would be more likely to be feasible in a spatially focused study. Economic surplus analysis would probably be most feasible for certain aspects of research in bovine quantitative genetics and genomics, like evaluating the economic impact of research on dairy or beef output.

In the application of standard economic methods, attribution issues would be important. In the NDL case, for example, there are many steps, both in research and policy, between the Standard Reference and the major impact area--human health. There are also many other contributing factors to this ultimate impact area. In the water quality and watersheds research case, there are other research actors with related research programs, and other

Table 10
Case study comparison—economic evaluation methods

	Bovine Quantitative Genetics and Genomics	Water Quality and Watersheds	Nutrient Data Laboratory
Quantitative economic evaluation feasibility	Yes, with qualifications	• Possible with a spatial focus for more complete accounting • Methodology for hard-to-measure benefits has greater costs, data needs • Benefit attribution could be either incomplete (some missed) or "superattributed" if multiple actors are indispensable	• No, costs easily measured but most benefits are diffuse, nonmarket goods • Attribution issues would be significant
Econometric analysis feasibility	Less feasible because of likely attribution issues (e.g., genetic vs. management research)	Probably less feasible, even in a spatially focused study	See above—unlikely to be feasible
Economic surplus analysis feasibility	More likely to be used, both retrospectively and prospectively	More likely to be feasible in a spatially focused study	See above—unlikely to be feasible
Attribution issues	• Other research actors with related research programs • Other actors using or contributing to ARS' research • Contributions of genetic vs. management research	• Other research actors with related research programs • Many other contributing factors to ultimate impact area, particularly environmental policy	• Many steps (both in research and policy) between Standard Reference (SR) and major impact area, e.g., human health • Many other contributing factors to ultimate impact area

Assessing the Benefits of Public Research Within an Economic Framework / ERR-95
Economic Research Service / USDA

contributing factors to water quality, especially environmental policy.[49] In bovine quantitative genetics and genomics, there are other research actors, like universities, with related research programs, and other actors who use or contribute to ARS' research.[50] In addition, this case study illustrates the uncertainty of technical attribution within the research itself. To what degree is an increase in beef or dairy productivity attributable to genetic improvement versus management innovations?[51] Accounting for different pathways to economic impact would be an important part of analysis in all three case studies, whether or not the analysis was quantitative.

Other Evaluation Methods and the Case Studies

We briefly considered other, non-economic evaluation methods for the three case studies. Certain features of these methods and of the cases themselves make them more or less feasible (table 11). For example, one non-economic quantitative assessment tool, bibliometrics, could be applied in both the water quality/watersheds case and the bovine quantitative genetics/genomics case. It would be less feasible in the NDL case. Citations to its major output, the Standard Reference, are often omitted from subsequent scientific publications that use the database, or subsumed under citations of secondary scientific use.

Some form of benchmarking, the measurement of progress against internally determined performance measures, would be possible in all three case study areas. A modified form of benchmarking, in which the research program is compared to similar research programs, would likely be feasible for some aspects of water quality/watersheds research and bovine quantitative genetics/genomics. It would be less feasible for the NDL, because a research product comparable to the Standard Reference does not exist.

Other qualitative assessment tools such as narrative assessments or retrospective analysis might be applied to all three case study research programs, with qualifications. Narrative assessments might have less credibility than other forms of evaluation because they are less likely to follow a prescribed approach to evaluation. Retrospective analysis might overlap with economic analysis, particularly for the bovine quantitative genetics and genomics case,

[49] Research and environmental policy may be particularly likely to complement each other in the water quality case.

[50]To the extent that ARS takes a leadership role in providing forums for this larger research community to exchange ideas and plan broad research areas, part of the value of ARS' research could stem from this coordinating role. In a number of research areas, like agricultural animal genomics, this appears to be likely.

[51]Even within genetic improvement, there are two pathways to greater productivity—through production traits and through improved lifecycle and reproductive performance.

Table 11
Case study comparison—other evaluation methods

	Bovine Quantitative Genetics and Genomics	Water Quality and Watersheds	Nutrient Data Laboratory
Feasibility of other quantitative evaluation methods (e.g., bibliometrics)	Bibliometrics feasible	Bibliometrics feasible	Bibliometrics limited—NDL has primarily one product, citations of Standard Reference are often omitted or subsumed
Peer review	Normal practice	Normal practice	Normal practice
Narrative assessments	Yes, but may be less credible	Yes, but may be less credible	Yes, but may be less credible
Benchmarking	Possible application for some aspects of the program	Possible application for some aspects of the program	Comparable product does not exist
Retrospective analysis	• Might overlap with economic analysis • More feasible when attribution is not a major issue	• More feasible when attribution is not a major issue • Options for counterfactuals may be limited	• Options for counterfactuals may be limited • Survey may be necessary

where some standard economic methodology might be more feasible. In the bovine and water cases, retrospective analysis might be more feasible for aspects of the research program where attribution is less of an issue. And in the cases with largely nonmarket benefits—water quality and watersheds, and NDL—options for counterfactuals might be limited.

Finally, peer review is a standard part of research evaluation for all ARS research programs, as it is for research in other Federal agencies. The NDL case offered a particular example in which a more intensive peer review process initiated by ARS 12 years ago—the National Food and Nutrient Analysis Program—led to substantial changes and improvements in the NDL research program. Furthermore, one way that economic rigor might be incorporated into the peer review process would be to recruit economists in both prospective and retrospective peer review panels, with a particular focus on the provision of public goods and identification of unique roles for Federal research with poor substitutes.

Conclusions

Analysts from many different disciplines, including economics, have applied a wide variety of tools to the evaluation of scientific research. Nearly all of these applications, including economic approaches, have been directed to either or both of two broad purposes, accountability and planning. Institutions that use public funds for scientific research, including Federal research agencies such as ARS, are responsible for providing evidence that such investments have been socially beneficial. When planning research, governments that fund research, research managers, and researchers attempt to anticipate the directions of future scientific advances. In many cases, funding agencies and research managers also decide how best to use available resources to attain advances that will provide social benefits, and not duplicate research investments that private firms might make.

We reviewed the reasons for economic evaluation of agricultural research (chapters 1 and 2); surveyed available methods for research evaluation, with particular focus on economic methods (chapters 3 and 4); and summarized the protocols by which ARS develops and reviews its research portfolio (chapter 4). We conducted three case studies of ARS research programs (best viewed as National Program components, rather than entire ARS National Programs) to determine how research evaluation, and particularly economic evaluation, might be applied in each case (chapter 5). Our goal was to choose a portfolio of cases that would convey the wide range of research that ARS performs. We now summarize some broad findings from the study.

- **Peer Review**

Peer review is the most widely applied method of Federal research evaluation, and it is widely used throughout the ARS planning and evaluation cycle. Peer review is critically important for monitoring and improving the scientific content of Federal research, but it may not provide complete information concerning market acceptance of technology or ultimate social benefits flowing from Federal research.

Peer review is sometimes called the "gold standard" of research evaluation methods. Peer review's strength is assessing scientific merit, but in most cases it does not address the market impact or ultimate social benefits of scientific research programs. A carefully selected and well-prepared peer review panel can provide either prospective or retrospective analysis on the scientific merit of a particular research program. Furthermore, such a panel might even be able to contribute preliminary ideas about certain impact areas, such as economic impact, even though peer review is not specifically designed to do this.

Assessing the economic and social impact of past research usually requires analysis of a longer time horizon than is undertaken in a typical peer review. Understanding the likelihood of economic and social impact, and where that impact might be felt, could strengthen the peer review process.

• Econometric and Economic Surplus Analysis

Standard tools of economic evaluation such as econometric or economic surplus analysis have usually been applied retrospectively. They have also been most frequently used to analyze economic impacts at more aggregated levels than a single research project.

The two major tools for analyzing the economic impact of agricultural research are econometric methods and economic surplus analysis. These methods have been applied most often retrospectively; in fact, econometric analysis requires historical data. Both econometric and economic surplus analyses have well-known strengths and weaknesses. A number of challenges are relevant regardless of which form of standard analysis is applied. These include the lag between research investments and the realization of research benefits, and accounting for research spillovers and other issues of attribution.

Econometric analysis has usually been applied to more aggregate research investment, like agricultural research in general, crops research, livestock research, or research on a particular commodity. Economic surplus analysis may be applied to research aggregates, but variants of the approach can also be applied to a particular research project. For Federal research evaluation, cost-benefit analysis, which is closely related to economic surplus analysis, has been used primarily for research programs that are near enough to market application that specific technologies deriving from the research program can be identified. Cost-benefit analysis is less applicable to more basic research or to research where many of the benefits are nonmarket goods. Furthermore, high-quality cost-benefit analyses are usually expensive.

• Economic Reasoning in the Absence of Formal Methods

Even when formal economic methods cannot be applied directly to research evaluation, economic reasoning can inform evaluation and analysis of past or potential research performance.

Recognition of important economic variables can be part of research evaluation, whether it is retrospective or prospective. If research is targeting a specific industry sector, the relative economic size of that industry can be identified, as it may be of some importance. Simple scenarios can be constructed to illustrate future impacts with or without the research in question. Similarly, an approximate breakdown of benefit shares to different groups might provide useful information to research managers.

A principal economic justification for public research is market failure in the creation of knowledge. One of the most fundamental contributions of an economic perspective is a more precise characterization of the concept of "market failure." Economic analysis can help to depict the nature of the market failure that public research needs to address. Understanding the nature of the market failure addressed by research can not only provide justification for past investments, but also help target future research investments.

• Distinctive Characteristics of Agriculture

Agriculture has some distinctive characteristics that influence the creation and evaluation of research benefits. The biological, structural, and spatial

characteristics of agricultural production, along with the nonmarket nature of some agriculturally related research, suggest a "one size fits all" evaluation methodology is inappropriate across all public sector agricultural research programs.

The reproductive characteristics of domesticated crop and livestock species, the importance of pests and diseases, and the influence of weather are all factors in determining the role of public sector research in solving agricultural problems. Although changes in technology and intellectual property protection increased opportunities for private sector investment in agricultural research, biological factors impair the ability of private firms to capture all the benefits of the agricultural research they perform. Furthermore, the spatially diverse nature of farming and the predominance of relatively small business entities in production agriculture means that the socially optimal role for the public sector in production-related research is more uncertain than in other industries.

There are also important areas of mutual interest and overlap between agricultural research and other research areas, particularly environmental quality and human health. The nonmarket nature of such research benefits makes measuring them and attributing them more difficult.

- **Public Goods Aspects of Research**

In all three cases we studied, ARS provides research outputs that would not be provided at all, or would be more incomplete or costlier, were they acquired from alternate sources. This suggests that ARS research in these three cases has clear public goods components.

All three ARS case study research programs exist in a complex research environment featuring multiple institutional actors. However, the private sector provides little comparable research in all three case study areas. Furthermore, ARS research often fills a niche that other public sector research institutions do not. Some ARS research also contributes to regulatory and policy functions of the Federal Government. These findings suggest the public goods nature of ARS research. When economists value information or technology related to the ARS program components in these case studies (like the nutrient database), the social benefits appear to be large.

- **Nonmarket Benefits and Attribution Issues**

As expected, measuring research impacts is more difficult when important research benefits are nonmarket, as with water quality or nutrient databank research. In addition, proper attribution of research results is particularly important for economic research evaluation, and attribution issues are considerably more complex than simply identifying potential alternative sources of the research in question.

In both the water resources case and the nutrient databank case, most ultimate benefits—environmental quality and human health, respectively—are nonmarket and thus difficult to measure. Even in the third case, bovine quantitative genetics and genomics, production-oriented research might have environmental impacts. Economists have developed methods for measuring nonmarket environmental and human health goods. However, applying these

measures requires more data, involving more assumptions, than is usually necessary for measuring benefits in commodity markets with observable prices. As a result, data collection and analysis are costlier in nonmarket cases.

Proper attribution of research benefits to research investments and research programs is an important concern in any evaluation of scientific research. "Superattribution" of benefits might occur if many institutions claim credit for the same research impact or if multiple actors are indispensible to ultimate research impacts. However, additional factors besides parallel research complicate the attribution of research benefits. For example, nutrition is an important component of human health, but certainly not the only cause of human health outcomes. Environmental quality can be affected by the application of scientific research results, but environmental policy and farm operator decisions are important determinants of environmental quality. Even when the link between a particular research investment and a particular economic impact is pronounced, there may be multiple means to achieving that impact.

In summary, economic evaluation of agricultural research in general and ARS research in particular requires a diverse set of research activities, data, methods, and analysis. The method of economic evaluation depends on resources both human and financial. Available tools might be applied with greater success in some research areas than in others. In some cases, it might be desirable to apply sophisticated economic methods in more detailed studies. Even when this is not feasible, some aspects of economic reasoning could be applied in more rapid impact appraisal. This might involve, for example, characterizing the target market, analyzing the nature of the market failure addressed by the research, or even making naïve benefits forecasts for priority setting purposes (Mullen, 2008). When using more formal economic tools, and even in rapid impact appraisal, careful targeting of economic evaluation exercises might lead to more useful results than simply relying on targets of opportunity where evaluation appears to be relatively "easy."

From a management and assessment standpoint, economic analysis of ARS research may help assess research priorities and guide prospective R&D investment decisions. Integration of economic evaluation with ARS National Program planning cycles would probably have more impact than assessment of economic benefits as a stand-alone exercise. For example, economic evaluation might help determine whether a National Program has had significant impacts. Economic analysis could also suggest appropriate economic and social measures that could be included in evaluating a National Program, or inform programmatic decisions on how best to revise or redirect the program. Specific project plans could be selected within a program for quantitative, net return, or qualitative benefit analysis. Before the retrospective review process (possibly a year in advance of NP review), components or thematic areas within appropriate NPs—similar to the case studies described in this report—could be selected for economic analysis.

Potential benefits of economic evaluations to ARS administrators and managers include reducing the risk and uncertainty that may exist in both basic and applied research; better balancing the prospective and retrospective review process; determining more easily identifiable accomplishments and

outcomes; and communicating more forcefully to customers, stakeholders, partners, and policymakers (primarily OMB, GAO, and Congress) the value of ARS research. Scientists might also benefit in terms of understanding the importance of an economic evaluation, knowing that the scientific community is competing for finite funds, and realizing that their project can be improved based upon the prospective and retrospective review process. As economic evaluation tools and results of ARS research evolve, so will the ways in which ARS establishes its short- and long-term research agendas, but the mandate will always remain—to use Federal resources in the most effective way for the greatest benefit of American citizens.

References

Abler, D.G., and J.S. Shortle. 1991. "The Political Economy of Water Quality Protection from Agricultural Chemicals." *Northeastern Journal of Agricultural and Resource Economics* 20(1): 53-60.

Abler, D.G., and J.S. Shortle. 1995. "Technology as an Agricultural Pollution Control Strategy." *American Journal of Agricultural Economics* 77(1): 20-32.

Adams, J.D., and R. Clemmons. 2006. *Science and Industry: Tracing the Flow of Basic Research Through Manufacturing and Trade*. NBER Working Paper 12459. Cambridge, MA: National Bureau of Economic Research (NBER).

AIPL (Animal Improvement Programs Laboratory, USDA/ARS). 2007. "History of USDA Dairy Evaluations" and "History of Dairy Evaluations at USDA." http://aipl.arsusda.gov/aipl/history/hist_eval.htm and http://aipl.arsusda.gov/aipl/history/hist_lab.htm

Alston, J.M. 2002. "Spillovers." *Australian Journal of Agricultural and Resource Economics* 46(3): 315-346.

Alston, J.M., C. Chan-Kang, M.C. Marra, P.G. Pardey, and T.J. Wyatt. 2000. *A Meta-Analysis of Rates of Return to Agricultural R&D: Ex Pede Herculem?* IFPRI Research Report 113. Washington, DC: International Food Policy Research Institute (IFPRI).

Alston, J.M., B. Craig, and P.G. Pardey. 1998. *Dynamics in the Creation and Depreciation of Knowledge and the Returns to Research*. EPTD Discussion Paper No. 35. Washington, DC: International Food Policy Research Institute (IFPRI).

Alston, J.M., G.W. Norton, and P.G. Pardey. 1995. *Science Under Scarcity: Principles and Practice for Agricultural Research Evaluation and Priority Setting*. Ithaca, NY: Cornell University Press.

Alston, J.M., and P.G. Pardey. 1996. *Making Science Pay: The Economics of Agricultural R&D Policy*. Washington, DC: The American Enterprise Institute Press.

Alston, J.M., and P.G. Pardey. 2001. "Attribution and Other Problems in Assessing the Returns to Agricultural R&D." *Agricultural Economics* 25(2-3): 141-152.

Alston, J.M., and P.G. Pardey. 2006. "Developing-Country Perspectives on Agricultural R&D: New Pressures for Self-Reliance?" In *Agricultural R&D in the Developing World: Too Little, Too Late?* P.G. Pardey, J.M. Alston, and R.R. Piggott (eds.). Washington, DC: International Food Policy Research Institute (IFPRI).

Arrow, K.J. 1962. "Economic Welfare and the Allocation of Resources for Invention." In *The Rate and Direction of Inventive Activity: Economic and Social Factors*. Princeton, NJ: Princeton University Press.

Audretsch, D.B., A.N. Link, and J.T. Scott. 2002. "Public/Private Technology Partnerships: Evaluating SBIR-Supported Research." *Research Policy* 31(1): 145-158.

Barkley, A.P. 1997. "Kansas Wheat Breeding: An Economic Analysis." Paper presented to the WAEA Annual Meeting, Reno, NV, July 13-16, 1997.

Bateman, I.J., R. Brouwer, H. Davies, B.H. Day, A. Deflandre, S. Di Falco, S. Georgiou, D. Hadley, M. Hutchins, A.P. Jones, D. Kay, G. Leeks, M. Lewis, A.A. Lovett, C. Neal, P. Posen, D. Rigby, and R.K. Turner. 2006. "Analyzing the Agricultural Costs and Non-Market Benefits of Implementing the Water Framework Directive." *Journal of Agricultural Economics* 57(2): 221-237.

Bator, F. 1958. "The Anatomy of Market Failure." *Quarterly Journal of Economics* 72(3): 351-379.

Bickman, L. (ed.). 1987. *Using Program Theory in Evaluation, New Directions in Program Evaluation, No. 33*. San Francisco: Jossey Bass.

Boadway, R.W., and N. Bruce. 1984. *Welfare Economics*. Oxford, UK: Basil Blackwell.

Bodner-Montville, J., J.K.C. Ahuja, L.A. Ingwersen, E.S. Haggerty, C.W. Enns, and B.P. Perloff. "USDA Food and Nutrient Database for Dietary Studies: Released on the Web." *Journal of Food Composition and Analysis* 19 (supplement 1): S100-S107.

Bovine Genome Sequencing and Analysis Consortium, C.G. Elsik, R.L. Tellam, and K.C. Worley. 2009. "The Genome Sequence of Taurine Cattle: A Window to Ruminant Biology and Evolution." *Science* 324 (24 April 2009): 522-528.

Bovine HapMap Consortium. 2009. "Genome-Wide Survey of SNP Variation Uncovers the Genetic Structure of Cattle Breeds." *Science* 324 (24 April 2009): 528-532.

Bozeman, B. 2000. "Technology Transfer and Public Policy: A Review of Research and Theory." *Research Policy* 29(4-5): 627-655.

Bozeman, B. 2002. "Public-value Failure: When Efficient Markets May Not Do." *Public Administration Review* 62(2): 145-161.

Bozeman, B. 2003. *Public Value Mapping of Science Outcomes: Theory and Method.* Center for Science, Policy and Outcomes, Washington, DC. http://www.cspo.org/home/cspoideas/know_flows/Rock-Vol2-1.PDF

Brennan, J.P. 1984. "Measuring the Contribution of New Varieties to Increasing Wheat Yields." *Review of Marketing and Agricultural Economics* 52: 175-195.

Brown, E.A. 1996. *Applying the Principles of the Government Performance and Results Act to the Research and Development Function: A Case Study Submitted to the Office of Management and Budget by the Army Research Laboratory*. Washington, DC: The Congressional Institute.

Bush, V. 1945. *Science: The Endless Frontier*. Washington, DC: U.S. Government Printing Office.

Byerlee, D., and R. Murgai. 2001. "Sense and Sustainability Revisited: The Limits of Total Factor Productivity Measures of Sustainable Agricultural Systems." *Agricultural Economics* 26(3): 227-236.

Byerlee, D., and G. Traxler. 2001. "The Role of Technology Spillovers and Economies of Size in the Efficient Design of Agricultural Research Systems." In *Agricultural Science Policy: Changing Global Agendas*, J.M. Alston, P.G. Pardey, and M.J. Taylor (eds.). Baltimore and London: The Johns Hopkins University Press.

Cahill, P. 2000. "Fast Track: Is It Speeding Commercialization of the Department of Defense Small Business Innovation Research Projects?" In *The Small Business Innovation Research Program: An Assessment of the Department of Defense Fast Track Initiative*, C.W. Wessner (ed.). Washington, DC: National Academy Press.

Capalbo, S.M., and J.M. Antle. 1989. "Incorporating Social Costs in the Returns to Agricultural Research." *American Journal of Agricultural Economics* 71(2): 458-463.

Caswell, M., and K. Day-Rubenstein. 2006. "Agricultural Research and Development." In *Agricultural Resources and Environmental Indicators, 2006 Edition*, K. Wiebe and N. Gollehon (eds.). U.S. Department of Agriculture, Economic Research Service. http://www.ers.usda.gov/publications/arei/eib16/eib16_3-2.pdf

Chavas, J.-P., and T.L. Cox. 1992. "A Nonparametric Analysis of the Influence of Research on Agricultural Productivity." *American Journal of Agricultural Economics* 74(3): 583-591.

Claassen, R. 2003. "Emphasis Shifts in U.S. Agri-Environmental Policy." *Amber Waves* 1(5): 39-44. U.S. Department of Agriculture, Economic Research Service.

Cockburn, I.M., and R.M. Henderson. 2000. "Publicly Funded Science and the Productivity of the Pharmaceutical Industry." In *Innovation Policy and the Economy 1*, A.B. Jaffe, J. Lerner, and S. Stern (eds.). Cambridge, MA: National Bureau of Economic Research and the MIT Press.

Cohen, L.R., and R.G. Noll. 1996. "The Future of the National Laboratories." *Proceedings of the National Academies of Sciences USA*. 93: 12678-12685.

Cozzens, S.E. 1995. *Assessment of Fundamental Science Programs in the Context of the Government Performance and Results Act (GPRA)*. RAND Monograph Report MR-707/0-OSTP. Santa Monica, CA: The RAND Corporation.

Crow, M., and B. Bozeman. 1998. *Limited by Design: R&D Laboratories in the U.S. National Innovation System*. New York: Columbia University Press.

CVision Corporation. 2006. *Benefits and Costs of a Subsurface Agricultural Drainage Water Management System to Improve Water Quality and Increase Crop Production in a Public-Private Partnership*. Available online from the Agricultural Drainage Management Coalition at http://www.admcoalition.com/complete_analysis.pdf.

Dasgupta, P., and J. Stiglitz. 1980. "Uncertainty, Industrial Structure, and the Speed of R&D." *Bell Journal of Economics* 11(1): 1-28.

Day-Rubenstein, K., P.W. Heisey, R. Shoemaker, J. Sullivan, and G. Frisvold. 2005. *Crop Genetic Resources: An Economic Appraisal*. Economic Information Bulletin No. 2. U.S. Department of Agriculture, Economic Research Service.

Dietary Guidelines Advisory Committee. 2004. *Report of the Dietary Guidelines Advisory Committee on the Dietary Guidelines for Americans, 2005*. U.S. Department of Health and Human Services and U.S. Department of Agriculture.

Dolan, P. 2000. "The Measurement of Health-Related Quality of Life for Use in Resource Allocation Decisions in Health Care." In *Handbook of Health Economics, Volume 1*, A.J. Culyer and J.P. Newhouse (eds.). Amsterdam: Elsevier Science Publishers.

Ershow, A.G. 2003. "Research Science, Regulatory Science, and Nutrient Databases: Achieving an Optimal Convergence. *Capstone Lecture*." *Journal of Food Composition and Analysis* 16(3): 255-268.

Evenson, R.E. 1989. "Spillover Benefits of Agricultural Research: Evidence from U.S. Experience." *American Journal of Agricultural Economics* 71(2): 447-452.

Evenson, R.E. 2001. "Economic Impacts of Agricultural Research and Extension." In *Handbook of Agricultural Economics, Vol. 1A: Agricultural Production*, B.L. Gardner and G.C. Rausser (eds.). Amsterdam: Elsevier Science.

Feather, P., D. Hellerstein, and L. Hansen. 1999. *Economic Valuation of Environmental Benefits and the Targeting of Conservation Programs:*

The Case of the CRP. Agricultural Economic Research Report No. 778. U.S. Department of Agriculture, Economic Research Service.

Feyerherm, A.M., G.M. Paulsen, and J.L. Sebaugh. 1984. "Contribution of Genetic Improvement to Recent Wheat Yield Increases in the USA." *Agronomy Journal* 76: 985-990.

Fitzsimmons, C.B. 2001. *Knowledge Spillovers from Joint Government-Industry Supported Research: A Case-Study from the Automotive Industry*. Ph.D. Dissertation. Fairfax, VA: George Mason University.

Fuglie, K., N. Ballenger, K. Day, C. Klotz, M. Ollinger, J. Reilly, U. Vasavada, and J. Yee. 1996. *Agricultural Research and Development: Public and Private Investments Under Alternative Markets and Institutions*. Agriculture Economic Report Number 735. U.S. Department of Agriculture, Economic Research Service.

Fuglie, K.O., and P.W. Heisey. 2007. "Economic Returns to Public Agricultural Research." Economic Brief Number 10. U.S. Department of Agriculture, Economic Research Service.

Fuglie, K.O., J.M. MacDonald, and E. Ball. 2007. "Productivity Growth in U.S. Agriculture." Economic Brief Number 9. U.S. Department of Agriculture, Economic Research Service.

Gal-Or, E. 1986. "Information Transmission—Cournot and Bertrand Equilibria." *Review of Economic Studies* 53(1): 85-92.

Gaughan, M. 2003. *Public Value Mapping Breast Cancer Case Studies.* Center for Science, Policy and Outcomes, Washington, DC. http://www.cspo.org/products/rocky/Rock-Vol2-2.PDF

Gopinath, M., and T.L. Roe. 2000. "R&D Spillovers: Evidence from U.S. Food Processing, Farm Machinery and Agriculture." *Economics of Innovation and New Technology* 9(3): 223-243.

Green, R.D., M.A. Qureshi, J.A. Long, P.J. Burfening, and D.L. Hamernik. 2007. "Identifying the Future Needs for Long-Term USDA Efforts in Agricultural Animal Genomics." *International Journal of Biological Sciences* 3(3): 185-191.

Griliches, Z. 1957. "Hybrid Corn: An Exploration in the Economics of Technological Change." *Econometrica* 25(4): 501-522.

Griliches, Z. 1958. "Research Costs and Social Returns: Hybrid Corn and Related Innovations." *Journal of Political Economy* 66(5): 419-431.

Griliches, Z. 1986. "Productivity, R&D, and Basic Research at the Firm Level in the 1970's." *The American Economic Review* 76(1): 141-154.

Hall, B. 1993. "Industrial Research During the 1980s: Did the Rate of Return Fall?" *Brookings Papers on Economic Activity* 2: 286-343.

Hansen, L., and M. Ribaudo. 2008. *Economic Measures of Soil Conservation Benefits: Regional Values for Policy Assessment*. TB-1922. Washington, D.C.: U.S. Department of Agriculture, Economic Research Service.

Harper, D.A. 2001. *Changing Works: Visions of a Lost Agriculture*. Chicago: The University of Chicago Press.

Heisey, P.W., and M.L. Morris. 2006. "Economic Impact of Water-Limited Conditions on Cereal Grain Production." In *Drought Adaptation in Cereals*, J.-M. Ribaut (ed.). New York: The Haworth Press.

Heisey, P.W., J.L. King, K. Day-Rubenstein, and R. Shoemaker. 2006. *Government Patenting and Technology Transfer*. Economic Research Report Number 15. U.S. Department of Agriculture, Economic Research Service.

Heisey, P.W., C.S. Srinivasan, and C. Thirtle. 2001. *Public Sector Plant Breeding in a Privatizing World*. Agriculture Information Bulletin Number 772. U.S. Department of Agriculture, Economic Research Service.

Hicks, D., P. Kroll, F. Narin, P. Thomas, R. Ruegg, H. Tomizawa, Y. Saitoh, and S. Kobayashi. 2002. *Quantitative Methods of Research Evaluation Used by the U.S. Federal Government*. NISTEP Study Material No. 86. Tokyo, Japan: National Institute of Science and Technology Policy (NISTEP), Ministry of Education, Culture, Sports, Science and Technology.

Hounshell, D. 1984. *From the American System to Mass Production, 1800-1932*. Baltimore, MD: The Johns Hopkins University Press.

Huffman, W.E., and R.E. Evenson. 2006. *Science for Agriculture: A Long-Term Perspective* (2nd Edition). Ames, IA: Blackwell Publishing.

Hurley, J. 2000. "An Overview of the Normative Economics of the Health Sector." In *Handbook of Health Economics, Volume 1*, A.J. Culyer and J.P. Newhouse (eds.). Amsterdam: Elsevier Science Publishers.

Just, R.E., D.L. Hueth, and A. Schmitz. 1982. *Applied Welfare Analysis and Public Policy*. Englewood Cliffs, NJ: Prentice-Hall.

Kappes, S.M., J.W. Keele, R.T. Stone, R.A. McGraw, T.S. Sonstegard, T.P. Smith, N.L. Lopez-Corrales, and C.W. Beattie. 1997. "A Second-Generation Linkage Map of the Bovine Genome." *Genome Research* 7(3): 235-249.

Klotz, C., K. Fuglie, and C. Pray. 1995. "Private-Sector Agricultural Research Expenditures in the United States, 1960-92." Staff Paper No. AGES9525. U.S. Department of Agriculture, Economic Research Service.

Kostoff, R.N. 1995. "Federal Research Impact Assessment: Axioms, Approaches, Applications." *Scientometrics* 34(2): 163-206.

Lambert, D., G. Schaible, R. Johansson, and U. Vasavada. 2007. "The Value of Integrated CEAP-ARMS Survey Data in Conservation Program Analysis." *Journal of Soil and Water Conservation* 62(1): 1-10.

Larsen, J., and R. Wigand. 1987. "Industry-University Technology Transfer in Microelectronics." *Policy Studies Review* 6(3): 584-595.

Lichtenberg, F.R., and D. Siegel. 1991. "The Impact of R&D Investment on Productivity—New Evidence Using Linked R&D-LRD Data." *Economic Inquiry* 29(2): 203-229.

Link, A.N. 1995. "Evaluating Program Performance: The Case of Federally-funded Collaborative Research." *International Journal of Technology Management* 19(7/8): 847-852.

Link, A.N., and J.T. Scott. 2004. "Evaluating Public Sector R&D Programs: The Advanced Technology Program's Investment in Wavelength References for Optical Fiber Communications." *The Journal of Technology Transfer* 30(1-2): 241-251.

Link, A.N., and G. Tassey. 1987. *Strategies for Technology-Based Competition: Meeting the New Global Challenge*. Lexington, MA: D.C. Heath.

Logsdon, J.M., and C.B. Rubin. 1988. "Research Evaluation Activities of Ten Federal Agencies." *Evaluation and Program Planning* 11(1): 1-11.

Loury, G.C. 1979. "Market Structure and Innovation." *Quarterly Journal of Economics* 93(3): 395-410.

Lubowski, R.N., M. Vesterby, S. Bucholtz, A. Baez, and M.J. Roberts. 2006. *Major Uses of Land in the United States, 2002*. Economic Information Bulletin Number 14. U.S. Department of Agriculture, Economic Research Service.

Lynam, J.K., and R.W. Herdt. 1989. "Sense and Sustainability: Sustainability as an Objective in International Agricultural Research." *Agricultural Economics* 3(4): 381-398.

Macauley, M.K. 2006. *Ascribing Societal Benefit to Environmental Observations of the Earth from Space: The Multi-angle Imaging Spectroradiometer (MISR)*. RFF Discussion Paper 06-09. Washington, DC: Resources for the Future.

Mansfield, E. 1980. "Basic Research and Productivity Increase in Manufacturing." *The American Economic Review* 70(5): 863-873.

Mansfield, E. 1991. "Academic Research and Industrial Innovation." *Research Policy* 20(1): 1-12.

Mansfield, E. 1998. "Academic Research and Industrial Innovation: An Update of Empirical Findings." *Research Policy* 26(7-8): 773-776.

Maredia, M.K., and D. Byerlee. 2000. "Efficiency of Research Investments in the Presence of International Spillovers: Wheat Research in Developing Countries." *Agricultural Economics* 22(1): 1-16.

Marglin, S.A. 1963. "The Opportunity Costs of Public Investment." *Quarterly Journal of Economics* 77(2): 274-289.

Martin, B.R., and J. Irvine. 1983. "Assessing Basic Research: Some Partial Indicators of Scientific Progress in Radio Astronomy." *Research Policy* 12(2): 61-90.

McCunn, A., and W.E. Huffman. 2000. "Convergence in U.S. Productivity Growth for Agriculture: Implications of Interstate Research Spillovers for Funding Agricultural Research." *American Journal of Agricultural Economics* 82(2): 370-388.

Mendelsohn, R. 1981. "The Choice of Discount Rates for Public Projects." *American Economic Review* 71(1): 239-241.

Morris, M.L., J. Rusike, and M. Smale. 1998. "Maize Seed Industries: A Conceptual Framework." In *Maize Seed Industries in Developing Countries*, M.L. Morris (ed.). Boulder, CO: Lynne Rienner Publishers.

Mullen, J.D. (ed.). 2004. *Evaluations in 2003 of Five Areas of Investment in R&D by NSW Agriculture: Summary*. Economic Research Report No. 22. Orange, New South Wales, Australia: NSW Department of Primary Industries.

Mullen, J.D. 2008. "Evaluation of Agricultural Research at the Sub-Sector Level in Australia." Presentation to the ERS/USDA Workshop *Assessing the Benefits of Agricultural Research Service (ARS) R&D Within an Economic Framework*, Washington, DC, March 10.

Murphy, K.M., and R.H. Topel. 2007. "Social Value and the Speed of Innovation." *The American Economic Review* 97(2): 433-437.

Narrod, C.A., and K.O. Fuglie. 2000. "Private Investment in Livestock Breeding with Implications for Public Research Policy." *Agribusiness: An International Journal* 16(4): 457-470.

NAS/COSEPUP (National Academy of Sciences/Committee on Science, Engineering, and Public Policy). 1999. *Evaluating Federal Research Programs: Research and the Government Performance and Results Act*. Washington, DC: National Academy Press.

Nelson, R.R. 1959. "The Simple Economics of Basic Scientific Research." *Journal of Political Economy* 67(3): 297-306.

Newbery, D.M.G., and J.E. Stiglitz. 1981. *The Theory of Commodity Price Stabilization: A Study in the Economics of Risk*. Oxford, UK: Clarendon Press.

Nielson, J., and J. Brazzel. 1980. *Evaluation as an Aid to Decision-Making in the Food and Agricultural Sciences*. Joint Planning and Evaluation Staff Paper No. 80-DD-02, Science and Education Administration, U.S. Department of Agriculture.

NRC (National Research Council) Committee on Science, Engineering, and Public Policy. 2008. *Evaluating Research Efficiency in the U.S. Environmental Protection Agency*. Washington, DC: National Academies Press.

NRC (National Research Council) Board on Agriculture and Natural Resources. 2003. *Frontiers in Agricultural Research: Food, Health, Environment, and Communities*. Washington, DC: National Academies Press.

NSF (National Science Foundation), Division of Science Resources Statistics. 2007. *Federal R&D Funding by Budget Function: Fiscal Years 2006-2008*. NSF 07-332. Arlington, VA.

OECD (Organization for Economic Cooperation and Development). 2006. *Workshop on Science of Science Policy: Developing our Understanding of Public Investments in Science*. Draft findings. Helsinki, Finland: July 12.

Oehmke, J.F., and D.E. Schimmelpfennig. 2004. "Quantifying Structural Change in U.S. Agriculture: The Case of Research and Productivity." *Journal of Productivity Analysis* 21(3): 297-315.

OMB (Office of Management and Budget, Executive Office of the President). 2006. *Assessment of USDA In-House Research for Nutrition and Health.* http://www.whitehouse.gov/omb/expectmore/detail/10003019.2006.html

OMB (Office of Management and Budget, Executive Office of the President). 2007. *Guide to the Program Assessment Rating Tool (PART)*. http://www.whitehouse.gov/omb/part/fy2007/2007_guidance_final.pdf

OTA (Office of Technology Assessment, U.S. Congress). 1986. *Research Funding as an Investment: Can We Measure the Returns? A Technical Memorandum*. OTA-TMSET-36.

Pannell, D.J. 1999. "On the Balance Between Strategic-Basic and Applied Agricultural Research." *Australian Journal of Agricultural and Resource Economics* 43(1): 91-113.

Pardey, P.G., J.M. Alston, C. Chan-Kang, E.C. Magalhães, and S.A. Vosti. 2006. "International and Institutional R&D Spillovers: Attribution of Benefits Among Sources for Brazil's New Crop Varieties." *American Journal of Agricultural Economics* 88(1): 104-123.

Pardey, P.G., J.M. Alston, J. Christian, and S. Fan. 1996. *Hidden Harvest: U.S. Benefits from International Research Aid*. Food Policy Report. Washington, DC: IFPRI (International Food Policy Research Institute).

Patton, M.Q. 1990. *Qualitative Evaluation and Research Methods, Second Edition*. Newbury Park, CA: SAGE Publications.

Pehrsson, P.R., D.B. Haytowitz, J.M. Holden, C.R. Perry, and D.G. Beckler. 2000. "USDA's National Food and Nutrient Analysis Program: Food Sampling." *Journal of Food Composition and Analysis* 13(4): 379-389.

Pehrsson, P.R., D.B. Haytowitz, and J.M. Holden. 2003. "USDA's National Food and Nutrient Analysis Program: Update 2002." *Journal of Food Composition and Analysis* 16(3): 331-341.

Peretz, J.H., S. Das, and B.E. Tonn. 2005. "Evaluating the Short-Run Benefits of Six Department of Energy R&D Projects." *The Journal of Technology Transfer* 30(3): 287-301.

Popp, D. 2006. "R&D Subsidies and Climate Policy: Is There a 'Free Lunch?'" *Journal of Climate Change* 77(3-4): 311-341.

Romer, Paul M. 1994. "The Origins of Endogenous Growth." *Journal of Economic Perspectives* 8(1): 3-22.

Rosenberg, N. 1990. "Why Do Firms Do Basic Research (With Their Own Money)?" *Research Policy* 19(2): 165-174.

Ruegg, R., and I. Feller. 2003. *A Toolkit for Evaluating Public R&D Investment: Models, Methods and Findings from ATP's First Decade*. NIST GCR 03-857. Gaithersburg, MD: National Institute of Standards and Technology.

Ruegg, R., and G. Jordan. 2007. *Overview of Evaluation Methods for R&D Programs: A Directory of Evaluation Methods Relevant to Technology Development Programs*. Prepared for U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy.

Salasin, J., L. Hattery, and T. Ramsay. 1980. *The Evaluation of Federal Research Programs*. NTIS Order Number PB 81234106. McLean, VA: MITRE Corporation.

Salter, A.J., and B.R. Martin. 2001. "The Economic Benefits of Publicly Funded Basic Research: A Critical Review." *Research Policy* 30(3): 509-532.

Samuelson, P.A. 1954. "The Pure Theory of Public Expenditure." *Review of Economics and Statistics* 36(4): 387-389.

Sarewitz, D., G. Foladori, N. Invernizzi, and M.S. Garfinkel. 2004. "Science Policy in its Social Context." *Philosophy Today* 48 (2004 Supplement): 67-83.

Sarewitz, D., and R.A. Pielke, Jr. 2007. "The Neglected Heart of Science Policy: Reconciling Supply of and Demand for Science." *Environmental Science and Policy* 10: 5-16.

Shapiro, C. 2007. *Patent Reform: Aligning Reward and Contribution*. NBER Working Paper 13141. Cambridge, MA: National Bureau of Economic Research (NBER).

Sobel, D. 1995. *Longitude: The True Story of a Lone Genius Who Solved the Greatest Scientific Problem of His Time*. New York: Walker and Company.

Stokes, D.E. 1997. *Pasteur's Quadrant: Basic Science and Technological Innovation*. Washington, DC: Brookings Institution Press.

Strauss, A., and J. Corbin. 1998. *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory* (Second Edition). Thousand Oaks, CA: SAGE Publications.

Tassey, G. 1997. *The Economics of R&D Policy*. Westport, CT: Quorum Books.

Tassey, G. 2003. *Methods for Assessing the Economic Impacts of Government R&D*. NIST Planning Report 03-1. Gaithersburg, MD: National Institute of Standards and Technology (NIST), U.S. Department of Commerce.

U.S. Census Bureau, Department of Commerce. 2009. *2007 Economic Census*. http://www.census.gov/econ/census07/.

U.S. Census Bureau, Department of Commerce. 2004. *2002 Economic Census*. http://www.census.gov/econ/census02/index.html.

U.S. Department of Agriculture, Agricultural Research Service, Nutrient Data Laboratory. n.d. *National Food and Nutrient Analysis Program: Technical Proposal*. Beltsville, MD: Human Nutrition Research Center.

U.S. Department of Agriculture; Cooperative State Research, Education, and Extension Service. 2008. Current Research Information System. http://cris.csrees.usda.gov/fsummaries.html.

U.S. Department of Agriculture, National Agricultural Statistics Service. 2004. *2002 Census of Agriculture*. http://www.agcensus.usda.gov/Publications/2002/index.asp.

U.S. Department of Agriculture, National Agricultural Statistics Service. 2009. *2007 Census of Agriculture*. http://www.agcensus.usda.gov/Publications/2007/index.asp.

U.S. Department of Agriculture, National Institute of Food and Agriculture. 2010. *Current Research Information System*. http://cris.csrees.usda.gov/fsummaries.html.

U.S. Department of Agriculture, National Agricultural Statistics Service. Various years. *Agricultural Statistics*.

U.S. Department of Agriculture, Natural Resources Conservation Servic). 2007. *2003 Annual National Resource Inventory*.

Van Tassell, C.P., T.S. Sonstegard, M.S. Ashwell, and E.E. Connor. 2001. "Quantitative Genetics and Genomics: Lessons from Dairy Cattle." Paper Presented at the Biotechnology-Aquaculture Interface: The Site of Maximum Impact Workshop, USDA/ARS and The Oceanic Institute, Shepherdstown, WV, March 5-7, 2001.

Variyam, J.N., and J. Cawley. 2006. *Nutrition Labels and Obesity*. NBER Working Paper No. W11956. Cambridge, MA: National Bureau of Economic Research (NBER).

Vives, X. 1984. "Duopoly Information Equilibrium: Cournot and Bertrand." *Journal of Economic Theory* 34(1): 71-94.

Walker, T.S., and K.O. Fuglie. 1999. "The Economics of Generating International Public Goods from Investing in Potato Plant Breeding." In *Impact on a Changing World: CIP Program Report 1997-1998*. Lima, Peru: International Potato Center (CIP).

Walter, T., M. Dosskey, M. Khanna, J. Miller, M.D. Tomer, and J. Wiens. 2007. "The Science of Targeting Within Landscapes and Watersheds to Improve Conservation Effectiveness." In *Managing Landscaping for Environmental Quality: Strengthening the Science Base*, M. Schnepf and C. Cox (eds.). Ankeny, IA: Soil and Water Conservation Society, pp. 63-89.

Weller, J.I. 1994. *Economic Aspects of Animal Breeding*. London: Chapman & Hall.

Weller, J.I., Y. Kashi, and M. Soller. 1990. "Power of Daughter and Granddaughter Designs for Determining Linkage Between Marker Loci and Quantitative Trait Loci in Dairy Cattle." *Journal of Dairy Science* 73(9): 2525-2537.

Wigand, R., and G. Frankwick. 1989. "Interorganizational Communication and Technology Transfer: Industry-Government-University Linkages." *International Journal of Technology Management* 4(1): 63-76.

Wright, B.D. 1983. "The Economics of Invention Incentives: Patents, Prizes, and Research Contracts." *American Economic Review* 73(4): 691-707.

Xia, Y., and S. Buccola. 2005. "University Life Science Programs and Agricultural Biotechnology." *American Journal of Agricultural Economics* 87(1): 229-243.

Zentner, R.P., and W.L. Peterson. 1984. "An Economic Evaluation of Public Wheat Research and Extension Expenditures in Canada." *Canadian Journal of Agricultural Economics* 32: 327-353.

Appendix: Challenges of and Potential Extensions to Standard Economic Methods

Given the data-intensive and time-consuming nature of both econometric and economic surplus analyses, a whole host of factors may contribute to the reliability of their results. In either approach, results are sensitive to data availability and model specification. Many important decisions have to be made when an economic analysis is undertaken. What is the appropriate scale and scope of coverage? How can the analysis be linked to reducing market failure or risk and uncertainty, since it is dependent on measuring observable impacts in markets? Should the analysis be *ex ante*, *ex post*, or, as in some cases, a hybrid that looks both at historical and prospective research benefits and costs? What is the appropriate counterfactual (without research) scenario, especially if the research interacts with policies aiming at some of the same economic and social objectives? How robust is the analysis to measurement and forecast errors?

Estimating the importance of *lags* in research benefits illustrates the importance of model specification. The expected benefits of research can only occur after research results have been obtained and implemented. This delay means that the economic value of research benefits should exceed their costs in order to justify the alternative use of funds for the time leading up to the first research benefits. But delays also create challenges for identifying research impacts. Analysts may find it relatively straightforward to observe and model the time pattern of research costs and benefits for well-defined economic surplus studies of particular technological changes or research programs. On the other hand, the lag structure of more aggregated econometric studies in which research costs enter as an explanatory variable requires additional judgment on the part of the analyst.

Evenson (2001) discusses potential patterns for flows of research benefits, and in many cases a particular model (e.g., a trapezoidal pattern) for research lags is simply assumed. Alston and Pardey (2001) summarize the results of a number of studies of returns to research by the different assumed lag structures. They argue that the problem of estimating research lag structure is one example of attribution problems in agricultural research. Other analysts (Oehmke and Schimmelpfennig, 2004; Alston et al., 1998) argue that assumed lags do not have sufficient grounding in economic theory, but statistical methods might be used to help determine the relevant lags in a particular analytical situation. Unfortunately, statistical methods to help determine lag structure usually have low statistical power and require very long data time series.

Knowledge spillovers are another persistent feature in the economic evaluation of research. Spillovers arise when productivity is altered by indirect participation, as when scientists draw on previously established research. *Geographic spillovers* occur when an activity is specific to a certain location or climate but nonetheless has an effect in a different geographic zone. Even though much agricultural research is location-specific, Evenson (1989) and Byerlee and Traxler (2001) have analyzed geographic spillovers in agricultural research. Interstate or international R&D spillovers could account for half or more of measured aggregate productivity growth in agriculture,

so accounting for geographical spillovers is clearly an important, if not well understood, issue (Alston, 2002).

Geographical spillovers are yet another example of the general problem of *attribution* in research (Alston and Pardey, 2001). More general forms of knowledge spillovers may come from the basic sciences to applied agricultural research. In any particular area of applied research, different public institutions, as well as private firms, may contribute to specific new technologies for agriculture. Plant breeding is an example of research for which attribution is an important issue. How much of the benefits of new crop varieties should be attributed to the plant breeders who developed the final varieties, how much to other plant breeders who improved the source germplasm, and how much to farmers who may have contributed the original genetic resources (Day-Rubenstein et al., 2005)? In some cases, attribution rules relating to plant breeding techniques can be formulated (Pardey et al., 1996), but these are choices of the analyst, and not based on particular economic or genetic criteria. For example, Pardey and colleagues (2006), in analyzing R&D for upland rice, soybeans, and edible beans in Brazil, contrasted attribution of all benefits from crop improvement for these crops to Embrapa, Brazil's primary public agricultural research agency, with attribution of benefits to other sources of crop improvement (Brazilian state-level research agencies, private firms, other domestic germplasm sources, and foreign research contributors and collaborators).

Potential Economic Extensions

The economic approaches to research evaluation outlined above have focused on market impacts. However, economics does indicate several ways that these approaches might be extended.

First, we have argued that looking at market impacts might be accomplished more readily for applied research than for basic research, despite the difficulties in defining and distinguishing these concepts. It is possible that market failure in the production of basic research is particularly likely because of high fixed costs, low reproduction costs, and limited appropriability. If market failure diminishes incentives for private basic research, is it possible to measure the economic impact of basic research in the public sector?

Many economists start with the presumption that measuring the benefits from basic research poses particular difficulties (Rosenberg, 1990; Alston et al., 1995). Salter and Martin (2001) note three analytical approaches to analyzing the benefits of publicly funded basic research: econometric studies, surveys, and case studies. Salter and Martin sometimes muddy the distinction between "publicly funded" research and "basic" research, however, and in fact *econometric studies* of returns to basic research (Mansfield, 1980; Griliches, 1986; Lichtenberg and Siegel, 1991) attempt to link basic research expenditures to productivity at the level of the private firm or industry, rather than focusing on public research per se. Although some studies of this nature distinguish between internal private research funding and Federal funding of private firm research, they do not address public basic research in general. Most find strong positive returns to basic research.[1]

[1] As Griliches (1986) notes, the direction of causation is not always clear. Firms with higher rates of productivity growth might have more resources to invest in basic research.

Surveys or other studies of knowledge flows attempt to measure the number of innovations that depend on academic, presumed basic research (Mansfield, 1991, 1998) or to trace the use of academic research in industry through citation analysis (Xia and Buccola, 2005; Adams and Clemmons, 2006). In general, these studies find that basic research, or its proxies, do have a significant impact on technological change, but they tend not to quantify the economic component of this impact. Finally, *case studies* attempt "to examine directly the innovation process and the historical roots of a particular technology" (Salter and Martin, 2001). Although some studies (Xia and Buccola, 2005) do look at impacts of basic research for particular technologies, many analyze the impacts of basic research in general, even when the unit of observation may be the individual firm, as it is in many econometric studies of basic research.

Despite the difficulties of measuring basic research and its impacts, economists have begun to propose models to help determine justifiable amounts of investment in basic research. Murphy and Topel (2007) propose a model for determining the balance of basic and more applied research in medicine and health, but do not have the data to quantify the parameters of their model. Pannell (1999) similarly uses sensitivity analysis in his model of basic and applied agricultural research, once again since the "true" model parameters are unknown. Pannell argues that within a broad range of possible allocations, the balance between basic and applied research may be near optimal as long as the research portfolio remains diversified.

Besides the question of basic versus applied research, standard economic analysis has focused particularly on impacts of research in defined markets. Increasingly, however, it is likely that calls will be made to incorporate valuation of nonmarket impacts of agricultural R&D into economic analysis (Capalbo and Antle, 1989; Lynam and Herdt, 1989; Byerlee and Murgai, 2001). Past studies have kept from addressing these issues because of measurement difficulties and the fact that research impacts directly observable in commodity and factor markets were abundantly available. Research systems, however, are called upon more and more to provide positive nonmarket environmental or health benefits as well as mission-driven benefits, and so valuation of these benefits will become more important to impact analysis.

Agricultural research may provide positive benefits in environmental, health, or other noncommodity areas through a variety of means. Some may come through direct technology development, such as research to improve plant nutrient management on farms, to absorb fertilizer runoff before it reaches water resources, or to improve the nutritional content of crop varieties. Other research may develop tools, such as better methods of predicting soil loss or nutrient runoff, that can be used to establish counterfactuals, direct policy, and guide allocation of program resources, all of which can improve environmental management. Social science research may be complementary to applied scientific research by demonstrating which combinations of technology and policy are more likely to achieve desired environmental or health impacts.

A few studies—for example, studies on the environmental effects of aquaculture or the health effects of pesticide use—have incorporated environmental

or health effects directly into economic estimates of costs and benefits. Most economic analyses do not, however, attempt to incorporate these costs or benefits. Instead, other methods are used to assess nonmarket impacts of agricultural R&D. Even though a wide variety of agricultural research is likely to have environmental impacts, Alston and colleagues (2000) found that economic studies of natural resources research have been primarily confined to forestry or fisheries research—in other words commodity-like research—and not to the effects of more broadly defined environmental research.

A number of economic tools are, in fact, available to measure nonmarket environmental benefits and costs. Feather et al. (1999) summarize the various approaches as follows:

- Averting or defensive expenditures—the measurement of expenditures "made by individuals to reduce or negate pollution damages."

- Changes in production costs—"observing changes in firm profits, input costs, or output prices due to changes in environmental quality."

- Revealed preference—inferring the demand for environmental quality by observing individual behavior. Typically, recreational trips are used to measure the demand for environmental quality.

All of these approaches estimate direct-use value of environmental benefits or costs. A fourth measurement approach estimates total value, that is both direct-use value and other values such as the value of preserving the environment for future generations or the value of maintaining a resource so that it can be used at some future date. This approach is:

- Stated preference—directly asking individuals either their willingness to pay for changes in environmental quality or asking them to order various scenarios with different prices and different amounts of environmental quality.

In principle, these methods could be used to evaluate the environmental benefits and costs of agricultural R&D.[2] Obviously, they add considerable measurement complexity to the analysis and bring up a host of methodological concerns.

Similarly, measuring nonmarket health benefits in an economic framework can be undertaken by applying concepts such as "quality adjusted life years" or "healthy year equivalents." However, these measures are often obtained only through the fourth method used to measure environmental benefits and costs—the stated preference method. This has led many economists studying human health to concentrate primarily on cost-effectiveness measures (for example, ranking different treatments for the same disease by dollars per life year saved), rather than the direct measurement of benefits and costs (Hurley, 2000; Dolan, 2000). Thus, direct measurement of health benefits and costs from agricultural R&D may be even more difficult than direct measurement of environmental benefits and costs.[3]

Attribution issues also complicate the incorporation of nonmarket environmental/health benefits into economic analysis. A large number of factors, ranging well beyond agricultural R&D, influence environmental, health,

[2] See Bateman et al. (2006) for an example of the application of various measures to the nonmarket benefits of water quality policy in the European Union.

[3] These measurement difficulties may be one reason why studies of returns to health R&D are considerably less common than studies of returns to agricultural R&D, even though investments in health research worldwide are very large components of scientific research.

and other nonmarket outcomes. In developed countries, application of these nonmarket measures has traditionally been to evaluate the impact of environmental policy, not the environmental impacts of agricultural R&D.[4] In some cases, it might be best to view policy as the prime instrument for influencing environmental quality, with agricultural R&D potentially providing scientific results that are complementary to policy. Thus, economic benefits or costs that do occur might be rightly attributed primarily to policy, with the value of agricultural R&D resting on the degree to which it provides information that increases the benefit of policy decisions and outcomes.

Economic experiments are another possible extension that can assist research evaluation. In experimental economics, test subjects are provided information and asked to make decisions in a simulated market or other economic setting. Typically, real payments are made to subjects who achieve higher simulated economic gains. By varying aspects of the simulation, these experiments can provide insight into economic decisionmaking. Economic experiments have been used to elicit valuation of nonmarket goods and to describe the value of information in decisionmaking. As economic experimental instruments are developed and refined, they could provide information that is helpful for assessing research on public goods, creation of new data sources, and research in mission areas.

Despite these many measurement and attribution difficulties, incorporation of nonmarket impacts into economic analysis to justify and demonstrate the public role may become increasingly important. Private sector research investments have recently outpaced investments by the public sector (Caswell and Day-Rubenstein, 2006). As research systems mature, the private sector might continue to focus on research with direct application to commodities or factors of agricultural production. As a result, agricultural research conducted by the public sector may be called upon increasingly to demonstrate its public goods characteristics. This may be one reason agricultural research is now often expected to have a broader range of positive impacts beyond simple increases in agricultural production, including mitigation of past negative environmental or health impacts, or provision of new, positive benefits in these or other areas.

Understanding the Social Impacts of Public R&D

Social outcomes that are more broadly defined than economic outcomes may be important reasons for public research funding. Furthermore, more attention to the societal context of research may be crucial in understanding the multitude of actors and actions that lead to particular social outcomes of research. Science is only one of the institutions determining social outcomes and often not the most important one. Other participants use the knowledge produced by scientists but also shape it in different ways through their response to this knowledge, and thereby can influence the context and method for producing scientific knowledge.

Bozeman (2003) addresses these issues with a discussion of public value mapping (PVM). A critical aspect of PVM is that it avoids a focus on market value as a means to assess the social value of public R&D activities. Instead, it attempts to discern broad public values and goals and measures impact relative to these. Therefore, the focus is on stated policy goals or bench-

[4]For examples of studies that consider the joint economic evaluation of policy, technology, and other factors that influence environmental quality, see Abler and Shortle (1991, 1995).

marks, including those contained within budgets, strategic plans, and documents related to the Government Performance and Results Act (GPRA) and Program Assessment Rating Tool (PART) (Bozeman, 2003). This defines success as intrinsic progress toward scientific achievement, an alternative to direct valuation methods (see box, "Public Value Mapping and Breast Cancer Research in the United States").

Within the PVM framework, Bozeman posits the notion of a knowledge value collective (KVC) that "…is a set of individuals connected by their production and uses of a body of scientific and technical information." The "collective" aspect of the KVC incorporates the concept of spillovers and joint production of research. The KVC includes scientists and their research teams, but also includes those who use the knowledge created to produce other information, technologies, or products (i.e., first-order users). KVC does not include "second-order users" or those who consume the final products from the process.

Using the case of the U.S. Department of Agriculture's Nutrient Data Bank to illustrate, first-order users within the KVC might be other scientific research agencies that use the data in their studies of the interactions between nutrition and health. First-order users might also be those intermediary firms that obtain and repackage the data to sell to researchers or other interested parties, or nongovernmental organizations that use the data to inform their

Public Value Mapping and Breast Cancer Research in the United States

Gaughan (2003) applied some of the concepts of public value mapping (PVM) to breast cancer research in the United States. She did this with a pair of case studies, one of federally sponsored breast cancer research, and one of the nascent Georgia Cancer Coalition. First, these studies included an analysis of stated legislative and executive objectives for scientific achievement, particularly with respect to expected outcomes. This included not only consideration of the language of legislation, but also examination of administrative documents such as strategic plans. Second, Gaughan did a thorough inventory, both current and historical, of the organizational mechanisms developed or modified to implement research objectives. Third, she presented relevant data, for example, on breast cancer incidence and mortality over time, broken down by population characteristics such as race, as well as data on proportions of women who had received mammograms within the previous 2 years. Fourth, she identified the knowledge value collective (KVC) essential to the success of the larger social objectives, including not only the scientific institutions but also the business community, the nonprofit community, and the intended beneficiaries.

Data presented showed that despite research efforts, trends in age-adjusted breast cancer incidence and mortality were either upward or flat between 1973 and 1996. She argued that the Federal breast cancer research effort had focused on biomedical investigation, to the detriment of "social, behavioral, and population-based studies to examine how to avoid and limit cancer in the first place." It was too early to reach conclusions about the Georgia Cancer Coalition (GCC), but based on the Federal experience, the GCC aimed to create greater linkages between academic researchers and clinicians and the appropriate clinical populations in the State of Georgia.

education outreach and lobbying activities. Second-order users outside the KVC might be consumers who read the nutrient data on food packaging to make better informed dietary choices.

In addition to *ex post* evaluations of research, Federal agencies are often concerned with the potential value of planned research. A model of the social impact of public R&D amenable to *ex ante* evaluation is to view science policy decisions as made within a supply and demand framework. That is, public expenditures on science are justified because of the potential benefits derived from the knowledge and information produced. It is often argued that since the benefits derived from scientific inquiry are often unexpected or serendipitous, the most efficient method of supplying this knowledge and information is through the autonomous research programs of scientists. The Organization for Economic Cooperation and Development's (OECD) Global Science Forum (OECD, 2006) refers to this as a "top down" approach in which expert groups are convened to systematically identify a consensus set of research priorities, based on explicit policy goals (for example, national defense or environmental protection), in contrast to a "bottom up" approach in which scientific priorities respond to citizen preferences.

Sarewitz and Pielke (2007) argue that a solely "top down" approach is unrealistic for several reasons, including:

- There are limited funds available for investment in public R&D.

- Since World War II, major scientific and technological changes have come about from public R&D investments in particular areas.

- Social groups and movements as well as industrial interests have had well-documented influence on public research agendas and even outcomes. Examples of the former include increases in resources for HIV/AIDS efforts.

Therefore, the ideal role for science policy is to link potential users of scientific data and information with producers in such a way as to facilitate the most socially optimal allocation of research funds to the most useful research portfolios. Sarewitz and Pielke conclude that "the challenge [for science policy scholarship] is to identify particular cases where the promises upon which scientific funding is predicated are not being met. The challenge for research policy-makers is to use the findings to allocate limited resources."