

This paper was presented at a colloquium entitled “Science, Technology, and the Economy,” organized by Ariel Pakes and Kenneth L. Sokoloff, held October 20–22, 1995, at the National Academy of Sciences in Irvine, CA.

Evaluating the federal role in financing health-related research

ALAN M. GARBERT^{†‡§} AND PAUL M. ROMER^{§¶||}

[†]Veterans Affairs Palo Alto Health Care System, Palo Alto, CA 94304; [‡]Stanford University School of Medicine, [§]National Bureau of Economic Research, and [¶]Graduate School of Business, Stanford University, Stanford, CA 94305; and ^{||}Canadian Institute for Advanced Research, Toronto, ON Canada MST 1X4

ABSTRACT This paper considers the appropriate role for government in the support of scientific and technological progress in health care; the information the federal government needs to make well-informed decisions about its role; and the ways that federal policy toward research and development should respond to scientific advances, technology trends, and changes in the political and social environment. The principal justification for government support of research rests upon economic characteristics that lead private markets to provide inappropriate levels of research support or to supply inappropriate quantities of the products that result from research. The federal government has two basic tools for dealing with these problems: direct subsidies for research and strengthened property rights that can increase the revenues that companies receive for the products that result from research. In the coming years, the delivery system for health care will continue to undergo dramatic changes, new research opportunities will emerge at a rapid pace, and the pressure to limit discretionary federal spending will intensify. These forces make it increasingly important to improve the measurement of the costs and benefits of research and to recognize the tradeoffs among alternative policies for promoting innovation in health care.

In this paper, we address three general questions. What role should the federal government play in supporting scientific and technological progress in health care? What information should the federal government collect to make well-informed decisions about its role? How should federal policy toward research and development respond to scientific advances, technology trends, and changes in the political and social environments?

To address these questions, we adopt a societal perspective, considering the costs and benefits of research funding to American society as a whole. Both in government and in the private sector, narrower perspectives usually predominate. For example, a federal agency may consider only the direct costs that it bears. A device manufacturer may weigh only the direct costs and benefits for the firm. Both organizations will thereby ignore costs and benefits that accrue to members of the public. The societal perspective takes account of all costs and benefits. Although alternative perspectives are appropriate in some circumstances, the comprehensiveness of the societal perspective makes it the usual point of departure for discussions of government policy. Much of our discussion focuses on decisions that are made by the National Institutes of Health (NIH), the largest federal agency devoted to biomedical research, but our comments also apply to other federal agencies sponsoring scientific research.

The approach we adopt is that of neoclassical, “Paretian” welfare economics (1). This approach dictates that potential

changes in policy should be evaluated by comparing the total costs and benefits to society. It suggests that only those policies whose benefits exceed their costs should be adopted. When they are accompanied by an appropriate system of transfers, these policies can improve the welfare of everyone. As is typical in cost–benefit analysis (CBA), we focus on total costs and benefits and do not address the more detailed questions about how gains should be distributed among members of the public. By adopting this approach to measuring policy-making, we simplify the analysis and can draw upon a well-developed intellectual tradition (2).

We start by outlining a theoretical framework for organizing the discussion of these issues. The usual analysis of government policy toward science and technology marries the notion of market failure—the failure of markets to satisfy the conditions necessary for economic efficiency—and the notion of a rate of return to research. These concepts have helped to structure thinking about these issues, but they are too limiting for our purposes. We propose a broader framework that compares the benefits from more rapid technological change with the costs associated with two possible mechanisms for financing it: expanded property rights (which creates monopoly power) and tax-financed subsidies. Expanded property rights could take the form of longer patent life or more broadly defined patent and copyright protection for intellectual property. Tax-financed subsidies could take the form of government-funded (extramural) research, government-performed research (e.g., intramural research at NIH), government subsidies to private research, and government-subsidized training. Optimal policy, we claim, uses a mix of expanded property rights and subsidies. Thus, policymakers must address two distinct questions. Is the total level of support for research and development adequate? Is the balance between subsidies and monopoly power appropriate?

These questions arise in any setting in which innovation is a concern. After we define the fundamental concepts used in discussions of technology policy, we show that the choice between monopoly power and subsidies arises within a private firm just as it does at the level of the nation. After describing this analytical framework, we then ask how it can be used to guide government policy decisions. Specifically, what kinds of data would policymakers need to collect to make informed decisions about both questions? Such data would enable a government agency engaged in research funding to set and justify overall spending levels and to set spending priorities across different areas of its budget. It would also be able to advise other branches of government about issues such as patent policy that can have far-reaching implications for the health care sector.

Theoretical Framework

Market Failure and Public Goods. The central theme of microeconomic analysis is the economic efficiency of the

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: NIH, National Institutes of Health; CBA, cost–benefit analysis.

idealized competitive market. There are many forms of market failure—departures from this ideal. Two of the most important are monopolistic control of specific goods and incomplete property rights. Many discussions treat research as a public good and presume that the underlying market failure is one of incomplete property rights. This suggests that if we could make the protection for intellectual property rights strong enough, we could return to the competitive ideal. In fact, a true public good is one that presents policymakers with an unavoidable choice between monopoly distortions and incomplete property rights.

There are two elements in the definition of a public good. It must be nonrival, meaning that one individual's consumption of the good does not diminish the quantity available for others to use. It must also be nonexcludable. Once it is produced, anyone can enjoy the benefits it offers, without getting the consent of the producer of the good (3).

Incomplete excludability causes the kind of market failure we expect to observe when property rights are not well specified. When a rival good such as a common pasture is not excludable, it is overused and underprovided. Society suffers from a "tragedy of the commons." The direct way to restore the conditions needed for an efficient outcome is to establish property rights and let a price system operate. For example, it is possible to divide up the commons, giving different people ownership of specific plots of land. The owners can then charge grazing fees for the use of the land. When there are so many landholders that no one person has a monopoly on land, these grazing fees give the owners of livestock the right incentives to conserve on the use of the commons. They also give landowners the right incentives to clear land and create new pasture. When it is prohibitively expensive to establish property rights and a price system, as in the case of fish in the sea, the government can use licensing and quotas to limit overuse. It can also address the problem of underprovision by directly providing the good, for example by operating hatcheries.

For our purposes, the key observation is that these unmitigated benefits from property rights are available for rival goods. Nonrival goods pose a distinct and more complicated set of economic problems that are not widely appreciated. Part of the difficulty arises from the obscurity of the concept of rivalry itself. The term rival means that two persons must vie for the use of a particular good such as a fish or plot of land. A defining characteristic of research is that it produces nonrival goods—bits of information that can be copied at zero cost. It was costly to discover the basic information about the structure of DNA, but once that knowledge had been uncovered, unlimited numbers of copies of it could be made and distributed to biomedical researchers all over the world. By definition, it is impossible to overuse a nonrival good. There is no waste when every laboratory in the world can make use of knowledge about the structure of DNA. There is no tragedy in the intellectual commons. For a detailed discussion of nonrivalry and its implications for technology development, see Romer (4).

Some of the most important science and technology policy questions turn on the interaction of excludability and rivalry. As noted above, for a rival good like a pasture, increased excludability, induced by stronger property rights, leads to greater economic efficiency. Stronger property rights induce higher prices, and higher prices solve both the problem of overuse and the problem of underprovision. However, for a nonrival good, stronger property rights may not move the economy in the right direction. When there are no property rights, the price for a good is zero. This leads to the appropriate utilization of an existing nonrival good but offers no incentives for the discovery or production of new nonrival goods. Higher prices ameliorate underprovision of the good (raising the quantity supplied) but exacerbate its underutilization (diminishing the quantity demanded). If scientists had to pay a royalty fee to *Waston and Crick* for each use that they made of the

knowledge about the structure of DNA, less biomedical research would be done.

The policy challenge posed by nonrival goods is therefore much more difficult than the one posed by rival goods. Because property rights support an efficient market in rival goods, the "theory of the first best" can guide policy with regard to such goods. The first best policy is to strive to establish or mimic as closely as possible an efficient market. For nonrival goods, in contrast, policy must be guided by the less specific "theory of the second best." For these goods, it is impossible, even in principle, to approach an efficient market outcome. A second best policy, as the name suggests, is an inevitable but uneasy compromise between conflicting imperatives.

The conceptual distinction between rivalry and excludability is fundamental to any discussion of policy. Rivalry is an intrinsic feature of a good, but excludability is determined to an important extent by policy decisions. Under our legal system, a mathematical formula is a type of nonrival good that is intentionally made into a public good by making it nonexcludable. Someone who discovers such a formula cannot receive patent or copyright protection for the discovery. A software application is another nonrival good, but because copyright protection renders it excludable, it is not a public good. It is correct but not very helpful to observe that the government should provide public goods. It does not resolve the difficult question of which nonrival goods it should make into public goods by denying property rights over these goods.

Beyond "the Market Versus the Government." In many discussions, the decision about whether a good should be made into a public good is posed as a choice between the market and the government. A more useful way to frame the discussion is to start by asking when a pure price system (which may create monopoly power) is a better institutional arrangement than a pure tax and subsidy system, and vice versa.

By a pure price system we mean a system in which property rights are permanent and owners freely set prices on their goods. Under such a system, a firm that developed a novel chemical with medicinal uses could secure the exclusive rights to sell the chemical forever.

A pure tax and subsidy system represents a polar opposite. Under this system, the good produced is not excludable, so a producer cannot set prices or control how their output is used. Production would be financed by the subsidy. Produced goods are available to everyone for free. To clarify the policy issues that arise in the choice between these two systems, our initial discussion will be cast entirely in terms of a firm making internal decisions about investment in research, avoiding any reference to the public sector.

Financing Innovation Within the Firm. Picture a large conglomerate with many divisions. Each division makes a different type of product and operates as an independent profit center. It pays for its inputs, charges for its outputs, and earns its own profits. Senior managers, who are compensated partly on the basis of the profits their division earns, have an incentive to work hard and make their division perform well.

To make the discussion specific, imagine that many of the products made by different divisions are computer controlled. Suppose also that some divisions within the firm make software and others manufacture paper products, such as envelopes. Both the software goods and the paper products may be sold to other divisions. The interesting question for our purposes is how senior managers price these internal sales.

Producing Paper Products. For rival goods like envelopes, the "invisible hand" theorem applies to an internal market within the firm just as it would to an external market: a pure price system with strong property rights leads to efficient outcomes. An efficient firm will tell the managers of the envelope division that they are free to charge other divisions whatever price they want for these envelopes. Provided that the other divisions are free to choose between buying internally or

buying from an outside seller, this arrangement tends to maximize the profits of the firm. It is efficient for the internal division to make envelopes if it can produce them at a lower cost than an outside vendor. If not, the price system will force them to stop. If senior management did not give the division property rights over the envelopes and allowed all other divisions to requisition unlimited envelopes without paying, envelopes would be wasted on a massive scale. The firm would suffer from an internal version of the tragedy of the commons.

Producing Software with a Price System. Now contrast the analysis of envelopes with an analysis of software. Almost all of the cost of producing software is up-front cost. When a version of the computer code already exists, the cost of an additional copy of the software is nearly zero. It is nearly a pure nonrival good.

Suppose that one division has developed a new piece of software that diagnoses hardware malfunctions better than any previous product. This software would be useful for all of the divisions that make computer-controlled products. Senior managers could give property rights over the software to the division that produced it, letting it set the price it charges other divisions for the use of the software. Then the producer might set a high price. Other divisions, however, will avoid using this software if the price is so high that it depresses their own profits. They might purchase a less expensive and less powerful set of software diagnostic tools from an outside vendor. Both of these outcomes lead to reductions in the conglomerate's overall profits. They are examples of what economists term monopoly price distortions—underuse induced by prices that are higher than the cost of producing an additional unit.

It would cost the shareholders of the conglomerate nothing if this software were made freely available to all of the divisions, and profits decrease if some divisions forgo the use of the program and therefore fail to diagnose hardware malfunctions properly or if they pay outside suppliers for competing versions of diagnostic software.

Producing Software Under a Tax and Subsidy System. Because software is a nonrival good, the best arrangement for allocating an existing piece of software is to deny the division that produced it internal property rights over it. This avoids monopoly price distortions. Senior management could simply announce that any other division in the conglomerate may use the software without charge. But this kind of arrangement for distributing software gives each division little incentive to produce software that is useful to other divisions within the firm. It solves the underutilization problem but exacerbates the underprovision problem.

Senior management, foreseeing this difficulty, might therefore establish a system of taxes and subsidies. They could tax the profits of each division, using the proceeds to subsidize an operating division that develops new software for internal use. They could even set up a separate research and development division funded entirely from subsidies provided by headquarters. This division's discoveries would be given to the operating divisions for free. Despite the statist connotation associated with the concepts of taxes and subsidies, the managers and owners of a private firm may adopt them because they increase efficiency and lead to higher profits.

These arguments show that, in principle, taxes, subsidies, and weak property rights can be an efficient arrangement for organizing the production and distribution of goods like software. However, subsidies have implicit costs. Managers must ensure that the software produced under the terms of the subsidy actually meets an important need within the other divisions of the firm. To supervise a subsidized operation, they must estimate the value of its output in the absence of any price signals or arms-length transactions that reveal information about willingness to pay. Operating divisions will accept any piece of software that is offered for free, so the fact that a subsidized software group seems to have a market for its goods

within the firm reveals almost nothing. This division might write software that is worth far less to the conglomerate than its cost of production. Thus, a subsidy system poses its own risk to the profitability of the firm. Avoiding these risks imposes serious measurement and supervisory costs on senior management, costs they need not incur when a division produces a rival good and runs as a profit center. To supervise the envelope division, senior managers only need to know whether it earns a profit.

The taxes that headquarters imposes on the operating divisions also create distortions. If the workers in a division keep only a fraction of the benefits that result from their efforts, they will not work as hard as they should to save costs and raise productive efficiency. Taxes weaken incentives. If it is too difficult for senior management to supervise the activities of software workers who receive subsidies, the distortion in incentives resulting from a system of taxes and subsidies may be more harmful than distortions resulting from operating the software division as a monopolistic profit center.

The problem for this firm is a problem for any economic entity. For rival goods like envelopes, a price system offers a simple, efficient mechanism for making the right decisions about production and distribution. For nonrival goods like software, there is no simple, efficient system. Both price systems and tax and subsidy systems can induce large inefficiencies. In any specific context, making the right second-best choice between these pure systems or some intermediate mixture requires detailed information about the relative magnitudes of the associated efficiency costs.

Financing Innovation for the Nation as a Whole. At the level of the nation, just as at the level of the firm, relative costs drive choices between price systems and tax subsidy systems. The major cost of the price system is monopoly price distortion, which occurs when a good is sold at a price that exceeds marginal cost.

Fig. 1 illustrates monopoly price distortion. The downward-sloping demand curve shows how the total quantity purchased varies with the price charged. The demand curve can also be interpreted as a schedule of the willingness to pay for an additional unit of the good as a function of the total number of units that have already been purchased. As the number already sold increases, the willingness to pay for one more unit falls. The figure also charts the marginal cost of producing additional units of output, assumed here to be constant, as well as the price p^* and quantity q^* purchased when a monopolist is free to set prices to maximize profits.

The triangle marked "Deadweight loss" represents the dollar value of the welfare loss that results from setting price above marginal cost: some people are willing to pay more than

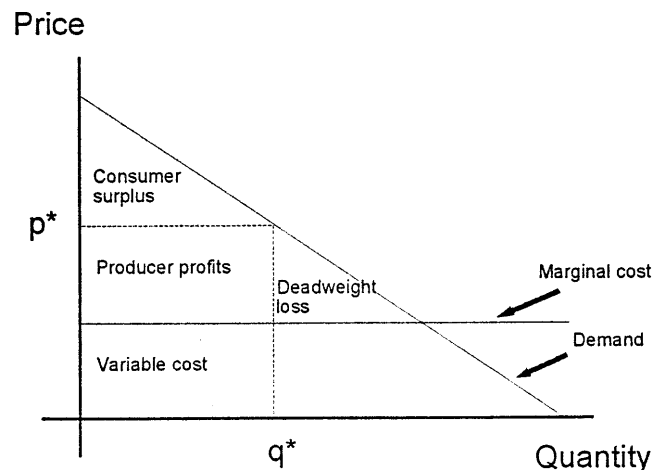


FIG. 1. Monopoly price distortion.

the cost of producing one more unit but less than the monopoly price. The resulting underconsumption can be overcome by reducing the monopolist's price to the level of marginal cost. Expiration of a patent, by eliminating monopoly after a fixed time, eventually solves this problem.

Monopoly pricing can cause another problem. The total value to society of the good depicted here is the total area under the demand curve less the cost to society of the units that are produced. In the figure, the rectangle below the marginal cost line marked "Variable cost" represents the production costs for q^* units of the good. The total value to society is the sum of the willingnesses to pay of all people who purchase, or the area under the demand curve up to the quantity q^* . The net value to society is the difference between these two areas, which is equal to the rectangle marked "Producer profits" plus the triangle marked "Consumer surplus." This rectangle of profits is the difference between the revenue from sales and the variable cost of the goods produced. The surplus is a pure gain captured by those consumers who pay less than the goods are worth to them. Firms compare the profit rectangle to the fixed research and development cost of introducing the good when they evaluate a new product. They neglect the consumer surplus that the new good will generate for purchasers. Thus even under conditions of strong property rights and high monopoly prices, there will be a tendency for the market to underprovide valuable new goods.

When policymakers weigh the use of property rights and monopoly power to finance the introduction of new goods, they must consider two other aspects of monopoly pricing that change the size of the total distortions it creates. On the one hand, price discrimination—the strategy of charging different customers different prices—can mitigate or eliminate the efficiency losses due to monopoly. On the other hand, the efficiency losses from monopoly power become worse when one monopolist sells to another.

A surprising implication of economic theory is that a perfectly price-discriminating monopolist (i.e., one that charges each consumer his exact willingness to pay) produces the efficient (i.e., perfectly competitive) quantity of output. By charging each consumer the exact amount that he would be willing to pay, the monopolist continues to produce up to the point where the value to the last consumer is equal to the marginal cost of an additional unit of output. Thus, price discrimination mitigates or completely solves the problem of underuse. In addition, it helps solve the problem of underprovision because it increases the total profit that a supplier of a new good can capture. Price discrimination is widely used in air travel (airlines usually charge more for the changeable tickets likely to be used by business travelers) and telephone services (which throughout the world charge businesses more than individuals). Price discrimination also occurs in physician and hospital services and in pharmaceutical and laboratory supply sales. Recent legal challenges to the use of price discrimination by pharmaceutical companies in their sales to managed care organizations may unfortunately have limited the use of this promising strategy for minimizing the losses from monopoly pricing.

Monopoly distortions can become larger when production of a good involves a chain of monopolists. For example, suppose that one monopolist invents a new laboratory technique, and a second develops a new drug whose production uses this technique. When two or more monopolists trade in this kind of vertical chain, the welfare losses do not just add up, they multiply. The problems of underuse and failure to develop the good both become worse than they would be if a single monopolist invented the technique, developed products from it, and priced the goods to the final consumers. This is the justification that economists typically offer for vertical integration of an upstream and a downstream firm into a single firm. However, in an area that is research intensive and subject

to uncertainty, and where there are many possible users of any innovation, vertical integration is often unfeasible. Chiron, which held a monopoly in the use of a critical enzyme for PCR, would have been unable to identify, much less integrate into a single firm, all of the possible firms that could use PCR before it made its decisions about developing this technique. On these grounds, theory suggests that a single monopolist in a final-product market will induce smaller social losses than a monopolist that will play a crucial supplier role to other firms, which are themselves monopolists in downstream markets.

Taxes and Subsidies Cause a Different Set of Distortions. As we have noted, the polar alternative to a pure price system is an allocation mechanism that relies on subsidies to finance innovation. The funds required for a system of government subsidies can be raised only by taxation, which harms incentives. For example, raising the income tax diminishes the incentives to work. In addition, subsidies replace a market test with a nonmarket system that rewards a different set of activities. If these activities are not useful or productive, the subsidies themselves induce distortions and waste. To describe the costs associated with a subsidy system, recall the case of the subsidized software-producing division of the conglomerate. It is costly to design and operate an administrative system that tries to identify useful activities. Failures in such a system also impose costs. Suppose that many of the projects that are subsidized produce no value; suppose further that a system which relied on a market test of value produced fewer such failures. Then funds allocated to the additional wasteful projects must be counted as part of the cost of operating the subsidy system. Under conditions of uncertainty, any allocation system will produce some failures; in this example, we assume that a subsidy system would produce more of them.

Peer review of university-based research grants is widely regarded as an unusually efficient and effective mechanism of subsidy allocation. Its effectiveness derives partly from the details of its structure, such as the anonymous reviews by panels of disinterested experts. However, it also benefits from the limited problem that it is trying to solve. Research review groups make decisions at a high level of abstraction; they do not need to forecast the precise consequences of pursuing a line of research, and do not need much information about the "market demand" for the good they are ultimately subsidizing. Consider, for example, the information necessary to make a good decision about subsidizing different research proposals for research on computer-human interfaces. Then contrast this with the substantially greater amount of information that would be necessary for selecting among several proposals to develop new software applications that will be sold to the public. The information needed to make decisions about final products is extensive, including detailed information about characteristics and their value in the myriad ways that consumers might put them to use. Surely a market test by people who spend their own money is the most efficient mechanism for selecting products in this setting. (To push this point to an extreme, imagine what recreational reading would be like if the government did not offer copyright protection of books, so that the only people who could make a living as authors were people who received grants awarded on the basis of relevancy by university professors!)

Debate about technology policy programs often turns on disagreements about the cost of setting up and operating a system for allocating subsidies. Views in this area are often polarized, but there is little disagreement that it is much harder to establish effective subsidies for narrowly defined final products from an industry than it is to subsidize flexible inputs for that industry and let a market test determine how they are allocated to produce the final product mix. Arguably, the most important contribution the federal government made to the development of the biotechnology industry was to promote training for people who went to work in molecular biology and

related fields. Similarly, government subsidies for training in computer science, which provided the software industry with a pool of talented developers and entrepreneurs, have probably been more effective mechanisms for promoting the development of the software industry than government attempts to promote specific computer languages. The one possible exception to this rule arises when the government is an important user of the good in question, as for example in the case of military equipment. In this case, users within the government have much information about the relevant market demand and can be more successful at selecting specific products to subsidize.

Measuring the Gains from Research and the Costs of Financing It. To make informed decisions about research support, and to strike an appropriate balance between expanded property rights and subsidies, policymakers need quantitative information that will enable them to answer three questions. (i) What are the benefits of an additional investment in research? (ii) What are the costs of financing research through a system of property rights that depends on monopoly profits as the principal incentive? (iii) What are the costs of financing research through a system of taxes and subsidies? We discuss each of these questions, then address some of the pitfalls that may arise in making decisions based on incomplete or misleading information.

Measuring Benefits. The problem of measuring the benefits from research expenditures can be readily posed in terms of the demand curve of Fig. 1. The full benefit to society from research leading to a new discovery is represented by the area under the demand curve up to the quantity of goods sold. If we subtract the variable costs of producing the units sold, we have a measure of benefits that can be compared with the research costs needed to generate this benefit. There are then two ways to proceed. Policymakers can use an estimate of profits to firms as a crude underestimate of the total gains to society. Alternatively, they can try to measure these gains directly by looking at the benefits enjoyed by users of the goods.

Profits as a Proxy for Social Benefits. To keep the discussion simple, assume that a firm made a fixed investment in research sometime in the past. Each year, it earns revenue on sales of the product produced from this research and pays the variable costs of goods produced. The difference, the annual accounting profits of the firm, appears as the profit rectangle in Fig. 1. These profits change over time. The value of the innovation will change as substitute goods are developed, prices for other goods rise or fall, and knowledge about the innovation grows. Accounting profits of firms can thus be used as a lower-bound estimate of the welfare gains from innovation.

In practice, there are several obvious problems with this approach. First, by ignoring consumer surplus, this measure underestimates the benefits of a good. Second, it may be impossible for a government agency (unlike the manufacturer) to estimate the revenues attributable to a single product. Third, until a product has run the course of its useful life, its entire revenue stream will be highly uncertain. At an early stage in the life of a new product, such as a patented drug, the stock market valuation of the company may be taken as an indication of the best estimate of the present value of all the revenue streams held by the firm, and changes in stock market valuation when a new product is approved may give some indication of the present value of the anticipated revenue stream from the good. But if the possibility of approval is anticipated by the stock market, the change in stock market value at the time of approval will be an underestimate of the full value of this revenue stream.

Finally, market transactions will not give an accurate indication of willingness to pay if demand for a good is subsidized. Traditional fee-for-service medical insurance acts as such a subsidy (5). Then patients bear only a fraction of the cost directly, and consume drugs and health services whose value

falls short of the true social cost. In this situation, the monopolist's profits from the sale of the innovation overstate the magnitude of the benefits to society of a newly invented medical treatment.

Cost-Benefit Approach. A more complete picture of the benefits to society can be painted using cost-benefit measures of the total value to consumers of a new good. Consider the value of the discovery that aspirin prevents myocardial infarction (6). What is the information worth? To answer this question, one begins by considering the size of population that would benefit from the therapy, followed by the change in the expected pattern of morbidity and mortality attributable to adoption, and finally the dollar valuation of both the survival and quality-of-life effects. This would represent the potential return and could be calculated on an annual basis, but the potential return would likely overestimate the actual surplus. Some people in the group at risk, for example, might have been taking aspirin before the information from the studies became available. Furthermore, not everyone who could potentially benefit would comply with treatment. Thus, it is necessary to estimate the increment in the number of people using the therapy rather than the potential number of individuals taking it. In addition, there would likely be reductions in expenditures for the treatment of heart attacks, which, after all, would be averted by use of the therapy. Essentially, the estimate of the surplus would be based on a CBA, perhaps conducted for the representative candidates for treatment, multiplied by the number of people who undergo treatment as a direct consequence of the information provided by the clinical trial.

Although the techniques of CBA have been adopted in many areas of public policy, most "economic" analyses of health care and health practices have eschewed CBA for the related technique of cost-effectiveness analysis, which, unlike CBA, does not attempt to value health outcomes in dollar terms (7). Instead, outcomes are evaluated as units of health (typically life expectancy or quality-adjusted life years). The lack of a dollar measure of value of output means that cost-effectiveness analysis does not provide a direct measure of consumer surplus. However, if the cost-effectiveness analysis is conducted properly, it is often possible to convert the information from a cost-effectiveness analysis into a CBA with additional information about the value of the unit change in health outcomes. For example, suppose that the value of an additional year of life expectancy is deemed to be \$100,000, and that a patient with severe three-vessel coronary artery disease treated with bypass surgery can expect to live two years longer at a cost (in excess of medical management) of \$45,000. The cost-effectiveness ratio of surgery, or the increment in costs (\$45,000) divided by the increment in health effects (two years), is \$22,500. The net benefit of surgery is the dollar value of the increased life expectancy (\$200,000) less the incremental cost (\$45,000), or \$155,000. Calculations like these are a central feature of the field of medical technology assessment (8).

Usually the information needed to construct exact measures of the value of medical research will not be available, but crude calculations can be illuminating. Moreover, basic investments in information collection, for example, surveys of representative panels of potential consumers, might greatly improve the accuracy of these estimates. Simple calculations like these, together with more systematic data on health outcomes for the population at large, are among the prerequisites for better decision-making by the government.

Measuring the Cost of Using the Price System and Monopoly Profits. Benefit measures comprise only part of the information needed for good decision-making. Suppose, for example, that policymakers anticipate a large benefit from research directed toward the prevention of a specific disease. They must also decide whether this research should be subsidized by the government or financed by granting monopoly power to private sector firms.

The theoretical discussion in the last section has already identified some of the factors that can influence the social cost of using monopoly power to motivate private research efforts. Monopoly will be more costly if there are many firms with some monopoly power that sell to each other in a vertical chain. In principle, this second problem might be serious for an industry that is research-based, particularly if current trends toward granting patents on many kinds of basic and applied scientific knowledge continue. For example, a drug may be produced by the application of a sequence of patented fundamental processes that results in production of a reagent. The reagent may then be combined with other chemicals to produce a drug. If access to the process is sold by a monopoly, the reagent is sold by another monopoly, and the drug is sold by a third monopoly, the distortion due to monopoly will be compounded. Strengthened property rights can mean, in the limit, that an arbitrarily large number of people or firms with patent rights over various pieces of knowledge will each have veto power over any subsequent developments. If one firm had control of all these processes and carried out all these functions, the price distortion for the final product would be smaller, but as we have indicated, in a research-intensive field characterized by much uncertainty and a large number of small start-up firms, this arrangement may not be feasible.

Yet as we have also indicated, monopoly will be less costly to society as a whole if firms can take advantage of price discrimination. Because the cost of more reliance on monopoly makes this issue so important to a research-intensive field such as pharmaceuticals and because so little is known about the net effect of these two conflicting forces, we believe that it would be valuable to collect more information about the magnitude of monopoly distortions in fields closely related to health care.

Below, we describe feasible mechanisms that could be used to collect more of this kind of information. There are real challenges to collecting the information, because much of it—such as the prices that hospitals and health care networks pay for drugs—is a trade secret.

Monopoly distortions are not the only costs incurred when the private sector finances research; the cost of establishing and maintaining property rights may be substantial. Enforcement of property rights is inexpensive for most physical objects, such as cars or houses. But for nonrival goods that can readily be copied and used surreptitiously, it is much more costly to extend property rights, and more subtle mechanisms may be needed to do so. Initially, software publishers relied on copy protection schemes to prevent revenue losses from unauthorized copying. Over time, they have developed less intrusive techniques (such as restricting technical assistance to registered customers) that achieve the same end.

Sometimes the costs of enforcing property rights are so high that a system based on private incentives will be infeasible. These cases will therefore have high priority for scarce taxpayer-financed government subsidies. Suppose that a private firm decided to sponsor a trial of aspirin to prevent colon cancer and sought the permission of the Food and Drug Administration (FDA) to have exclusive rights to market the use of aspirin for this purpose. Although the company might establish effectiveness and obtain exclusive rights from the FDA, the availability of aspirin from many producers and without a prescription, along with the large number of indications for its use, would make it nearly impossible to enforce market exclusivity for this indication. In such an extreme case, measuring the costs of enforcement is unnecessary, but there often will be instances in which such estimates will be needed because enforcement of property rights is worthwhile but costly. Moreover, as the software example suggests, there is much room for experimenting with different systems to protect property rights.

Cost of Using Taxes and Subsidies. Most of the field of public finance is concerned with quantifying the losses and

gains that occur with government activity, such as taxation. Every form of taxation alters behavior by distorting economic incentives; for example, taxes on bequests reduce the desired size of a bequest, reduce national savings, and increase transfers of wealth during life. Income taxation modifies the relative attractiveness of time devoted to leisure and time devoted to paid work. Traditional calculations of the benefits of government programs in health care, however, ignore the “deadweight” losses due to the behavioral distortions induced by taxation. These losses can be substantial, although their exact magnitude depends on the form of the tax and the economic activity to which it applies. According to the recent estimates, the 1993 personal tax rate increases raised the deadweight loss by about two dollars for every additional dollar of tax revenue (9). These are part of the costs of a tax and subsidy system.

A government subsidy system, like that of a large firm, can generate extensive administrative costs. It can also cause large quantities of resources to be wasted on poorly selected projects. A government agency that dispenses research dollars must devote substantial time and effort to choosing among several competing projects. The market directly produces a mechanism (albeit a Darwinian mechanism that may not be costless) to sort among competing uses of resources. Little is known about the costs of a system to administer subsidies. However, as the previous discussion suggested, qualitative evidence suggests that subsidy systems work better when they make general investments in outputs that are flexible and have many uses. They are less suitable for specific, inflexible investments that require extensive, context-specific information about benefits and willingness to pay.

Making Decisions with Incomplete Information. Because many of the pieces of information that we have outlined above are not available or are available only in the form of qualitative judgments made by experts, it is tempting to substitute surrogate measures for which the information is available. For example, one might simply give up any hope of making judgments about the magnitudes of costs and benefits of various strategies for supporting advances in health care. The NIH might simply try to produce the best biomedical science possible and assume that everything else will follow. But as Rosenberg has noted (10), a country’s success in producing Nobel Prizes in scientific fields is inversely correlated with its economic performance!

More seriously, other seemingly reliable measures could significantly bias government decisions. For example, because profits are observable and salient in political debates, a government agency that subsidizes research may want to maximize profits earned by firms that draw on their research. For NIH, this might mean adopting a strategy that maximizes the profits earned by biotechnology and pharmaceutical firms in the United States. Because a substantial portion of the demand for medical care is still subsidized by a system of fee-for-service insurance, this strategy could lead to large social losses. Even under the paradigm of managed care, which removes or decreases the implicit subsidy for medical care services, profits can be a poor guide to policy. The highest payoff to government spending on research may come from funding research in areas where it is prohibitively expensive to establish the system of property rights that makes private profit possible.

A prominent example of this phenomenon, mentioned above, is the discovery that aspirin can prevent heart attacks and death from heart attacks (11). It is difficult to conceive of realistic circumstances in which a producer of aspirin could gain exclusive rights to sell aspirin for this indication, and it is unlikely that the discovery that aspirin had such beneficial effects markedly increased the profits of its producers. Moreover, since aspirin is produced by many firms, no one of them had much to gain by financing this kind of research. But if the increased profit in this case was small, the consumer’s surplus

may have been extremely large. As our previous discussion noted, it is precisely in circumstances under which a producer cannot recoup the fixed costs of investment in developing a technology that government research may have its greatest payoffs. (In this context, the research established aspirin's beneficial effects on heart disease rather than proving the safety and efficacy of the drug more generally.) In such circumstances, it is imperative to go beyond profits and measure consumer's surplus, but the usual market-based proxies may provide very little information about any such benefits.

Public good features also lent a strong presumption that it was appropriate for the government to sponsor research on the value of beta-blockers after myocardial infarction (12, 13). In the influential NIH-sponsored Beta-Blocker Heart Attack Trial, propranolol reduced mortality by about 25%. The excludability and property rights problems characteristic of aspirin would seem to have been less important for propranolol, but the combination of looming patent expiration and the availability of a growing number of close substitutes diminished the incentives for a private company to sponsor such a trial. Any increase in demand for beta-blockers resulting from the research would likely have applied to the class generally. Although strengthened property rights (such as lengthened market exclusivity) might have made it possible for a private company to capture more of the demand increase resulting from such research, problems with enforcement are similar to those of aspirin: it would be difficult to ensure that other beta-blockers would not be prescribed for the same indication, diluting the return to the manufacturer of propranolol.

The drug alglucerase, for the treatment of Gaucher disease, has characteristics almost opposite to those of aspirin and beta-blockers. Gaucher disease is caused by deficient activity of the enzyme glucocerebrosidase, and NIH-sponsored research led to discovery of the enzyme defect and the development of alglucerase, a modified form of the naturally occurring enzyme.

Subsequently, a private corporation (Genzyme) developed alglucerase further and received exclusive rights to market the compound under the provisions of the Orphan Drug Act. Thus in this instance, both a tax subsidy and a strong property rights approach facilitated the development of the drug.

The high price of alglucerase attracted substantial attention, particularly because most of the drug's development had been sponsored or conducted by the government. The standard dosage regimen devised by the NIH cost well over \$300,000 per year for an adult patient, and therapy is lifelong. According to the manufacturer, the marginal cost of producing the drug accounted for more than half the price, a ratio that is unusually high for a pharmaceutical product (14). Although drug-sparing regimens that appear to be as effective have since been tested, the least expensive of these cost tens of thousands of dollars annually (15). The supplier was able to charge high prices because there is no effective substitute for the drug. This meant that nearly all insurers and managed care organizations covered the drug at any price the manufacturer demanded. Insurance coverage meant that demand would not fall significantly with increases in price, so that monopoly would not cause as much underutilization as would be typical if demand were highly price-responsive. With the insurance subsidy, there would be overconsumption, and expenditures on the drug could exceed the value of benefits it provided.

At current prices, alglucerase is unlikely to be cost-effective compared with many widely accepted health care interventions. An exploration of the federal role in the development of alglucerase revealed the hurdles to be overcome in obtaining the information needed to guide public decision-making—it was possible to obtain rough estimates of the private company's research and development investment but not the investment made by the federal government. Nevertheless, precise information about the costs of research are often, as in this

case, unnecessary to make qualitative decisions about the appropriateness of the taxation and subsidy approach (14).

More detailed information about the relative costs of public and private support for various forms of research can be valuable for many reasons. It may overturn long-standing presumptions about the best kind of research for the government to support. The traditional view is that Nobel Prize-winning science is the area where government support is most important. However, as the case of PCR demonstrates, it is now clear that it is possible to offer property rights that can generate very large profits to a firm that makes a Nobel Prize-winning discovery. It may not be as costly to set up a system of property rights for basic scientific discoveries as many people have presumed. If so, we must still verify whether the costs of relying on monopoly distortions for this kind of discovery are particularly high. At present, we have little basis for making this judgment.

In an era when research budgets are stagnant or shrinking, circumstances will force this kind of judgment. Much population-based research, including epidemiological research and social science research, could provide valuable information (providing insights in such areas as etiologic factors in human disease, biological adaptations to aging, and understanding of the economic consequences of disease and its treatment). This information could inform both public policy and individual planning. All such information is nonrival, and much of it may be inherently nonexcludable because it would be so costly to establish a system of property rights. It is precisely in the areas of research that produce knowledge which is not embodied in a specific product that the benefits from federal investment are likely to be greatest, but most difficult to measure. With a fixed budget, a decision to fund work that could be financed in the private sector—such as sequencing the human genome—means that competing proposals for population-based or epidemiological research cannot be funded. The choice between these kinds of alternatives should be based on an assessment of the best available evidence on all the benefits and costs.

Using Experimentation to Inform Decisions About Research Financing. In many studies of clinical interventions, it is feasible to construct rough estimates of the social returns to government investment in research. As we noted above, it is considerably more difficult to estimate the costs of different systems for financing research. Does this mean that no measurement is possible and that debates about financing mechanisms should be driven by tradition, belief, and politics rather than by evidence? Undoubtedly, measurement can be improved by devoting more resources to it and engaging more intensively in standard activities to measure proxies for research productivity (citation analysis, tracing the relationship of products to research findings, and so on). Even if such activities result in credible estimates of the benefits of research, they tend not to address the principal policy issue: what mix of private and public financing is best? To answer this question, consideration should be given to the collection of new kinds of data and even to feasible large-scale social experiments.

A provocative experiment that could be designed along these lines would be one that "auctioned off the exploration rights" along a portion of the human genome. Another portion of the genome could be selected for comparison; here the government could refuse to allow patent protection for basic results like gene sequences, and would offer instead to subsidize research on sequencing and on genetic therapies. If two large regions were selected at random, the difference in the rate of development of new therapies between the privately owned and the public regions and the differences in the total cost of developing these therapies could give us valuable information about the relative costs and social benefits of different financing mechanisms.

The experimental approach is unlikely to settle all issues about the appropriate federal role in funding research. In a

gene-mapping experiment, with one region assigned to the private sector and the other to federally sponsored researchers, differences in outcomes could be due to characteristics of the regions that were randomly assigned (and random assignment would not eliminate chance variation if the regions were too small). But many insights might emerge from such an effort, including the identification of cost consequences, the effect of funding source on ultimate access to resulting technological innovations, the dissemination of research results, the effectiveness of private sector firms in exploiting price discrimination, and so on. The scope for conducting such experiments might be large, and should be targeted toward those areas of research in which there is genuine uncertainty about the appropriate allocation between property rights and taxes and subsidy.

A more conservative strategy would be to collect detailed information about natural experiments such as the discovery and patenting of PCR. It would be very useful to have even ballpark estimates of the total monopoly price distortions induced by the evolving pricing policy being used by the patent holder.

Conclusions

Federal agencies often use estimates of industry revenues or consumer surplus to make claims about benefits or returns to their investments in scientific and technological research. Though these components of research productivity are important, they are inadequate as a basis for setting and evaluating government policy toward research. Our discussion has emphasized the choice between property rights and a system of taxes and subsidy (i.e., government sponsorship) for research. This decision is not made at the level of NIH or any other agency that sponsors and conducts research, but it is fundamental to public policy.

It may be tempting to dismiss these issues because it is so difficult to estimate the quantities that we identify as being central to decisions about government support for research. However, it would not be difficult to make rough estimates of these quantities and to begin to use them in policy discussions. Undoubtedly, it is difficult to select among the alternative mechanisms for supporting research. Nevertheless, decisions about the use of these mechanisms are made every time the government makes spending and property rights decisions relevant for science and technology policy issues. The effort required to obtain the needed information and consider these issues systematically might pay a large social return.

In coming years, three forces will increase the importance of taking this broad perspective on the federal role in supporting research. First, voters and politicians are likely to attribute a higher cost to taxes and deficit finance. As a result, in future years all federal agencies will likely be forced to rely less on the tax and subsidy mechanism for supporting technological progress than they have in the past.

Second, a dramatic reduction in the cost of information processing systems will increasingly affect all aspects of economic activity. This change will make it easier to set up new systems of property rights, which can be used to give private firms an incentive to produce goods that traditionally could be

provided only by the government. The rapid development of the Internet as a medium of communication may ultimately lead to advances in the ability to track and price a whole new range of intellectual property. The success of the software industry also suggests that other kinds of innovations in areas such as marketing may make it possible for private firms to earn profits from goods even when property rights to the goods they produce seem quite weak.

At the same time, a third force—the move toward managed care in the delivery of health care services—pushes in the other direction. This change in the market for health care services is desirable on many grounds, but to the extent that it reduces utilization of some medical technologies, it will have the undesirable side effect of diminishing private sector incentives to conduct research leading to innovations in health care. Everything else equal, this change calls for increased public support for biomedical research. In the near term, the best policy response may therefore be one that combines expanded government support for research in some areas with stronger property rights and a shift toward more reliance on the private sector in other areas. Further work is needed to give precise, quantitative guidance to striking the right balance. In the face of stagnant or declining resources, we will have to make increased efforts to gather and analyze the information needed to target research activities for subsidy and to learn which areas the private sector is likely to pursue most effectively.

A.M.G. is a Health Services Research and Development Senior Research Associate of the Department of Veterans Affairs.

1. Phelps, E. S. (1973) in *Economic Justice*, ed. Phelps, E. S. (Penguin Books, Baltimore), pp. 9–31.
2. Mishan, E. J. (1988) *Cost-Benefit Analysis* (Unwin Hyman, London).
3. Cornes, R. & Sandler, T. (1986) *The Theory of Externalities, Public Goods, and Club Goods* (Cambridge Univ. Press, Cambridge, U.K.).
4. Romer, P. (1993) *Brookings Pap. Econ. Act. Microecon.* **2**, 345–390.
5. Pauly, M. V. (1968) *Am. Econ. Rev.* **58**, 531–536.
6. Steering Committee of the Physicians' Health Study Research Group (1989) *N. Engl. J. Med.* **321**, 129–135.
7. Gold, M. R., Siegel, J. E., Russell, L. B. & Weinstein, M. C., eds. (1996) *Cost-Effectiveness in Health and Medicine* (Oxford Univ. Press, New York).
8. Fuchs, V. R. & Garber, A. M. (1990) *N. Engl. J. Med.* **323**, 673–677.
9. Feldstein, M. S. & Feenberg, D. (1996) in *Tax Policy and the Economy*, ed. Poterba, J. (MIT Press, Cambridge, MA), Vol. 10.
10. Rosenberg, N. (1994) in *Economics of Technology*, ed. Granstrand, O. (North-Holland, New York), pp. 323–337.
11. Antiplatelet Trialist Collaboration (1994) *Br. Med. J.* **308**, 91–106.
12. Yusuf, S., Peto, R., Lewis, J., Collins, R. & Sleight, P. (1985) *Prog. Cardiovasc. Dis.* **27**, 336–371.
13. Beta-Blocker Heart Attack Trial Research Group (1982) *J. Am. Med. Assoc.* **247**, 1707–1714.
14. Garber, A. M., Clarke, A. E., Goldman, D. P. & Gluck, M. E. (1992) *Federal and Private Roles in the Development and Provision of Alglucerase Therapy for Gaucher Disease* (U. S. Office of Technology Assessment, Washington, DC).
15. Beutler, E. & Garber, A. M. (1994) *PharmacoEconomics* **5**, 453–459.