CHAPTER III

CHRISTOPHER A. SIMS

I. Introduction

In recent years there has been increased emphasis in government on careful assessment of expenditures, in order to eliminate waste and redundancy and to improve effectiveness. To support this emphasis analytic tools for quantifying the benefits and costs of government activities have been brought in to play, thus reducing the need for subjective and qualitative judgment in budget decisions. Such tools have had important, though often controversial, impacts on several kinds of government programs.

Some have called for the application of cost-benefit techniques to statistical programs. The 1976 National Research Council report Setting Statistical Priorities argued that statistical budgets should not be exempt from quantitative analysis, that in principle the tools of cost-benefit analysis can be applied to statistical budgets as to any other budget, and that increased resources should be devoted to such quantitative analysis of statistical budget decisions.

This committee believes that the presumption that statistical budget decisions are in principle reducible to quantitative comparisons of costs and benefits, and that agencies should seek steady progress toward this type of decision making, has been counterproductive. As a relatively young statistical agency, EIA has been subject to more concentrated external scrutiny on this point than have other statistical agencies. Our review of EIA's efforts at evaluating its programs and of previous outside criticism of those efforts has convinced us that misplaced emphasis on quantitative precision in program evaluation has led to identifiable bad effects on evaluation programs. We document this point in the latter part of this chapter.

We believe that the persistent lack of success in providing quantitative bases for statistical budget decisions reflects not just a lack of analytical resources devoted to the task or a need for more methodological research in the area, but also a fundamental difference between statistical programs and other types of government program. Our

This material is a near-final draft of what appeared in the 1985 report of the National Research Council panel on statistics on natural gas.

reasons for this conclusion are spelled out in detail in the next section, but we can summarize them here. [1]

Cost-benefit analysis seeks to do for a government activity what an ordinary profitability calculation does for a private business activity. To be sure, most government activity in non-socialist countries occurs in the context of some kind of market failure: certain commodities may be improperly priced by the market or not marketed at all. Cost-benefit analysis differs from profitability calculations in that it confronts market failure with a set of techniques for constructing appropriate valuations for commodities which the market fails to price or prices incorrectly. Once the appropriate valuations have been constructed, cost-benefit analysis proceeds as would a profitability calculation.

But we are so used to taking profitability as a guide to resource allocations that we may forget that its validity depends on some assumptions. The notion that commodities have well-defined relative market values depends on the existence of unfettered competitive markets; such markets are defined in part by the costless flow of information about prices to all participants in them. Where some commodities are not properly priced by markets, the notion that there is a way to construct an appropriate value for them depends on costless flow of information about production and consumption opportunities. Obviously, a set of techniques grounded on assumptions of costless information flow may turn out to be unusable or unhelpful when applied to valuing an activity which, like statistical programs, has the objective of producing an information flow.

The logical incoherence of attempting to apply methods which assume information flow is free to evaluate the benefits of programs whose benefits are themselves information flows shows itself in three concrete problems in the application of cost-benefit analysis to statistical programs:

(1) Discovering who has obtained information and used it is practically impossible. Thus the usual first step in measuring benefits, identifying the beneficiaries, turns out to be infeasible. Usually information flow is rapid and cheap and occurs through diverse channels, including market transactions themselves;

[1]We do not mean to suggest that quantitative cost-benefit calculations are necessarily the best route to evaluating other kinds of government programs. The benefits of, say, clean air, well-maintained public gardens, or an informed electorate may be impossible in practice to quantify. If so, focusing attention on deeply flawed quantitative analyses of these programs may hinder good decision making, just as with statistical programs. But for these other kinds of program there is no logical contradiction in supposing that benefits can in principle be determined by summing up well-defined dollar values. In that sense these programs are fundamentally different from information programs.

furthermore, information can flow to affect a person's decision via prices or via setting norms of good practice without the person's consciously absorbing the information generating the effect. The diversity of channels of flow means that even tracking down everyone who read the publications or used the tapes in which a certain kind of data appeared would not come close to finding all the users, or even the most important users of data. The possibility of information influencing decisions without being consciously absorbed means that even locating everyone who "knew" the information would not suffice to locate all those whose decisions or behavior the information affected.

(2) Determining the value of data to those who benefit from it is ordinarily infeasible. The value of a statistical program depends in part on how much information would be available, and to whom, in the absence of the program. Since information flow without a government statistical program would ordinarily widespread, cheap, and through diverse channels, determining just what it would be is not in practice possible.

(3) For statistical programs it will often, perhaps even usually, be true that a useful cost-benefit analysis would have to be at least the same order of expense as the program itself. Locating beneficiaries of data programs, determining how they use the data, and determining what they would know in the absence of the data program, is itself a data-gathering activity. For most government programs, the usual approximation that information gathering and processing is of negligible cost is valid, protecting us against the absurdity of spending as much money on evaluating a program as executing it. There is no presumption that we can avoid such absurdity in analyzing a data program.

These difficulties face private sector firms as well as government, with understandable results. If cost-benefit analysis of government programs is analogous to a private firm's profitability analysis of an investment project or a production activity, cost-benefit analysis of data programs is akin to attempting to analyze the profitability of a firm's information gathering and dissemination programs – its accounting department and long term R and D. We do not see firms attempting to determine the profitability of their accounting departments or applying the same kind of profitability calculations to long term R and D as to expenditures on plant and equipment.

Thus expert analysis will not produce one-dimensional answers concerning the "value" or "usefulness" of data programs. The best application of analytical talent will produce results which cannot usefully be summarized in a cost benefit ratio, or even a range of cost benefit ratios.

Our conclusion regarding cost benefit analysis has important implications for how scarce analytical resources should be allocated in evaluating data programs. It has been critical in determining the analytical framework for this committee's own work concerning EIA's natural gas programs; it also conditions our evaluation of EIA's previous internal efforts at program evaluation and our recommendations for future EIA efforts along this line.

The alternative approach we adopted, described in more detail in chapter I, was to identify important current and future issues related to natural gas and associated decisions by public and private sector groups. Based on our assessment of the data needed to address these issues, we sketched a baseline of continuing data on natural gas — a kind of accounting system for the nation's gas production and use — and specific data oriented to understanding the behavior of the industry as the regulatory environment changes. Our assessment that data is needed for these purposes is certainly subject to dispute, but the dispute would have to concern complicated, multidimensional policy issues and could not be reduced to numerical cost benefit analysis. Once the value of a baseline data system and of specific data for tracking the effects of regulatory changes is accepted, though, we think it leads us to useful conclusions about budget priorities.

Our recommendations about specific components of the EIA natural gas data program are in effect conditioned on the existence of a national program to collect and disseminate energy data at roughly the current budget level. In this context, and using our judgment as to the purposes for which these data are useful, we can make recommendations suggesting shifts in emphasis or increases in effort without having to reach a precise conclusion about the value of EIA's program as a whole. Our attempts to provide useful conclusions in this frame of reference should not be taken to imply that we think the broader question, weighing EIA's programs as a whole against other uses of resources, is unimportant. A study focused on natural gas statistics is not the appropriate place to address the question of whether the recent shrinkage in EIA's budget has been good social policy on the whole.

As we will document in the following section, EIA itself has devoted substantial resources to evaluating its programs, and what it has done along this line has for the most part been valuable. It has been criticized, however, for not providing "incisive" analysis of the "usefulness" of its programs, even by representatives of agencies who do much less analysis of the usefulness of their programs than does EIA. We believe that such criticism pushes EIA's data evaluation programs in unproductive directions.

For example, evaluation programs that help determine how data is used or what public and private decisions depend on data are an important contribution to reducing the need for subjective judgment in budget decisions, even if they clearly hold no promise for eliminating the need for such judgment. Yet EIA's critics have emphasized attempting to survey "representative samples of users" rather than interviewing informally selected groups of users or discussing programs with advisory committees. We see this emphasis as motivated in part by the mistaken belief that evaluations which do not produce quantitative results are inadequate and that a formal sample survey might be capable of producing quantitative measures of data usefulness.

Another type of effect of misplaced emphasis on quantitative precision occurred in at least one evaluation study done for EIA by an outside contractor, which proceeded with a quantitative determination of cost benefit ratios. In doing so it focused attention on numerical results so heavily conditioned by arbitrary assumptions that they were useless, while diverting attention from valuable qualitative information in the same report on how data were being used.

A more detailed discussion of cost-benefit analysis is contained in the following section, after which we discuss procedures used by EIA and other federal statistical agencies to evaluate programs and set priorities. Some critiques of EIA's program evaluation methods are considered, and finally our recommendations for an issue-oriented evaluation process are set forth in Section V.

II. Why the Benefits of Data Programs Are So Hard to Measure.

II.1. **The main argument.** No analysis of how information should be produced and distributed can correctly proceed as if information were a commodity like other commodities.[2] Of course access to information is in fact given a price and traded. But information is unlike other commodities in that it automatically flows from one person to another as prices are quoted. There can be no clean separation between the exchange of the "commodity" and observation of the market prices which economic theory presumes individuals take as "given" in deciding how much of each commodity to buy and sell.

Our conclusion that information cannot be treated as an ordinary commodity is important to our approach, but to support it in detail requires discussion somewhat more technical than in the rest of the report. A reader who is convinced by the sketch

[2]We use "commodities" to refer to all goods and services which are traded or allocated in the economy.

of our argument in the first chapter and the first section of this chapter may skip or skim the remainder of this section. [3]

When resources are allocated efficiently, every pair of commodities will have the same relative value in every use and to every user. Any discrepancy in relative values represents an opportunity to increase welfare. In a market allocation system such discrepancies represent opportunities for trade, so that a freely functioning market tends to eliminate the discrepancies. It is because of this result that the practice of treating a commodity as having a value independent of how or by whom it is used can be justified. Cost-benefit analysis characterizes an activity as a set of commodities produced (benefits) and a set consumed (costs), values each of these commodities, and aggregates the values to calculate total costs and benefits.

In a market system when information flow about prices is costly, not everyone will in general know all prices. This will imply missed trading opportunities, which in turn means that the assumption that commodities have uniquely defined values breaks down. Many of the government's and EIA's most important statistical programs involve collecting and disseminating information about prices. The standard techniques of cost benefit analysis must fail in analyzing such a program. If the program has any value, then without it some information about prices would not be reaching everyone to whom it would be useful. That is, without the program, markets would not be functioning properly and commodities would be misallocated. The benefits of the program are precisely in its prevention of such misallocation. But reducing misallocation is not equivalent to producing and consuming some list of commodities, so the basic analytic framework of cost-benefit techniques does not apply. To determine the effects of the statistical program one has to analyze the degree of imperfection of non-market allocation mechanisms. This is a more difficult task than valuing a list of commodities, and a qualitatively different one.

The foregoing argument does not apply to price information alone; it extends to any data program which collects information from economically active agents and disseminates it to others who find it relevant to their economic decisions. In the last ten years or so economists have worked out the theory of market economies in which those in the economy start out with different information, but trade without restriction and with free flow of information about prices. The conclusion of this theory is that the free flow of information about prices generally suffices to transmit all information to everyone, so that the economy functions as if there were no private

[3]Readers familiar with cost benefit analysis or with statistical decision theory have taken issue with our argument, even in several cases when they agree with the conclusion that cost-benefit analysis of information programs is impractical. The appendix to this chapter describes some of the main objections we have encountered and gives our response.

information. (This is called the theory of "rational expectations equilibrium" and is surveyed in Grossman [1981].) Collecting economic data from people and making it publicly available can be a valuable activity only if without the activity markets would function imperfectly. Thus, like collecting and publishing price information, programs to disseminate other sorts of information will not yield to the standard tools of cost benefit analysis.

The argument applies to data collected for government decision-making with the same force as to data whose main benefit is to the private sector. The information collected will have value to the extent that it reduces or prevents misallocation of resources by the government. An efficiently functioning government will not misallocate resources and will leave commodities with unique values; analyzing the misallocation by a government agency with insufficient information raises the same problems for defining the value of commodities, and hence for cost benefit methods, as does analyzing misallocation by an imperfect market.

II.2. **An example.** We have noted that we believe there should be a base of data about reserves, production, and use of energy of all types, because of the special role of energy in the economy and the potential for disruption in petroleum supplies. One among many decisions to which such a data base is relevant is the decision as to whether to install fuel-switching capacity in a burner. Though many burners are privately owned, public regulatory policies are likely to affect the decision, so it has a mixed public-private character. The optimal decision depends on, among other things, how likely a supply disruption is, the costs of the switching capacity, the amount of switching capacity installed in other firms and industries, and the likely pattern of shifts in production and demand for final goods in the event of a disruption.

In the ideal world of allocation by markets, there would be markets in every commodity, including what are called "contingent commodities." One kind of contingent commodity would be "natural gas at location X contingent on the occurrence of a supply disruption of type Y at date T." A firm which could observe prices for gas and fuel oil and its own output contingent on every possible type of supply condition would have all the information it needed to make the decision about whether to install fuel switching capacity.[4] The rational expectations equilibrium result is that an EIA

[4]It may be worth noting that the firm would not, in a world of really complete markets, need to form judgments about the probability of supply disruptions. If it can buy inputs and sell its output contingent on disruption occurring or not occurring, it need only consider whether the value now of the additional sales it could achieve contingent on disruption if it had switching capacity exceed the costs now of that capacity and of the other inputs it would need contingent on disruption. The probability judgments would all be implicit in the prices of contingent commodities.

data dissemination program could not improve the resource allocation achieved by markets in this ideal world.

But of course contingent commodity markets are rare, so a burner owner must estimate for himself the likely supply and demand situations in the event of disruption. In doing this he is likely to find help in EIA and other data, in analyses and projections done by researchers, and in opinions of experts and business associates which are in turn based on the researchers' studies. The costs of decision-making with less EIA information will not be mainly in any evident mistakes in the decisions made by any one burner-owner, but in inconsistencies across burner owners. For example, one owner, having judged that a disruption is likely, demand for his product will be high despite the disruption, and that price differences between fuels are likely to be very large in the event of a disruption, may decide to install fuel-switching capacity. Another owner whose business is even more likely to have high demand in the event of a disruption and whose burner can be converted to fuel switching much more cheaply may judge that a disruption is unlikely or that fuel prices are likely to move in parallel during a disruption and hence decide not to install fuel switching. A complete market allocation would never invest in fuel switching in the first burner without investing in it also for the second, because the payoff, whatever it is, is certainly higher for the second. Yet, given his own judgments about the future, each owner is making a good decision. EIA data might, by improving the availability of high quality assessments of likely supply and demand conditions, reduce or eliminate the inefficiency.[5]

To put a value on the EIA data program's effects on fuel switching decisions, then, we have to analyze its effect on a pattern of misallocation. In how many cases, involving what dollar value of investment, would decisions show the type of inconsistency we describe? Answering this question is a tremendously complex problem when hundreds or thousands of burners are involved. To what extent do informal information flows and social consensus among businessmen create a uniformity in practices which prevents such misallocation? How far does the stock market alone, by increasing or decreasing firms' equity values as they install fuel switching capacity, go toward providing a sufficient signal as to which types of firms should install it?

[5]An ideal market could eliminate the discrepancy by allowing the first owner's beliefs about the likely value of fuel-switching to influence the second owner's investment decision. For example the second owner could sell rights to shares in his profits contingent on his installing fuel-switching and a disruption occurring. The first owner would then invest in these rights, financing the fuel-switching capacity in the second burner, before he began investing in his own fuel-switching. Both the first and second owners would probably modify their views about the future after seeing, via the market price of the contingent claims, more information about the views of others.

It is not that these questions have no answers at all; but answering them requires confronting sociological and psychological questions about human behavior that will not yield to the accounting schemes of cost benefit analysis. The critical commodities for which prices would have to be constructed in order to carry out a cost benefit analysis are contingent commodities. Where a government data program is potentially useful, contingent commodity markets will be incomplete and imperfect, meaning different firms and individuals will put different values on contingent commodities. Furthermore, since, both via prices and via other mechanisms (imitation, conversations over lunch), people's decisions may be influenced by information which they never consciously acquire themselves, it will be impossible to discover how information dissemination influences decisions by asking what information decision makers "know" or "use".

In this case it should be clear that determining the effects of data dissemination is itself an analytic problem of major dimensions, and in fact would require more data-collection, at a finer level of detail, than the data program whose effects were being measured.

II.3. **Where the ideas of cost benefit analysis can be useful for data programs.** There are circumstances where the tools of cost benefit analysis can, at least in principle, be helpful in evaluating data programs. While these circumstances apply to few if any EIA programs, we need to note their existence. Examples of this type are sometimes cited to prove that cost benefit tools can be applied to data programs, without recognition of their special character.

Cost benefit tools are certainly potentially useful in dealing with data programs whose value is primarily in creating information available nowhere in the economy rather than in transmitting information among economically active people. Examples would be programs of geological or oceanographic exploration or of scientific research. In this case it is reasonable to suppose that without the program, no one would have access to the information it generates, and that with the program everyone has access to it at negligible cost. (If in fact the cost of disseminating the information is a major fraction of the cost of the project, it does not fit this category.) The information generated cannot itself be treated as a commodity and priced, but it can be treated as generating or increasing the availability of a set of commodities.

Even where a data program involves mainly transmission of information, rather than creation of information new to everyone, cost benefit methods may apply at least in part when both of two special conditions hold:

(1) the barriers to information flow must be realistically treatable as isolated within a narrow production sector, with the rest of the economy operating with negligible costs of information flow; and
(2) we must be able to observe the performance of the sector both with and without the data program.

In appendix section A.4 we present examples of the special types of information programs for which cost benefit analysis of a sort might be possible.

EIA data programs for the most part assemble data from producers and consumers, summarize it, and transmit it. They are not mainly research programs producing new information. The benefits of EIA programs are not primarily in making a small sector – even the entire energy sector – more efficient. They are in making possible better decision-making and planning in every part of the economy which uses energy, which is nearly the entire economy. And in any case there is no possibility of making a "before and after" analysis of economic performance for the most important application of EIA statistics, improved ability of the economy to adapt to or forestall energy supply disruptions.

II.4. **The market test.** It might be thought that, whatever the difficulties of cost benefit analysis, we can at least be sure that whenever, in the absence of a government data program, the private sector would produce similar data, the activity is better left to the private sector. That this is not a generally applicable principle is important for this report, because in the case of data on reserves, an EIA program has displaced a similar private sector program and we do not recommend that the program revert to the private sector.

The fact that information does not behave like an ordinary commodity, which makes cost benefit analysis inapplicable to data programs, also implies that private sector profitability criteria are inappropriate for data programs. A program which assembles data and redistributes it can be profitable only to the extent that it involves barriers to information flow which are in some sense artificial. The social cost of transmission of data from one user to another is just the cost of reproduction. For a company to profit from selling data which it has generated at some cost, it must restrict the right of users to reproduce and distribute or resell the data. Copyright laws recognize this situation and represent a compromise between the need to guarantee low cost dispersal of information and the need to provide an incentive to collect it or generate it. But the market in access rights to information which such laws make possible is only a compromise. It is clear that generally a government program of information dissemination can charge lower fees and thereby disseminate the information more widely. That we have copyright (and patent) laws and do not have all information

dissemination in government hands reflects public concern that the government not be the only source of information dissemination and recognition of the fact that government, left to itself, is unlikely to do a good job of recognizing all the kinds of information which should be produced.

A profit-making data program also needs to preserve the uniqueness of the information it provides. Thus there is a general tendency for private providers of data to be less formal and explicit about the methods they use for gathering data than is standard in the government. If the statistical methodology underlying a private survey is completely explicit, and it develops a profitable market, then competitors can easily enter by meeting the same explicit statistical specifications. If the survey methodology is kept hidden competitors cannot easily imitate the survey. Users of the original survey will find its value only over time, with experience in using it for planning and prediction. Even a competitor who produces as good or better survey results will face a delay in convincing users that his survey is as useful as the original.

Of course some "private sector" data programs, particularly in the energy area, are run by nonprofit organizations, such as industry associations. Such organizations may run data programs much as the government does. However, such organizations seldom can maintain a really large scale data program because, as voluntary associations, they face the "free rider" problem in mounting costly programs even if they make all members better off. There can also be antitrust problems for extensive trade organization information dissemination programs. And finally, when the information collected is important outside the organization itself, problems arise. Most importantly, in public policy discussions of energy problems trade associations are seen as interested parties. Data they collect may therefore not be as useful in guiding policy formulation as government data, even when the latter is no more comprehensive or accurate, simply because the government data is not suspected of the same systematic bias.

Choice of whether a data program should be in the public or private sector does not rest, therefore, on whether it could be done profitably in the private sector. The choice is a difficult balancing of the benefits of cheap dissemination, explicit, formal methodology, and freedom from some kinds of political bias available in the public sector against the negative incentive effects on private initiatives to generate data if the government regularly displaces private data dissemination activities.

A data program which switches back and forth between sectors is likely to be worse than a stable program in either sector. A private program which faces the threat of government competition or takeover will have to focus more on short term profitability, exaggerating the disadvantages of private sector data programs. A public program

which may be impermanent will have difficulty attracting high quality professional staff. Because the allocation of data programs between public and private sectors is a difficult judgmental decision, there ought then to be a presumption in favor of leaving an existing program in the sector where it currently resides. If programs are moved from one sector to another on the basis of small shifts in the current view of which sector could best carry out the program, instability and inefficiency will result.

Appendix A

Here we discuss briefly what seem to us the most important counterarguments to our treatment of cost-benefit analysis of information systems. In each subsection below we first briefly present the counterargument, then provide our reaction.

A.1. **Rational Decision Making is Tautologically Cost Benefit Analysis.** The costs of statistical programs are not unusually difficult to measure. One way or another, decisions must be made as to whether to allocate available funds to pay those costs or instead to shrink or eliminate those costs and use the funds for another purpose. In making such choices we implicitly compare the benefits attainable for a given cost in the programs being compared. Why then object to cost benefit analysis of data programs? Budget decisions must be made, and they are bound to be based on cost benefit analysis, whether explicitly or implicitly.

Our view that cost benefit analysis does not apply to most data programs assumes that "cost benefit analysis" is not a mere label for any sort of rational decision making, but instead is a particular analytical approach. Cost benefit analysis, we take it, means describing a program as a set of commodities produced (benefits) and a set consumed (costs), and aggregating these using prices – market prices where possible, otherwise "shadow" prices which emerge from calculations based on assumptions of optimization, either by individuals or by components of a market economy. The problems of cataloging the costs and benefits, then constructing or finding the appropriate prices to apply to them, are treated as technical matters, and technical analysts are expected to produce cost benefit ratios unique up to sensitivity analysis in a few key dimensions.

With information dissemination programs, this analytical framework is not helpful. Technical analysts can determine some of the political and economic decisions to which the information is relevant, and they can look for alternative pathways through which the information might flow if the program were reduced or eliminated. But these efforts will involve tracing out the operation of incomplete and imperfect markets and of non-market information transfer mechanisms; the usual practices of relying on market prices and on the uniqueness of the values of traded goods will not

be available. Trying to proceed nonetheless to attach dollar values to the effects of the information will nearly always lead to a proliferation of guesswork and arbitrary assumptions which obscures, rather than clarifies, the analysis.

A.2. **The Value of Information in Decision Theory.** Statistical decision theory can be used to solve, for an individual decision-maker, the problem of putting a dollar value on information. Offered the right to observe a certain random variable at a certain price, the decision maker can determine whether his expected profit or utility is increased by the transaction. The value of the information in the random variable is just the highest price at which he is still willing to purchase the right to observe the random variable.

Given that this is so, why could we not value the benefits of a government information program by locating its users or potential users and determining how much the information generated by the program would be worth to them? Wouldn't the value of the government program be at least approximately the sum over users of the decision-theoretic value of information?

When the government program in fact creates information previously unavailable to anyone and makes it available to everyone, this line of reasoning is correct in principle.[6] There would be great difficulties in evaluating such a program, but they would be qualitatively similar to those in any other cost benefit analysis.

When the program is a data dissemination program, however, there will be unique problems in defining what it is that the program delivers to the individual. The program delivers a change in the information on which the individual's decisions are based, but he may not consciously use the information on which his decisions are based. To the extent the price mechanism functions properly, it may influence decisions, reflecting information present in the economy, without those influenced using the information directly. Where the price mechanism itself does not have this effect, other informal mechanisms for information transfer may create the same result. Firms or individuals may imitate the behavior of others whom they perceive to be better informed, say, without knowing what data or calculations determine the behavior they are imitating. The idea of using standard methods to value information at the level of the individual decision maker, then summing up, fails to apply to data

[6]One could not simply ask receivers of the information what it would be worth to them: if an individual received the information when no one else did, it would have speculative or resale value which would not be true social value; on the other hand if everyone received the information any one individual could hope to get the information without paying for it by watching the behavior of others. Nonetheless the principle that the benefit of the information to the individual is its dollar value to him, taking prices as given (to rule out speculation) and assuming there is no other route of acquiring the information than by paying for it, would be correct.

dissemination programs because of the difficulty of defining, at the individual level, the information which decision theory could value for us.

Because individuals will not necessarily know what information in the economy is influencing their behavior, there will be no way of interrogating them to determine the effect of an additional piece of information on their behavior. We will have to trace through the combination of imperfect market and non-market mechanisms by which information flows through the economy to influence behavior.

Consider again the burner-switching example discussed in the text. With an active stock market, firms may regard their proper objective as maximization of the market value of their shares. The stock market may give a clear signal that firms of a certain type should install fuel-switching capacity – such firms' equity values rise when they do so. The manager of such a firm might see no value to him in information relevant to the likelihood of a fuel supply disruption or to the likely relative prices of fuels in such a situation – the stock market signal might be sufficient to determine his decision. Yet the stock market signal might embody information available to others in the economy. In fact, the firm's manager himself, if he thought the stock market signal mistaken, would do best by adjusting his personal portfolio of security holdings, not by changing the behavior of his own firm. Giving him the results of a government survey might, if the survey information did not penetrate to most other participants in asset markets, lead the manager alter his personal portfolio without making any changes in his business decisions for the plant he runs.

If the market mechanism sufficed to transfer all available information to all firms, then a government program to, say, survey firms to discover the extent of existing fuel-switching capacity would not help decision-making. The information from the survey would already be reflected in market prices. Once we admit the possibility that the information would not be transmitted by the market, we would have to analyze the entire transmission mechanism in order to determine the survey's value — how many firms of which types are behaving in ways which would change if the government survey were introduced; what information is conditioning their behavior (not what information do they consciously use) in the absence of the survey; to what extent would barriers to information flow prevent the survey information from influencing behavior even of those to whom it was made available free. All this analysis, which would generally be more complex than cost benefit analysis itself, would be preliminary to any attempt at cost benefit analysis.

A.3. **All Cost Benefit Analysis is Imperfect.** Our claim that information flow is not free, market allocations therefore not efficient, and market prices therefore not unique, applies generally, not just when one thinks about information programs. Why

does our argument make cost benefit analysis any more inapplicable to information programs than to any other program?

The assumption that information flow is free is critical to cost benefit analysis of any program. But for many programs it will be reasonable to assume free information flow as an approximation. When a program's costs and benefits are large relative to the costs of information processing, the approximation will be good and may lead to useful results. For an information-dissemination program, the approximation leads immediately to the trivial result that the program is worthless.

A.4. **Examples Where It Has Been or Could Be Done.** *Setting Statistical Priorities* pointed out the dearth of serious attempts at cost benefit analysis of data programs. In the time since that report, there have been only a few new attempts at it. Bruce Spencer has done a cost benefit analysis of Census area statistics; the FTC did a cost benefit analysis of its "Line of Business" statistics. This is not the place for a careful discussion of the methodology and conclusions of those studies. Suffice it to say that we do not know of other important examples of such studies (though there may be a few), and we detect no professional consensus that either study displays a methodology that is widely and easily applicable.

We did point out in the text that cost benefit analysis is at least in principle possible for certain special types of data program. One of these is a program which generates new information and distributes the information to everyone at negligible cost.

For example, suppose there is a region in which mineral rights have diverse ownership and there is uncertainty about the geological structure. A government geological study could reduce that uncertainty, and, because the survey would have to apply to the whole region no private landowner would be ready to finance the study. The survey would in effect make available an array of commodities of the form "an acre of land in location X given that the study has results Y" which would not be available without the study. These commodities would have well defined prices if markets in them existed, so cost benefit methods apply in principle. Since markets probably do not exist for them, a cost benefit analyst needs to form "shadow prices" for them.

In theory the government could apply a pure market test to its decision to make the survey by organizing a special auction. It would have landowners and potential buyers get together and arrive at competitive market prices for all the land in the area contingent on their being no survey, and also contingent on their being a survey. The prospect of the survey, which would allow attempts at exploitation of the mineral rights just in cases where drilling or mining looked promising, would increase the value of the land. The survey would be worthwhile if the increase in the land value exceeded the cost of the survey.

Of course organizing such an auction would ordinarily be impossible, so to apply cost benefit methods one would have to find ways to estimate what the results of such an auction would be. The most likely method would be to assume that the absence of complete markets did not prevent approximately efficient behavior of those owning and exploiting the mineral rights. By attempting an assessment of how much uncertainty there is without the survey, and the likely cost savings and production improvements available from efficient use of the survey by an ideal social planner, one could make a start at putting a value on the survey. Solving this planning problem would be difficult indeed, but at least in this case, unlike the case of information-dissemination, the technique of treating the market mechanism as mimicking the solution of an optimization problem does apply.

An information-dissemination program could allow cost benefit analysis if it applied to a narrow sector of the economy and the performance of the sector could be measured directly with and without the program.

For example, consider a possible program to survey corn farms every week during the appropriate season to determine the extent and geographical pattern of corn borer infestation. Suppose the main potential benefit is that with accurate and timely information about infestation, farmers likely to be affected can apply expensive preventive control measures which are not cost-effective unless the likelihood of infestation is high. We suppose that farmers whose crop has already been infested know of the infestation and discuss it with their neighbors, insecticide salesmen, and the like, so that some information about infestation would spread with or without the survey. Estimating benefits of the program directly, before it went into effect, would therefore be a formidable problem, not amenable to cost benefit methods. But if the program were tried, and we had data on production during infestations before and after the program went into effect, we might well be able to estimate benefits. We would take the market value of corn and insecticides as given and value the program by the difference between the value of corn it saved and the sum of the cost of the program and the cost of control measures it induces farmers to take.

For this example to meet the two special conditions, we have to ignore the fact that the survey would also affect futures prices, and thereby decisions of people other than farmers.

Some may take the position that programs with the special characteristics we have set out here include most government data programs, and that therefore our claims are beside the point. While this position merits discussion, it seems not to apply to EIA's programs.

A.5. **Cost Benefit Analysis Is Impractical for Information Programs for the Same Reasons It Is Impractical for Many Other Programs.** Information programs tend to benefit large numbers of people, mostly indirectly and in ways of which they are unaware. This can be described by saying that information flow and use generates very strong "externalities." Where there are externalities, markets are tautologically imperfect or incomplete. Externalities are difficult to deal with in any cost benefit analysis, and information programs are special only in that they are associated with large and numerous externalities. This does not make them qualitatively different from other programs. We could base our claim that cost benefit analysis of information programs is impractical by pointing to these substantial, but standard, difficulties and to the absence of convincing examples of cost benefit analysis of data programs. Our appeal to an academic argument about "logical incoherence" of cost benefit analysis is unnecessary and needlessly provocative. Some would also argue that our distinction between data dissemination and data generation programs is misleading — in practice the difficulty of cost benefit analysis for information generation programs may match that for pure dissemination programs.

To some extent this argument is a matter of semantics and exposition. We do think it important that there is a general case against the practicality of cost benefit analysis of data dissemination. If the point is not made convincingly, some might argue that the appropriate reaction is to put more resources, not less, into cost benefit analysis of data programs, hoping to overcome the difficulties.

We do not intend to prejudge the issue of cost benefit analysis of data generation programs, with which we have not had to deal in discussing natural gas statistics. We think it is important to understand that cost benefit analysis of data generation programs, whether practical or not, can avoid a host of problems which beset such analysis for dissemination programs and which are by themselves sufficient to make cost benefit analysis of dissemination programs generally useless.