Risk Analysis, Vol. 26, No. 3, 2006 DOI: 10.1111/j.1539-6924.2006.00774.x

What Affects the Quality of Economic Analysis for Life-Saving Investments?‡

Robert W. Hahn,[1]* Katrina Kosec,[1] Peter J. Neumann,[2,3] and Scott Wallsten[1]

Economic analysis of life-saving investments in both the public and private sectors has the potential to dramatically improve longevity and the quality of life, but only if the analyses on which decisions are based are done well. In this article, we analyze a data set that provides information on the content and quality of journal articles that measure the cost-effectiveness of life-saving investments. Our study is the first to provide a detailed multivariate analysis of factors affecting objective measures of quality. We also explore whether a series of recommendations by an expert panel convened by the U.S. Public Health Service affect the way analyses of specific life-saving investments are done. Our results suggest that four factors are positively correlated with an index we construct to measure analytical quality: (1) having at least one author affiliated with a university, (2) publication in a journal that has experience in publishing these analyses, (3) if the life-saving investment is located in the United States, and (4) if the analysis considers a measure of social costs or benefits. Somewhat surprisingly, a study's funding source and whether it is affiliated with industry are not significantly correlated with the quality index. Finally, neither time nor the panel guidelines had an impact on the index.

KEY WORDS: Cost-benefit analysis; health economics; risk analysis

Governments around the world are experiencing increasing demands to address health, safety, and environmental concerns. At the same time, political leaders are seeing a need to use good science and economics in making policy choices aimed at saving lives and improving the quality of life.[4] Partly in response to this need, policy analysts are relying more heavily on two tools widely embraced by applied economists: cost-benefit analysis and cost-effectiveness analysis (CEA).

Cost-benefit analysis attempts to measure the costs and benefits of a decision and to monetize them where possible.[5] CEA typically monetizes costs, but leaves benefits in terms of other units, such as statistical lives or quality-adjusted life years.[6] It is widely recognized that cost-benefit analysis and cost-effectiveness analysis can play a critical role in informing decisions aimed at promoting pubic health and saving lives.[7]

1 AEI-Brookings Joint Center for Regulatory Studies, Washington, DC, USA.
2 Center for the Evaluation of Value and Risk in Health, Tufts-New England Medical Center, Boston, MA, USA.
3 Tufts University School of Medicine, Boston, MA, USA.
‡The views expressed in this article reflect those of the authors and do not necessarily reflect those of the institutions with which they are affiliated.
* Address correspondence to Robert W. Hahn, AEI-Brookings Joint Center for Regulatory Studies, 1150 17th Street, NW, Washington, DC 20036; tel: 202-862-5909; fax: 202-862-7169; rhahn@aei-brookings.org.
4 See, e.g., Blair (2005).
5 See, e.g., Arrow *et al.* (1996).
6 See Lave (1981) and Russell *et al.* (1996).
7 Tengs *et al.* (1995) develop an extensive data set on cost-effectiveness that can be used for making such calculations. Morrall (2003) provides a good overview of the potential to save

0272-4332/06/0100-0641$22.00/1 © 2006 Society for Risk Analysis

The potential for economic analysis to improve decision making was one of the main motivations for requiring such analysis for U.S. federal regulations.[8] Every president since Reagan has required executive regulatory agencies to perform cost-benefit analyses for most regulations estimated to impose annual costs of over $100 million.[9] These regulatory impact analyses inform important government decisions, such as whether to require airbags in vehicles or reduce particulate emissions. Such decisions are estimated to result in roughly $200 billion of expenditures annually.[10]

In some cases, economic analyses of public policy decisions have had a marked impact on health-related issues. For example, the decision to curtail the use of chlorofluorocarbons and the decision to phase out lead from gasoline were both influenced by the careful application of cost-benefit analysis.[11] And CEA played an important role in the decision to implement a market-based approach for reducing sulfur dioxide emissions.[12]

Because of the potential importance of cost-benefit analysis and CEA in informing policy decisions, a number of organizations have issued guidelines that researchers and government officials may want to follow. For example, the U.S. Office of Management and Budget issues guidelines to federal agencies on the development of regulatory analysis. These guidelines aim to improve the quality of regulatory analysis and to standardize the measurement and reporting of costs and benefits.[13] Similarly, the U.S. Public Health Service organized the Panel on Cost-Effectiveness in Health and Medicine to establish protocols for analyzing the cost-effectiveness of health interventions. Panel members included 13 scientists and scholars with expertise in CEA, clinical medicine, ethics, and health outcomes measurement. The panel met 11 times between 1993 and 1996. The panel offered recommendations directed toward both authors and users of CEAs for improving the quality as well as the comparability of CEAs. The major panel recommendations appeared in a three-part Consensus Statement published in the *Journal of the American Medical Association* in 1996 (Russell *et al.*, 1996) and in a book (Gold *et al.*, 1996). These recommendations broadly apply to all CEAs of interventions designed to improve health. As described in Russell *et al.* (1996), the panel's focus was on "policy decisions and resource allocation at a broad level."

Economic analysis of life-saving investments in both the public and private sectors has the potential to dramatically improve longevity and the quality of life, but only if the analyses on which decisions are based are done well. Currently, one of the key impediments to the effective use of such analyses is the lack of quality and comparability with other studies.[14]

There are three basic approaches to assessing the quality of such analyses. One is to conduct case studies of particular examples.[15] The main advantage of this approach is that detailed analysis of individual cases can highlight the strengths and weaknesses of the data, assumptions, and underlying models. A major weakness is that the analytical methods are not easily generalized and the results are not easily replicated. Furthermore, reasonable people may disagree about the relative quality of analyses because of the highly subjective nature of the case-study approach. A second approach uses estimates of a key parameter, such as net benefits or cost effectiveness, from studies done prior to and after the implementation of a policy. The idea is that the estimate done after a study may provide a better measure of the actual impact of a policy.[16] A third approach uses the degree of compliance with the objective recommendations

lives and/or reduce costs. Murphy and Topel (2003) evaluate the value of medical research.

8 See, e.g., Litan and Nordhaus (1983). For an insightful assessment of the importance of analysis, see Farrow (1991).

9 See Smith (1984) and Viscusi (1996) for illuminating discussions on the efforts to reform regulation.

10 See, e.g., Arrow *et al.* (1996).

11 See Nichols (1997) on the lead phase out and Hammitt (1997) on chlorofluorocarbons. See also Sunstein (2002) for a good overview. For an insightful study of whether regulators make rational decisions, see Viscusi and Hamilton (1999).

12 See, e.g., Schmalensee *et al.* (1998) and Stavins (1998).

13 Office of Management and Budget (2003). In 2003, OMB issued Circular A-4, which refines the OMB's regulatory impact analysis guidelines of 1996.

14 See, e.g., Russell *et al.* (1996). Even if CEAs are of high quality, however, there are other barriers to their use. For instance, political considerations may impede efficient decisions.

15 See, e.g., Morgenstern (1997).

16 See, e.g., Harrington *et al*. (2000). This approach poses a number of challenges from the standpoint of measuring quality. First, the quality of the analysis is contingent on the state of information available when the studies are done. Second, there is no reason to necessarily believe estimates done after the fact are necessarily more accurate, even though they presumably rely on more recent information. Third, there are relatively few studies of this kind up to this point, and they are costly to do. All of these problems can be remedied to varying degrees. For example, to address the information problem, one could use information markets. Indeed, these markets provide yet another way to make assessments of quality at a given point in time or over time. See, for example, Abramowicz (2004). The problem is that such markets are not widely available at this point.

of experts as a measure of the quality of an analysis.[17] An advantage of this approach is that it is easily generalized and replicated and allows comparison of a wide range of different analyses. A drawback is that satisfying all of the recommendations of experts does not guarantee a high-quality analysis. Thus, compliance with objective recommendations is likely to be an imperfect proxy for quality. Nonetheless, using objective recommendations to judge quality is perhaps the best way to draw conclusions based on hundreds of disparate analyses.

Few studies have used regression analysis to address factors affecting analytical quality and whether economic analysis has actually improved over time. In this article, we analyze a data set that provides information on the content and quality of journal articles that measure the cost effectiveness of life-saving investments. Our study is the first to provide a detailed multivariate analysis of factors affecting objective measures of quality. We also explore whether the recommendations by the expert panel convened by the U.S. Public Health Service affect the way analyses of specific life-saving investments are done.

Our results suggest that four factors are positively correlated with an index we construct to measure analytical quality: (1) having at least one author affiliated with a university, (2) publication in a journal that has experience in publishing these analyses, (3) if the life-saving investment is located in the United States, and (4) if the analysis considers a measure of social costs or benefits. We find that a study's funding source and whether it is affiliated with industry are not significantly correlated with the quality index. Finally, neither time nor the panel guidelines had an impact on the index.

The article is organized as follows. Section 1 offers background information on research regarding the quality of cost-benefit analysis and CEA. Section 2 reviews our basic approach and summarizes the data. Section 3 presents our empirical tests and describes the results. Section 4 concludes.

1. BACKGROUND

Only in the last decade have scholars systematically examined the quality of cost-benefit analyses and CEAs. This research has focused primarily on CEAs of health interventions and cost-benefit analysis of economic regulations. Both types of studies have identified factors that could affect analytical quality. Some studies find evidence that the quality of analyses has improved over time while others find improvements in certain areas but no overall improvement. Studies of the regulatory impact analysis literature have used multivariate analysis to determine the factors affecting quality while simultaneously controlling for other study factors. Studies focusing on the quality of CEAs, however, have not. Here, we present what is known about the factors affecting analytical quality and whether it has improved over time.

Studies focusing on the medical economics literature have identified a number of factors that may affect the quality of a CEA.[18] Udvarhelyi *et al.* (1992) find that articles in general medical journals are more likely to use analytic methods appropriately than articles in general surgical or medical subspecialty journals. Gerard *et al.* (1999) find great variation in analytical quality among published CEAs, but find that analyses appearing in specialist medical journals tend to be worse than those in more general medical journals. Similarly, Neumann *et al.* (2000b) find that a journal's experience in publishing CEAs is more important than the type of journal in explaining quality, although general medical and surgical journals tend to have better reporting practices, conduct certain kinds of sensitivity analyses, and discuss ethical implications better than subspecialty journals. Using a larger data set, Neumann *et al.* (2005) find that quality increases with a journal's experience in publishing CEAs, but that industry funding does not affect analytical quality. Using bivariate analysis, these studies suggest that CEAs have improved over time and that industry funding does not affect the quality of an analysis.[19]

Studies focusing on the quality of regulatory impact analyses have also identified factors that likely affect quality. Hahn *et al.* (2000) reviewed 48 major environmental, health, and safety regulations and their associated regulatory impact analyses conducted by U.S. government agencies between 1996 and 1999. The authors use a number of objective measures of quality, such as whether the analysis provides a point estimate or a range estimate of net benefits and whether

[17] Our "objective" measures of quality are arguably subjective in certain ways. Nonetheless, they are measures of quality about which there is little debate or controversy; most would agree that a high-quality analysis should include them.

[18] Frequently, the analyses measure the cost per quality-adjusted life year. Such analyses are sometimes referred to as cost-utility analyses.

[19] One study, Neumann *et al.* (2000b), does include some multivariate analysis. This analysis, however, did not control for the myriad factors that could affect quality.

it uses a discount rate.[20] The study suggests several reasons for a relatively low level of quality, including resource constraints and the possibility that an agency may not want interested parties to know that the costs of a regulation exceed the benefits. Hahn and Dudley (2004) use multivariate analysis to review 74 regulatory impact analyses of health, environmental, and safety regulations spanning the Reagan, first Bush, and Clinton administrations. The authors find that measures of quality are generally unrelated to key factors such as the size of the regulation and under which administration it was implemented. They do not address the effects of funding source on quality because the government funds regulatory impact analyses.[21]

Studies of the health and medical CEA literature have also examined whether analytical quality has improved over time, but generally do not control for other factors. Udvarhelyi *et al.* (1992) find no significant improvement in the quality of CEAs between 1978 and 1987. Chapman *et al.* (2000) find some evidence suggesting that quality has improved over time for some factors, such as discounting. This suggests that quality, as measured by these criteria, has improved over time. Neumann *et al.* (2000b) find that the number of CEAs published annually has increased over time and that the average quality of those analyses has improved slightly over time.[22] Neumann *et al.* (2005) find evidence that CEA quality seems to be improving, but emphasize that many studies still fail to follow recommended protocols. The authors also note an increased use of a 3% discount rate and a societal perspective after the Panel on Cost-Effectiveness in Health and Medicine recommended them in 1996—measures that could make studies more easily compared to other studies and more useful for public policy.

To our knowledge, only one study examines changes in the quality of regulatory impact analyses over time. Hahn and Dudley (2004) use a multivariate analysis and find that measures of quality are generally unrelated to the year in which a regulation is implemented. This result is similar to the findings of Udvarhelyi *et al.* (1992), but different from the findings of more recent analyses, including Chapman *et al.* (2000), Neumann *et al.* (2000b), and Neumann *et al.* (2005). One reason for the apparently conflicting results could be that the studies focusing on the CEA literature do not generally use regression analysis while the study focusing on regulatory impact analyses does. It is possible that their results would evaporate upon controlling for other study and author factors.

As noted, only the literature reviewing the quality of regulatory impact analyses has controlled for multiple factors that could affect the quality of analysis in a multivariate framework. It is also the only literature that has used a rigorous multivariate analysis to test whether official guidelines actually affect the quality of economic analysis.[23] As a result, it is still unclear what factors affect the quality of CEAs and whether these analyses improved following the 1996 recommendations of the Panel on Cost-Effectiveness in Health and Medicine. It is to these issues that we now turn.

2. DATA

Our analysis is based on data from the Cost-Effectiveness Analysis Registry, which is maintained by researchers at the Tufts-New England Medical Center.[24] The registry, intended to be comprehensive,

20 In Hahn *et al.* (2000) and Hahn and Dudley (2004), scoring involved a primary scorer and a second individual who validated the findings of the first scorer. The primary scorer used the regulatory impact analysis to evaluate each item on the scorecard. The second researcher validated the first researcher's findings by reviewing the analysis and the completed scorecard. If the findings of the two researchers differed for any part of the scorecard, the researchers resolved the differences by discussion.

21 While regulatory impact analyses are usually funded by the government, they are sometimes carried out with the assistance of nongovernmental organizations and the private sector. Hahn and Dudley (2004) do not examine the effect of author affiliation on analysis quality.

22 In Neumann *et al.* (2000b) and Neumann *et al.* (2005), a number of readers, all with masters or doctoral degree training in decision analysis and CEA, scored the CEAs on a subjective, Likert scale ranging from 1 (low) to 7 (high). Scores were assigned based on overall study quality. Two readers independently read and scored each CEA.

23 See Hahn and Dudley (2004).

24 The CEA registry has been supported in the past by Grant R01 HS10919 from the Agency for Health Care Research and Quality and support from the Harvard Interfaculty Program for Health Systems Improvement. It is currently supported by the National Library of Medicine, Grant GO8LM008413. It is now housed at the Tufts-New England Medical Center. Registry investigators attempted to include every CEA published between 1976 and 2001 that met specified criteria. The criteria include the following: the analysis must be written in English, have been published in a peer-reviewed medical journal (ruling out government reports), specifically use the measure "$/quality adjusted life year" (or some other currency), and include original quality-adjusted life year measures (ruling out articles that do not provide original estimates). Tufts University-New England Medical Center (2005).

includes 533 CEAs published in peer-reviewed medical journals between 1976 and 2001.[25] The analyses fit into five main categories: health care, occupational or school-based, transportation, residential, and environmental.[26] The registry contains data on the methods used to estimate costs and effects as well as detailed information on the authors and analyses themselves. Two readers with masters or doctoral training in decision analysis or CEA independently reviewed each study, completed a data abstraction form, and then met to resolve any discrepancies in their ratings.

Table I defines the measures of the CEAs in the registry that we use in our analysis. We use these variables in several ways. First, we create dummy variables indicating whether the author included each of a number of recommended research methods in his or her CEA. Second, we create two indices that count the number of these methods that the author included in his or her analysis. We designed the first index, which we call the "quality index," to measure basic analytical quality. It ranges from zero to four, depending on the number of a set of four particularly well-defined methods the author included: whether the author performed incremental analysis, stated the year of currency used in the cost estimates, performed sensitivity analyses,[27] and discounted costs and quality-adjusted life years. The four components of the quality index represent what we believe to be the bare minimum requirements for an analysis to be useful.[28]

Our second index, which we call the "panel index," measures the compliance of CEAs with recommended protocols—whether they are intended to improve empirical quality (such as the components in the quality index) or the comparability of disparate analyses.[29] This index, which ranges from zero to eight, includes all measures related to recommendations by the panel. In addition to the measures in the first index, it also includes whether the authors included probabilistic sensitivity analysis, model validation, a 3% discount rate, and a societal perspective.[30]

A 3% discount rate, a societal perspective, or probabilistic sensitivity analysis are not necessarily superior to other discount rates, perspectives, and sensitivity analyses. Nonetheless, these were the recommendations of the Panel on Cost-Effectiveness in Health and Medicine, and our aim in this component of the analysis is to measure compliance with these recommendations. Our "quality index" includes fewer variables than the "panel index" because some

25 Our analysis uses CEAs in the Cost-Effectiveness Analysis Registry written between 1992 and 2001. This encompasses the vast majority of analyses (484 of the 533 total analyses currently in the registry).

26 The vast majority of the studies in the registry are in the health care category. There is a small number in each of the other categories. Health care studies cover a range of medical and health care interventions. An occupational/school-based CEA might examine the effectiveness of a particular vaccine in the military or the effect of a school-based tobacco-use prevention program on teen smoking. A transportation analysis might study the effectiveness of air bags, seat belts, or laws against the use of cell phones while driving in preventing traffic fatalities. A residential analysis might examine the effectiveness of a tuberculosis screening process in long-term care facilities or the effectiveness of insecticide-treated bed nets in avoiding child mortality.

27 The reliability of this measure may be affected by a slight change in the question asked by the registry team over time (in 1997 and subsequent CEAs, it was different than for earlier analyses). In addition, the registry team was occasionally not clear if what a study did should qualify as sensitivity analysis. In general, readers were generous in determining whether the authors incorporated sensitivity analysis.

28 While a large number of studies score highly on the index, our index encompasses fairly minimal requirements for a useful analysis. The index does not take into account many other factors that contribute to a study's quality, nor does it capture *how well* an author does the factors that are included in the index (such as sensitivity analysis and incremental analysis). Such additional data were not available. Nonetheless, we think the quality index does capture quality in a very basic sense; without these four components, few economic analyses could be considered to be of high analytical quality.

29 Others, such as Chiou *et al.* (2003), have devised systems for appraising the quality of CEAs. In the interest of clearly distinguishing measures of hard analytical quality (those in the 0–4 index) from more subjective measures of quality (such as many of the variables in the 0–8 index), we wanted to design our own instruments for measuring quality as well as compliance with recommended protocols.

30 We do not include performing model validation in the 0–4 index because of how loosely the variable was defined. As described in Table I, when the registry team compiled data on validation, they characterized validation very generously; if the authors compared modeling projections for disease progression to actual data (empirical work), contrasted results from different models, or even discussed "model validation," they were given credit for performing model validation. Some of this might be called "calibration" rather than "validation," and we were uncertain whether it was a consistently computed measure of analytical quality. Given that model validation is specifically advocated by the 1996 panel recommendations, however, we included this variable in the 0–8 index.

Table I. Definition of Variables

Variable	Meaning
Research sponsor	This is the source of funding of the research carried out by the authors of the cost-effectiveness analysis.
Authors' affiliation	Authors' affiliation indicates the authors' primary association or membership when the article was written. Where a single article has multiple affiliations, some of the authors had one affiliation (e.g., university) while other author(s) had another (e.g., government). The data do not indicate the affiliation of the primary author, so we only know that one or more authors have the listed affiliation(s) (not which ones).
Authors' perspective	Authors' perspective indicates the perspective from which the authors calculate costs and benefits. If the perspective is governmental, the analysis addresses how much something would cost the government (e.g., the cost of a proposed law for the government). A health care perspective looks at the costs that face HMOs, hospitals, etc. A patient's perspective looks at patient costs (e.g., the cost of a procedure and the amount of time the patient is unable to work). When authors take a societal perspective, they consider "everyone affected by the intervention, and all health effects and costs that flow from it are counted, regardless of who would experience them" (Russell *et al*., 1996). Authors declare their perspective, although about 35% of authors in the sample did not declare any perspective.
Reader's perspective	According to the registry team, sometimes an author would claim to have one perspective, but it appeared to the reader (a member of the registry team) that he or she had another. As a result, the researchers created two perspective variables: the authors' perspective and the reader's perspective.
Country of study	This is the country on which the study is based. In some cases, this is a region (such as sub-Saharan Africa or Western Europe) but usually it is a single country. All analyses fit into one of three categories: United States, non-U.S. Organization for Economic Cooperation and Development (OECD), and non-OECD studies.
Model validation	Model validation is generously characterized. If the author compared modeling projections for disease progression to actual data (empirical work), contrasted results from different models, or even "discussed" model validation, the registry team credited the analysis with including model validation. In some cases, a study's "model validation" might be more appropriately called "calibration."
Discount rate	The discount rate is the interest rate at which the authors discount future events. A present-oriented author discounts the future heavily, using a high discount rate.
Journal impact factor	A measure used to rank the importance of scientific journals. It is a measure of the frequency with which the "average article" in a journal has been cited in a particular year or period. Generally, the higher the impact factor, the higher the "impact" (i.e., "quality" or "prestige") of the journal.
Average analyses, current and previous 5 years	This is a number indicating the average annual number of CEAs published by the journal during the current and previous 5 years. It is a measure of journal experience with publishing CEAs. The higher this value, the more experienced.
Incremental analysis	The difference in the cost of two different courses of action is compared to their difference in outcomes by dividing the former by the latter. This ratio is known as the incremental cost-effectiveness ratio.
Sensitivity analysis	Sensitivity analysis is a procedure used to ascertain how closely a given model output depends upon the input parameters. Authors do this by varying model input parameters over a reasonable range and observing the relative changes in model response. If a small change in a parameter results in relatively large changes in the outcomes, the outcomes are said to be sensitive to that parameter. This may mean that the parameter has to be determined very accurately or that the alternative has to be redesigned for low sensitivity.
Probabilistic sensitivity analysis	Probabilistic sensitivity analysis is a type of sensitivity analysis for which outputs are calculated based on random assignment of values to inputs (drawn from a user-selected probability distribution).

Source: Conversations with the Cost-Effectiveness Analysis Registry Team (2005).

recommendations were intended to make CEAs more comparable while others were intended to increase analytical quality.

Finally, we also use a subjective score of overall analysis quality (on a scale of 1 (low) to 7 (high)), which the reader who catalogued the analysis for the database assigned. Table II shows the means of these variables.

The data set also included other useful information: the journal in which the CEA appeared,

Table II. Means, All Dependent Variables

Variable	Mean	*SD*
Subjective reader rating (0–7)	4.2	1.3
Number of quality measures that = 1 for the analysis (0–4)	3.2	0.94
Number of quality measures that = 1 for the analysis (0–8), using authors' perspective	4.2	1.6
Number of quality measures that = 1 for the analysis (0–8), using reader's perspective	4.1	1.5
Dummy — incremental analysis performed (included in 0–4 and 0–8 indices)	0.9	0.30
Dummy — discounts costs and quality-adjusted life years (included in 0–4 and 0–8 indices)	0.81	0.39
Dummy — year of currency for cost estimates stated (included in 0–4 and 0–8 indices)	0.76	0.43
Dummy — sensitivity analyses performed (included in 0–4 and 0–8 indices)	0.91	0.28
Dummy — analysis meets at least two of the quality measures in the 0–4 index	0.93	0.25
Dummy — analysis meets at least three of the quality measures in the 0–4 index	0.81	0.40
Dummy — analysis meets all four of the quality measures in the 0–4 index	0.49	0.50
Dummy — probabilistic sensitivity analysis performed (included in 0–8 index)	0.086	0.28
Dummy — model validation (included in 0–8 index)	0.15	0.35
Dummy — 3% discount rate (included in 0–8 index)	0.36	0.48
Dummy — societal perspective (authors') (included in 0–8 index)	0.39	0.49
Dummy — societal perspective (reader's) (included in 0–8 index)	0.27	0.45

Note: $n = 533$. Numbers are rounded to two significant digits. Variables are defined in the text.

the year it was published, primary research funders, authors' affiliations, authors' perspective(s),[31] and the country or countries on which the study focused (Table I).[32] Table III details the major categories of funding sources, affiliations, and authors' perspectives available in the data. The authors' perspective categories were designed to be exhaustive (if a perspective does not fit into a societal, governmental, health care sector, National Health Service, or unstated perspective, it is considered "other"), and the base group is a consultant or contract researcher. The five major types of funding (industry, government, foundation, university, and not stated) are nearly exhaustive, although there are a small number of studies that fit into none of these categories (e.g., a very small number were funded by private individuals). In addition, the funding source categories are not mutually exclusive. For example, studies could be funded by a university and a foundation. We also created three dummy variables indicating whether the study was conducted in

[31] The "authors' perspective" variable is a bit problematic. Sometimes, authors say they take a societal perspective when they do not, such as when they exclude costs that should be included in a societal perspective (such as non-health-care costs) and claim that they do not matter in the case at hand. In addition, there is no consensus regarding what, exactly, constitutes a societal perspective. In addition to these inherent difficulties, the CEA registry team changed over time, meaning that reader's own interpretations may have changed (while the team made efforts to standardize results, and while the leadership of the registry has remained constant, coding practices on this variable could have changed over time).

[32] CEAs can have multiple funding sources and, for analyses with multiple authors, multiple affiliations. In the data set, most studies take only one perspective, but a small number take a government perspective and an NHS perspective. Researchers on the registry team tried to assign a single perspective to most studies; as a result, some studies that really took two perspectives were assigned a single perspective. For example, studies that took both a health care perspective and a societal perspective were categorized as "societal."

Table III. Most Common Funding Sources, Affiliations, and Author Perspectives

Funding Source	Authors' Affiliation	Authors' Perspective
Industry	Industry	Societal
Government	Government	Government
Foundation	University	Health care sector
University	Health care organization	National Health Service
Not stated	Consultant/contracted researcher	Not stated
		Other

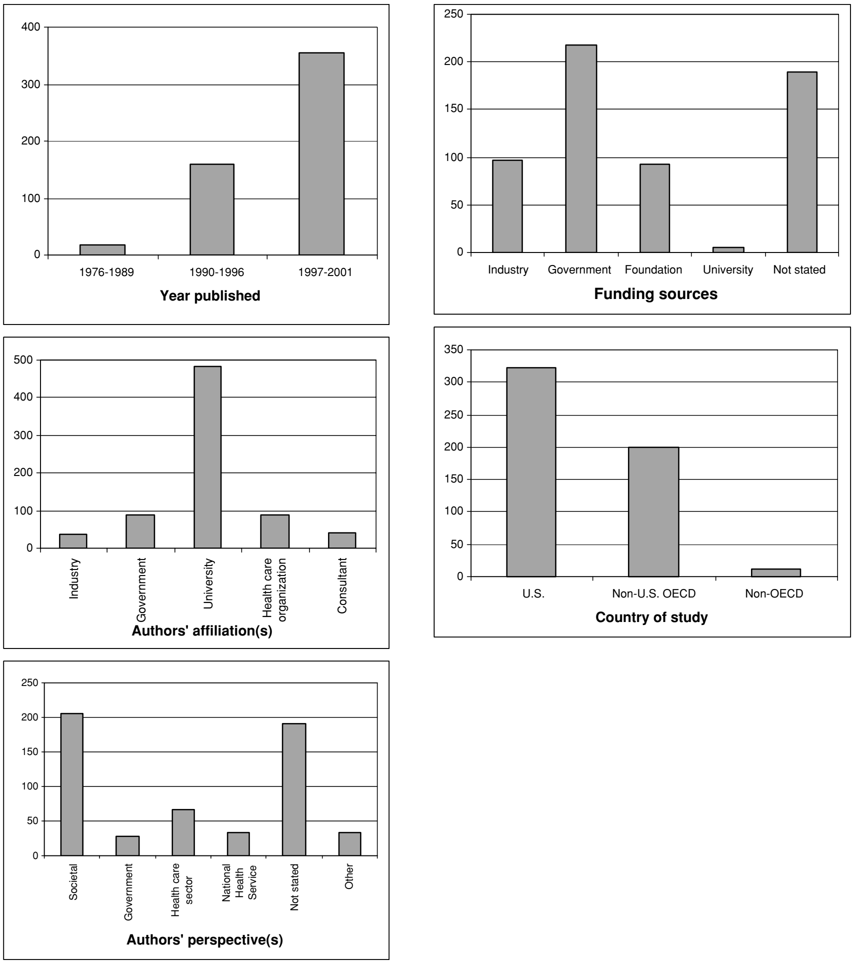

Fig. 1. Frequencies of key variables $n = 533$.

the United States, other Organization for Economic Cooperation and Development (OECD) countries, or non-OECD countries. In order to control for journal quality, we added data on each journal's annual "impact factor" over the years 1992–2001 to the main data set. Fig. 1 shows the relative frequencies of different years of publication, sponsors, affiliations, perspectives, and countries.

It is also important to control for each journal's experience and approach to CEAs. One approach would be to include journal fixed effects, which would control for the differential impact of each journal on quality. This approach, while theoretically appealing, is impractical because many journals appear only once in the data set. To avoid losing a large number of observations, we instead use a measure of journal experience with CEAs. Because the registry includes all CEAs published in peer-reviewed health and medical journals between 1976 and 2001, those journals appearing more frequently in the data have more experience in publishing CEAs. Journals publishing a large number of analyses may have editors and referees more familiar with what constitutes a high-quality analysis. Thus, we would expect journals that publish more analyses to be more adept at reviewing them and requesting relevant revisions. In fact, journals publishing and processing a large number of analyses may even have their own guidelines for performing CEA, or may specifically advocate those practices recommended by experts such as the Panel on Cost-Effectiveness in Health and Medicine.[33] In order to control for a journal's experience, we measure the journal's average annual output of CEAs each

[33] For example, in 1996, the *British Medical Journal* published guidelines; in 1995, *Annals of Internal Medicine* published guidelines; in 1994, the *New England Journal of Medicine* published an editorial calling for higher CEA quality; and in 1997, the *Journal of the American Medical Association* released a reader's guide to CEA.

year—based on the current and previous five years.[34] This number usually changes annually reflecting journals' decisions to publish more or fewer analyses over time.[35]

3. ANALYSIS AND RESULTS

Our analysis addresses two questions. First, what affects the analytical quality of a CEA, as measured by a quality index? Second, did the 1996 recommendations of the Panel on Cost-Effectiveness in Health and Medicine affect published CEAs?

3.1. Empirical Tests: Determinants of Analytical Quality

Equation (1) presents the general specification we used to test the effects of various factors on the quality of a CEA.

$$\begin{aligned} &\textit{Quality Measure}_i \\ &\quad = \beta_0 + \beta_1 \times (\textit{funding}_i) + \beta_2 \times (\textit{affiliation}_i) \\ &\qquad + \beta_3 \times (\textit{authors' perspective}_i) \\ &\qquad + \beta_4 \times (\textit{country studied}_i) \\ &\qquad + \beta_5 \times (\textit{journal impact factor}_i) \\ &\qquad + \beta_6 \times (\textit{journal experience with CEAs}_i) \\ &\qquad + \delta \times (time_i) + \varepsilon_i. \qquad (1) \end{aligned}$$

We expected that having an author take a societal perspective might improve quality; that studies focusing on the United States (and possibly other OECD countries) would be of higher quality;[36] that a journal's experience with publishing CEAs would improve quality; and that—after controlling for other factors including affiliation, funding source, journal experience, and journal quality—quality would improve with time. We were uncertain of the effect of industry funding on study quality; on the one hand, many studies find that industry studies are of lower quality and suffer from severe biases;[37] on the other hand, industry studies often enjoy generous funding budgets and have a strong economic incentive to follow protocols in order to improve their chances of getting published in reputable journals.

We estimate several versions of this equation, using different dependent variables and different variable definitions to test the robustness of the results. In total, we use nine definitions of *Quality Measure*. Our quality index is our principal measure of quality, but we perform robustness checks with eight other definitions of *Quality Measure*: each of the four individual measures of quality included in the 0–4 index; dummies for whether the study included at least two of the four measures, at least three of the four, and all four; and the subjective (1–7) index.

$\textit{Funding}_i$, $\textit{affiliation}_i$, $\textit{authors' perspective}_i$, and $\textit{country studied}_i$, are all vectors of dummy variables indicating the relevant information for study *i*. We model time three ways to test the robustness of our

34 For example, in 1999, a journal that has published three total articles between 1994 and 1999 has published an average of 0.5 CEAs per year. The variable includes analyses published in the same year. So, for example, because the *New England Journal of Medicine* published one CEA in 1983 and another three between 1976 and 1982 (four analyses over an 8-year period), the analysis published in 1983 takes on the value of 0.5 for the journal experience variable (average annual articles, to date).

We assume that CEAs published more than five years ago do not substantially influence a journal's current level of experience with CEAs. Consequently, regressions that include the journal experience variable can use data only from 1981 through 2001 (as 1976 was the initial year of data). While 5 years is an arbitrary amount of time, our results are robust to many other models of experience. These include modeling experience by the average annual number of articles published to date, measuring experience by the total number of articles published by the journal from 1976–2001, and measuring experience by the total number of articles published to date.

35 For instance, a journal may be experienced with publishing CEAs in the 1980s, have a period during which it published few analyses, and then become experienced again in the 2000s. Our journal experience variable captures the fact that experience changed. A more simplistic journal experience variable, such as the total number of CEAs published between 1976 and 2001, would not change over time and therefore fail to capture the journal's changing CEA publishing patterns.

36 We suspect that on average, U.S. studies will be of higher quality than other OECD studies. Nonetheless, because several OECD countries (such as the United Kingdom and the Netherlands, both of which have released their own panel recommendations) have been strong proponents of CEA and have been among the leaders in CEA methodology, it is possible that OECD-country studies are of comparable quality to U.S. studies. Because the authors of studies outside the United States may have fewer resources, we expect CEAs studying non-OECD countries to generally be of the lowest quality.

37 See Campbell *et al.* (2005) for a discussion of how relationships between government, universities, and industries have risks including misconduct, increased secrecy, and bias in the reporting of research results. Also, Angell (2004) expresses concerns about potential biases in industry research or industry-funded research. See also Azimi and Welch (1998), Friedberg *et al.* (1999), and Neumann *et al.* (2000a). These studies find that industry sponsorship is associated with more favorable results and with what might be interpreted as biased statements on whether interventions are cost effective. Hill *et al.* (2000), however, find "no basic intention to deceive," although they do identify that sometimes "company employees had optimistic views of their product's performance," resulting in "suboptimal and poorly designed studies."

results: using a linear time trend, a quadratic time trend, and year fixed effects.

3.2. Results: Determinants of Analytical Quality

Given our different independent variables and models of time, we ran 54 regressions—six regressions for each dependent variable. We focus here on the results from estimating Equation (1) with the quality index as the dependent variable, as it best captures the relevant quality measures. The additional regressions serve as sensitivity analyses. We run these regressions using ordinary least squares (OLS). Because the dependent variable is not strictly continuous, we also estimated regressions using different models. One might argue, for example, that the data are count data (i.e., counts of the number of factors an author included in the article). In that case, a Poisson regression may be appropriate. One might alternatively argue that the dependent variable is ordinal, suggesting an ordered probit or ordered logit. We rely on the OLS estimates given the relative ease in interpreting the results, but we also run the regressions using Poisson and ordered probit models. The qualitative results did not depend on the model used.

Table IV shows the results of estimating this equation using OLS.[38] None of the funding dummies is significantly correlated with the index, suggesting that study quality, measured by the quality index, may not depend on who funds the study. This finding is important and supports the hypothesis that authors of industry studies are interested in following protocols in order to improve their chances of getting published in reputable journals. Having at least one author affiliated with a university is positively correlated with quality. The main affiliation category excluded from the model, which serves as the base group, is a government affiliation. Thus, the results suggest that university researchers may produce higher quality CEAs than government researchers. The coefficient on industry is positive and the coefficient on health care affiliation is negative, but neither is statistically significant at conventional significance levels. While these results suggest that industry-affiliated analyses are not of lower quality than analyses with other affiliations, our indicator of quality only captures whether researchers follow basic and recommended practices. While these are important, CEAs may be of high analytical quality according to our quality index but may still suffer from biases that make the results less meaningful. Furthermore, an analysis may fail to include some recommended protocols but, nonetheless, remain useful because of the reasonableness of its assumptions and its helpfulness to decisionmakers. Our results, nonetheless, suggest that this area deserves closer scrutiny, and that authors affiliated with industry are no worse than other authors in following the basic, recommended practices captured by our quality index.[39] CEAs that focus on non-U.S. OECD countries tend to be of lower quality than CEAs that study interventions in the United States, and analyses that study non-OECD countries tend to be of even lower quality. This may reflect the fact that researchers in non-OECD countries have fewer resources. Another possibility is, because the vast majority of recommended practices originated from the United States and other OECD countries, that authors of studies outside of the United States and OECD countries have less knowledge of what procedures make for a high-quality CEA.

Having a societal perspective is positively correlated with the quality index. This result is statistically significant in all model specifications that use the authors' perspective, but is insignificant when we instead use the perspective assigned by the reader. A positive correlation between societal perspective and the quality index could reflect a more global view of the effects of a health intervention, which might necessitate a lengthier or more complete analysis.

The quality of a CEA rises with an increase in a journal's average annual output of analyses over the current and the previous 5 years. We expected this outcome given that a more experienced editorial staff and body of referees is likely to demand higher quality analyses. A journal's impact factor is also

[38] Results of all additional regressions are available upon request. The dependent variables in the additional regressions are binary variables, thus requiring limited dependent variable regression methods. We estimate those regressions with logit models and probit models. There is no *a priori* reason to believe that either a logit or probit model is more appropriate, but using probit instead of logit did not change the results. All regressions include heteroscedasticity-robust standard errors. Although it was not practical to use a model with journal fixed effects, we tried using random effects. This did not change the results; coefficients as well as standard errors were nearly identical with those presented here.

[39] Another possible explanation for the insignificance of the coefficient on industry affiliation is that industry-affiliated studies generally are of lower quality, but that journals set a higher standard for industry articles; that is, the journal may be more likely to reject an industry-funded study unless it is very high quality. This reflects an inherent problem with the data: we do not observe a journal's decisions to accept or reject an article; we only observe what a journal actually publishes.

Table IV. Determinants of Analytical Quality, Ordinary Least Squares Regression Results

Dependent Variable: Quality Index (0–4)	(1)	(2)	(3)	(4)	(5)	(6)
Industry funding	−0.161	−0.177	−0.176	−0.173	−0.187	−0.187
	(1.16)	(1.29)	(1.28)	(1.25)	(1.38)	(1.37)
Government funding	0.110	0.099	0.094	0.071	0.067	0.059
	(0.92)	(0.84)	(0.79)	(0.60)	(0.56)	(0.50)
University funding	0.106	0.076	0.080	0.154	0.147	0.151
	(0.60)	(0.45)	(0.48)	(0.66)	(0.64)	(0.67)
Foundation/society funding	0.093	0.088	0.082	0.119	0.121	0.114
	(0.85)	(0.81)	(0.75)	(1.09)	(1.10)	(1.03)
Undetermined funding source	−0.013	−0.010	−0.012	−0.038	−0.033	−0.036
	(0.10)	(0.07)	(0.08)	(0.28)	(0.24)	(0.26)
Authors' perspective is societal	0.353	0.357	0.355			
	(3.37)**	(3.43)**	(3.40)**			
Authors' perspective is government	0.070	0.107	0.076			
	(0.16)	(0.25)	(0.17)			
Authors' perspective is the NHS	0.410	0.366	0.387			
	(1.14)	(1.04)	(1.08)			
Authors' perspective is the health care sector	0.197	0.173	0.176			
	(1.28)	(1.15)	(1.17)			
Authors' perspective is "other"	−0.152	−0.139	−0.140			
	(0.83)	(0.77)	(0.78)			
Reader's perspective is societal				0.295	0.417	0.399
				(0.70)	(1.00)	(0.94)
Reader's perspective is government				0.354	0.448	0.422
				(1.32)	(1.72)†	(1.61)
Reader's perspective is the NHS				0.253	0.276	0.278
				(0.57)	(0.62)	(0.62)
Reader's perspective is the health care sector				0.215	0.339	0.323
				(0.52)	(0.82)	(0.76)
Reader's perspective is "other"				−0.401	−0.270	−0.302
				(0.89)	(0.61)	(0.67)
University affiliated	0.336	0.341	0.344	0.325	0.317	0.323
	(1.79)†	(1.80)†	(1.81)†	(1.76)†	(1.72)†	(1.74)†
Industry affiliated (includes consultants/contractors)	0.156	0.157	0.156	0.170	0.170	0.169
	(1.15)	(1.17)	(1.17)	(1.24)	(1.26)	(1.26)
Health care organization affiliated	−0.087	−0.113	−0.109	−0.079	−0.098	−0.093
	(0.70)	(0.95)	(0.91)	(0.63)	(0.81)	(0.77)
Non-U.S. OECD country studied	−0.277	−0.278	−0.279	−0.329	−0.319	−0.322
	(2.71)**	(2.70)**	(2.71)**	(3.06)**	(2.94)**	(2.96)**
Non-OECD country studied	−0.845	−0.847	−0.852	−0.876	−0.869	−0.878
	(1.92)†	(1.96)†	(1.93)†	(2.08)*	(2.08)*	(2.06)*
Journal impact factor	0.017	0.014	0.014	0.020	0.018	0.018
	(1.59)	(1.42)	(1.39)	(1.98)*	(1.85)†	(1.81)†
Analyses published, current and past 5 years	0.121	0.131	0.133	0.132	0.140	0.142
	(3.57)**	(3.98)**	(4.04)**	(3.85)**	(4.20)**	(4.26)**
Time trend		−0.008	−0.362		0.013	−0.409
		(0.25)	(1.00)		(0.40)	(1.09)
Time trend, squared			0.008			0.009
			(0.98)			(1.14)
Analysis was published in 1997 or after		0.130	0.209		0.084	0.180
		(0.70)	(1.01)		(0.45)	(0.84)
Constant	3.188	2.789	6.798	3.002	2.202	6.999
	(8.52)**	(4.15)**	(1.64)	(5.31)**	(2.61)**	(1.58)
Observations	431	431	431	431	431	431
R^2	0.21	0.20	0.20	0.20	0.19	0.19

†Significant at 10%, * significant at 5%, ** significant at 1%.
Notes: The 0–4 index includes: performing incremental analysis, discounting costs and quality-adjusted life years, stating the year of currency used in cost estimates, and performing sensitivity analysis.
Absolute value of *t* statistics in parentheses; all regressions include heteroskedasticity-robust standard errors; all models that do not include a time trend were estimated with year fixed effects.

positively correlated with the quality rating of a CEA published in it, but this result is not significant in all model specifications.

None of the coefficients on the time dummies is significant, nor are the coefficients on any of the time trend variables. This result suggests that the quality of analyses has not improved over time, at least when we describe quality with our index. The additional regressions reveal that the only measure of quality that may have changed over time is whether the analysis discounts costs and quality-adjusted life years. In all models that include a quadratic time trend, an analysis is more likely to discount costs and quality-adjusted life years after the 1996 recommendations than before. This result did not hold, however, when we instead used a linear time trend. Consequently, the results offer little support for the hypothesis that CEAs improved on our quality index following the 1996 panel recommendations. While these results suggest that the 1996 panel recommendations did not affect the quality index, they may have affected the extent to which analyses include other panel recommendations. To better answer this question, we turn to our second model, which includes the components of the previous index in addition to a number of more subjective recommendations that we did not include in the quality index.

3.3. Empirical Tests: Panel Impact on Adherence to Recommended Protocols

Equation (2) presents the general specification we used to test whether the 1996 panel recommendations affected the extent to which CEAs follow recommended protocols.

$$\begin{aligned} &\mathit{Compliance}_i \\ &\quad = \beta_0 + \beta_1 \times (\mathit{funding}_i) + \beta_2 \times (\mathit{affiliation}_i) \\ &\qquad + \beta_3 \times (\mathit{country\ studied}_i) \\ &\qquad + \beta_4 \times (\mathit{journal\ experience\ with\ CEAs}_i) \\ &\qquad + \beta_5 \times (\mathit{journal\ impact\ factor}_i) \\ &\qquad + \delta \times (\mathit{time}_i) + \varepsilon_i. \end{aligned} \tag{2}$$

We estimate several versions of Equation (2), using a somewhat different set of dependent variables and different variable definitions to test the robustness of the results. We also model time in several different ways.

Our dependent variable, *Compliance*, is a measure of the extent to which a CEA incorporates the recommendations of the Panel on Cost-Effectiveness in Health and Medicine. The recommendations on which we focus are the eight included in our panel index. In total, we use 11 definitions of *Compliance*: two versions of our panel index[40] and each of the eight individual measures of compliance included in that index.[41]

The right-hand side of Equation (2) is similar to Equation (1) except that it does not include authors' perspective[42] and we model time in a few additional ways. As in Equation (1), we model time in three ways to test the robustness of our results: using a linear time trend, a quadratic time trend, and year fixed effects. To examine the effects of the 1996 Panel on Cost-Effectiveness in Health and Medicine, we also include a dummy variable for being published after the panel recommendations diffused. It is uncertain how quickly the recommendations diffused. This depends on a number of uncertain factors, including the length of time between writing an article and publishing it and how quickly the panel recommendations became common knowledge.

To test the robustness of our results to changes in the amount of time it took these recommendations to become common knowledge, we used three different time dummies with each of our linear and quadratic time trends: after 1996, after 1997, and after 1998. The purpose of the original time trend is to de-trend the data and to test whether quality has changed over time. The purpose of the time dummies is to specifically test whether the panel recommendations had an impact on CEA quality that is different from the impact of time alone. By modeling time in several different ways, we rigorously test the robustness of any conclusions regarding the effects of time and the panel recommendations on CEA adherence to recommended protocols.

Again, $\mathit{funding}_i$, $\mathit{affiliation}_i$, *and* $\mathit{country\ studied}_i$, are all vectors of dummy variables indicating the relevant information for study *i*. We estimate the model with OLS when we use the panel index, although it is worth noting that our results did not qualitatively change when we instead used ordered probit and Poisson regression models. All other models are

40 One version uses the reader's perspective and the other uses the authors' perspective as the determinant of the study perspective.

41 Societal perspective appears twice—once for the authors' perspective being societal, and once for the reader's perspective being societal.

42 We removed authors' perspective because the panel recommended that authors take a societal perspective and this variable thus becomes one of our measures of quality and is incorporated into the new index.

Table V. Panel Impact on Compliance with Panel Recommendations, Ordinary Least Squares Regression Results

Dependent Variable: Panel Index (0–8)	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Industry funding	−0.105	−0.152	−0.156	−0.133	−0.132	−0.154	−0.161
	(0.46)	(0.68)	(0.69)	(0.59)	(0.58)	(0.70)	(0.72)
Government funding	0.290	0.238	0.233	0.260	0.249	0.240	0.227
	(1.48)	(1.21)	(1.17)	(1.32)	(1.25)	(1.24)	(1.16)
University funding	0.174	0.096	0.089	0.103	0.111	0.095	0.090
	(0.30)	(0.17)	(0.16)	(0.17)	(0.19)	(0.16)	(0.15)
Foundation/society funding	0.117	0.119	0.115	0.128	0.115	0.133	0.124
	(0.62)	(0.64)	(0.61)	(0.68)	(0.61)	(0.71)	(0.66)
Undetermined funding source	0.046	0.059	0.054	0.051	0.049	0.042	0.039
	(0.21)	(0.26)	(0.24)	(0.23)	(0.22)	(0.19)	(0.18)
University affiliated	0.106	0.123	0.122	0.098	0.105	0.117	0.124
	(0.41)	(0.46)	(0.46)	(0.38)	(0.40)	(0.45)	(0.47)
Industry affiliated (includes consultants/contractors)	0.045	0.066	0.066	0.045	0.043	0.068	0.073
	(0.20)	(0.28)	(0.28)	(0.19)	(0.19)	(0.29)	(0.31)
Health care organization affiliated	−0.286	−0.283	−0.280	−0.307	−0.298	−0.264	−0.254
	(1.48)	(1.50)	(1.47)	(1.60)	(1.55)	(1.39)	(1.33)
Non-U.S. OECD country studied	−0.715	−0.751	−0.754	−0.745	−0.746	−0.760	−0.764
	(4.69)**	(4.92)**	(4.93)**	(4.89)**	(4.91)**	(5.02)**	(5.06)**
Non-OECD country studied	−1.054	−1.140	−1.142	−1.083	−1.100	−1.106	−1.127
	(1.21)	(1.31)	(1.31)	(1.25)	(1.25)	(1.30)	(1.30)
Journal impact factor	0.068	0.066	0.066	0.065	0.064	0.067	0.067
	(4.05)**	(3.90)**	(3.89)**	(3.86)**	(3.81)**	(4.03)**	(4.01)**
Average analyses published, current and past 5 years	0.333	0.357	0.358	0.357	0.360	0.349	0.350
	(4.99)**	(5.32)**	(5.31)**	(5.37)**	(5.36)**	(5.20)**	(5.19)**
Time trend		0.213	0.016	0.158	−0.546	0.123	−0.231
		(5.08)**	(0.03)	(3.19)**	(1.14)	(2.47)*	(0.51)
Time trend, squared			0.004		0.015		0.008
			(0.30)		(1.45)		(0.76)
Analysis was published in 1996 or after		−0.200	−0.136				
		(0.79)	(0.40)				
Analysis was published in 1997 or after				0.206	0.361		
				(0.78)	(1.33)		
Analysis was published in 1998 or after						0.376	0.373
						(1.50)	(1.49)
Constant	3.291	−0.897	1.335	0.073	8.049	0.786	4.716
	(6.02)**	(0.98)	(0.18)	(0.07)	(1.51)	(0.75)	(0.94)
Observations	431	431	431	431	431	431	431
R^2	0.35	0.34	0.34	0.34	0.34	0.34	0.34

* Significant at 5%, ** significant at 1%.
Notes: The 0–8 index includes the variables in the 0–4 index plus performing probabilistic sensitivity analysis, model validation, using a 3% discount rate, and taking a societal perspective (authors' perspective).
Absolute value of *t* statistics in parentheses; all models that do not include a time trend were estimated with year fixed effects; all standard errors are robust to heteroskedasticity.

logit models because they contain binary dependent variables.[43]

3.4. Results: Panel Impact on Adherence to Recommended Protocols

Our various independent variables and models of time led us to run 55 regressions—five regressions for each of the 11 dependent variables. Table V shows the results of estimating Equation (2) where the panel index serves as the dependent variable.[44] We find some evidence that *Compliance* has improved over time, but no robust evidence that the panel guidelines, *per se*, directly affected our panel index. Time is significant at about the 1% level and has a positive coefficient

[43] We also ran probit models.

[44] The authors' perspective is used to indicate the study's perspective.

when the equation is estimated using a linear time trend. Time is insignificant when we use a quadratic time trend, although this may be due to multicollinearity, as time and time-squared are highly correlated. Whether we assume that the panel recommendations diffused rapidly or slowly, we found no evidence that they had a significant impact on our panel index. The coefficients on the dummy for 1997, the dummy for 1998, and the dummy for 1999 were insignificant in all regressions.

To explore this last result further, we estimated Equation (2) with each of our multiple definitions of *Compliance*. In none of these regressions did the panel recommendations have a clear impact on *Compliance*. The result on our panel index model suggest that studies improve over time in some respect, but the fact that time is significant in this regression but not in the regressions involving the quality index suggests that improvements have occurred on the more subjective recommendations (such as using a societal perspective and a 3% discount rate) but not on the measures in the quality index. This result is confirmed by the additional regressions in which societal perspective and the dummy for having a 3% discount rate were the dependent variables. In these models, time was significant when we used a linear time trend, regardless of our assumptions about the number of years it would take the recommendations to diffuse.

These results are not surprising. In many ways, the 1996 Panel on Cost-Effectiveness in Health and Medicine simply summarized and synthesized recommendations that had been around for some time.[45] In this case, our findings simply fail to indicate any added value of the panel recommendations in terms of CEA compliance with recommended protocols. Consensus statements, such as the panel recommendations, may simply reflect conventional practice rather than truly shape it.

4. CONCLUSION

Economic analysis of life-saving policy interventions has the potential to inform decisionmakers and improve the allocation of scarce public and private resources. These objectives can be achieved only if such analyses are done well.

In this article, we analyze a data set that provides information on the content and quality of journal articles that measure the cost effectiveness of life-saving investments. Our study is the first to provide a detailed multivariate analysis of factors affecting objective measures of quality. We also explore whether the recommendations by the expert panel convened by the U.S. Public Health Service affect the way CEAs for life-saving investments are done.

The analysis yields three key results. First, quality has not improved over time when measured by the quality index, but has improved over time when measured by the panel index of recommendations. The latter index includes subjective protocols, such as the particular discount rate to use and the perspective to take in conducting a study. That CEAs did not improve on the quality index over time is somewhat surprising. Yet, it is consistent with other studies of government regulatory impact analyses. One possible explanation for this result is that our approach to measuring quality has serious limitations. For example, our results do not measure the extent to which study quality has improved due to more sophisticated statistical approaches, enabled by the recent development of advanced statistical analysis programs and software packages.

Second, it is not clear that guidelines had much effect on analytical quality. Again, these results obtain both for CEAs and regulatory impact analyses.[46] A critical question this finding raises is whether there are conditions under which guidelines are likely to be more effective in improving analysis. In the federal government, we think that it is critical to have an enforcement mechanism. In the case of peer-reviewed work, the enforcement mechanism is less clear if referees and editors do not require it. In addition, peer review is likely to be effective only if reviewers and editors take their roles seriously and apply rigorous standards.

The third result, which was something of a surprise, relates to the impact of funding sources. We found that the source of study funding is not significantly correlated with the quality of CEAs as measured by the quality index. One possible explanation for our results is that it was not intended to pick up subtle forms of bias, but we still think it has relevance because it shows that on some general objective measure of quality, funding does not appear to make a difference. There is an important ongoing debate in the literature about conflicts of interest, the effect of disclosure, and the impact of funding sources. Our result should be considered in conjunction with other findings in devising sensible policies related to possible conflicts of interest.

45 See, e.g., Drummond *et al.* (1987). This is a widely used textbook on CEA. In addition, other consensus groups have also released guidelines (in Canada, Australia, and later Europe).

46 Hahn and Dudley (2004) show that these results obtain for regulatory impact analyses.

Social and private life-saving interventions represent a significant economic investment. Moreover, they are likely to become more important in magnitude as the demand for health-related services and their costs increase. We need to learn more about the factors that drive this investment, including any underlying analysis, if we wish to substantially improve it.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the assistance of James Fraumeni and Jenny Palmer and the comments of Emmett Keeler, Tim McDaniels, James Prieger, and two anonymous referees. This research was supported by the AEI-Brookings Joint Center.

REFERENCES

Abramowicz, M. (2004). Information markets, administrative decision-making, and predictive cost-benefit analysis. *University of Chicago Law Review*, *71*, 933–1015.

Angell, M. (2004). *The Truth about Drug Companies: How They Deceive Us and What to Do about It*. New York: Random House.

Arrow, K. J., Cropper, M. L., Eads, G. C., Hahn, R. W., Lave, L. B., Noll, R. G., Portney, P. R., Russell, M., Schmalensee, R., Smith, V. K., & Stavins, R. N. (1996). Is there a role for benefit–cost analysis in environmental, health, and safety regulation? *Science*, *272*, 221–222.

Azimi, N. A., & Welch, H. G. (1998). The effectiveness of cost-effectiveness analysis in containing costs. *Journal of General Internal Medicine*, *13*(10), 664–669.

Blair, T. (2005). Risk and the state. In *Policy Matters* (pp. 5–15). Washington, DC: AEI-Brookings Joint Center.

Campbell, E. G., Koski, G., & Blumenthal, D. (2005). Managing the triple helix in the life sciences. *Issues in Science and Technology*, *1*, 32–39.

Chapman, R., Stone, P., Sandberg, E., Bell, C., & Neumaan, P. (2000). A comprehensive league table of cost-utility ratios and a sub-table of "panel-worthy" studies. *Medical Decision-Making*, *20*(4), 451–467.

Chiou, C.-F., Hay, J. W., Wallace, J. W., Bloom, B. S., Neumann, P. J., Sullivan, S. D., Yu, H.-T., Keeler, E. B., Henning, J. M., & Ofman, J. J. (2003). Development and validation of a grading system for the quality of cost-effectiveness studies. *Medical Care*, *41*(1), 32–44.

Drummond, M. F., Stoddart, G. L., & Torrance, G. W. (1987). *Methods for the Economic Evaluation of Health Care Programmes*. Oxford: Oxford University Press.

Farrow, S. (1991). Does analysis matter? Economics and planning in the Department of the Interior. *Review of Economic Statistics*, *73*(1), 172–176.

Friedberg, M., Saffran, B., Stinson, T., Nelson, W., & Bennett, C. L. (1999). Evaluation of conflict of interest in economic analyses of new drugs used in oncology. *Journal of the American Medical Association*, *282*(15), 1453–1457.

Gerard, K., Smoker, I., & Seymour, J. (1999). Raising the quality of cost-utility analyses: Lessons learnt and still to learn. *Health Policy*, *46*(3), 217–238.

Gold, M. R., Siegel, J. E., Russell, L. B., & Weinstein, M. C. (1996). *Cost-Effectiveness in Health and Medicine*. Oxford: Oxford University Press.

Hahn, R., Burnett, J., Chan, Y.-H. I., Mader, E., & Moyle, P. (2000). Assessing regulatory impact analyses: The failure of agencies to comply with Executive Order 12,866. *Harvard Journal of Law and Public Policy*, *23*, 859–871.

Hahn, R., & Dudley, P. (2004). *How Well Does the Government Do Cost-Benefit Analysis*? Working Paper 04-01, AEI-Brookings Joint Center.

Harrington, W., Morgenstern, R., & Nelson, P. (2000). On the accuracy of regulatory cost estimates. *Journal of Policy Analysis and Management*, *19*(2), 297–332.

Hill, S. R., Mitchell, A. S., & Henry, D. A. (2000). Problems with the interpretation of pharmacoeconomic analyses. *Journal of the American Medical Association*, *283*(16), 2116–2121.

Lave, L. (1981). *The Strategy of Social Regulation: Decision Frameworks for Policy*. Washington, DC: Brookings Institution.

Litan, R. E., & Nordhaus, W. D. (1983). *Reforming Federal Regulation*. New Haven: Yale University Press.

Morgenstern, R. D. (Ed.). (1997). *Economic Analysis at EPA: Assessing Regulatory Impact*. Washington, DC: Resources for the Future.

Morrall, J. F. (2003). Saving lives: A review of the record. *Journal of Risk and Uncertainty*, *27*, 221–237.

Murphy, K. M., & Topel, R. H. (2003). The economic value of medical research. In K. M. Murphy & R. H. Toppel (Eds.), *Measuring the Gains from Medical Research: An Economic Approach* (pp. 41–73). Chicago: University of Chicago Press.

Neumann, P. J., Greenberg, D., Olchanski, N., Stone, P., & Rosen, A. (2005). Growth and quality of the cost-utility literature, 1976–2001. *Value in Health*, *8*(1), 3–9.

Neumann, P. J., Sandberg, E. A., Bell, C. M., Stone, P. W., & Chapman, R. H. (2000a). Are pharmaceuticals cost-effective? A review of the evidence. *Health Affairs*, *19*(2), 92–109.

Neumann, P. J., Stone, P. W., Chapman, R. H., Sandberg, E. A., & Bell, C. M. (2000b). The quality of reporting in published cost-utility analyses, 1976–1997. *Annals of Internal Medicine*, *132*(12), 964–972.

Nichols, A. L. (1997). Case study: The RIA for the 1985 lead-in-gasoline phasedown. In R. D. Morgenstern (Ed.), *Economic Analysis at EPA: Assessing Regulatory Impact* (pp. 49–80). Washington, DC: Resources for the Future.

Nichols, A. L. (1997), Lead in gasoline. In R. D. Morgenstern (Ed.), *Economic Analysis at EPA: Assessing Regulatory Impact*. Washington, DC: Resources for the Future.

Office of Management and Budget. (2003). *Circular A-4: Regulatory Analysis*. Washington, DC: Office of Management and Budget.

Russell, L., Gold, M., Siegel, J., Daniels, N., & Weinstein, M. (1996). The role of cost-effectiveness analysis in health and medicine. *Journal of the American Medical Association*, *276*(14), 1172–1177.

Schmalensee, R., Joskow, P., & Bailey, E. M. (1998). The market for sulfur dioxide emissions. *American Economic Review*, *88*(4), 669–685.

Smith, V. K. (Ed.). (1984). *Environmental Policy under Reagan's Executive Order: The Role of Benefit-Cost Analysis*. Chapel Hill, NC: University of North Carolina Press.

Stavins, R. N. (1998). What have we learned from the grand policy experiment: Lessons from SO2 allowance trading. *Journal of Economic Perspectives*, *12*(3), 69–88.

Sunstein, C. R. (2002). *Risk and Reason*. Cambridge: Cambridge University Press.

Tengs, T., Adams, M., Pliskin, J., Safran, D., Siegel, J., Weinstein, M., & Graham, J. (1995). Five-hundred life-saving interventions and their cost-effectiveness. *Risk Analysis*, *15*(3), 369–390.

Tufts University-New England Medical Center. (2005). *Cost-Effectiveness Analysis Registry*. Available at http://www.hsph.harvard.edu/cearegistry/.

Udvarhelyi, S., Golditz, G. A., Rai, A., & Epstein, A. M. (1992). Cost-effectiveness and cost-benefit analyses in the medical literature. *Annals of Internal Medicine*, *116*, 238–244.

Viscusi, K. W., & Hamilton, J. T. (1999). Are risk regulators rational. *American Economic Review*, *89*, 1010–1027.

Viscusi, W. K. 1996. Economic foundations of the current regulatory reform efforts. *Journal of Economic Perspectives*, *10*(3), 119–134.