

ARTICLES

Evaluating the Social Effects of Environmental Leadership Programs

by Jonathan C. Borck, Cary Coglianese, and Jennifer Nash

Editors' Summary: In the past decade, EPA and over 20 states have created voluntary environmental leadership programs designed to recognize and reward businesses that take steps that go beyond compliance with the strictures of environmental law. Environmental leadership programs seek not only to spur direct improvements to environment quality but also to advance broader social goals that may lead indirectly to environmental improvements, such as improving business-government relationships and changing business culture. Measuring progress toward leadership programs' social goals is a particularly challenging but essential task if researchers and decisionmakers are to understand the full impacts of these programs. In this Article, Jonathan C. Borck, Cary Coglianese, and Jennifer Nash present strategies for overcoming the three core challenges in evaluating the social effects of leadership programs and any voluntary environmental initiative: (1) defining appropriate measures of social goals; (2) inferring whether programs achieve those goals; and (3) linking social effects to environmental outcomes. Only through careful attention to these three empirical issues will it be possible to rule out alternative explanations and determine whether environmental leadership programs are truly generating their intended positive social effects as well as improvements to the environment.

Over the past decade, the U.S. Environmental Protection Agency (EPA) and more than 20 states have established environmental leadership programs (ELPs), a type of voluntary environmental program with the explicit goal of improving the environmental performance of private-sector facilities.¹ EPA's National Environmental Performance Track, considered by many to be the nation's flagship ELP, now boasts more than 500 member facilities that have voluntarily implemented environmental management systems and set goals for their environmental performance that

go beyond meeting legal requirements. Similar state programs include the Partnership for a Sustainable Georgia, Tennessee's Pollution Prevention Partnership, and Virginia's Environmental Excellence Program.

As a condition for the recognition and rewards that come with membership in ELPs, facilities must submit application materials to the government demonstrating that they meet specified entry criteria. As part of their applications, facilities show that they have implemented systematic management practices and are committed to improving their environmental performance in areas such as water and energy conservation and habitat protection. Many programs also require businesses to establish community outreach programs as a condition for membership, and nearly all programs require facilities to inform the public on a regular basis about the environmental impacts of their operations.²

Although ELPs primarily seek to promote better environmental performance by industry,³ they also have been designed to advance a variety of social goals, such as improv-

Jonathan Borck is an associate with the Analysis Group, Inc., Boston, Massachusetts. Cary Coglianese is the Edward B. Shils Professor of Law at the University of Pennsylvania Law School, Philadelphia, Pennsylvania. Jennifer Nash is the director of the Regulatory Policy Program at the Harvard Kennedy School, Cambridge, Massachusetts. The authors thank participants in the workshop, "Performance-Based Environmental Programs: Better Ways to Measure and Communicate Results," sponsored by the U.S. Environmental Protection Agency (EPA), the Multi-State Working Group on Environmental Performance, and the Environmental Council of the States, and held at the Harvard Kennedy School, Harvard University, Cambridge, Massachusetts, on March 11-12, 2008. We also received insightful comments from Angela Helman and Eric Ruder. Although the research described in this Article has been funded in part by EPA contract EP-W-05-047, it has not been subject to the Agency's review and therefore does not necessarily reflect the views of the Agency. No official endorsement should be inferred.

1. Jonathan C. Borck et al., *Environmental Leadership Programs: Toward an Empirical Assessment of Their Performance*, ECOLOGY L.Q. (forthcoming).

2. *Id.*

3. See, e.g., ENVIRONMENTAL COUNCIL OF THE STATES (ECOS), SURVEY OF STATE SUPPORT FOR PERFORMANCE-BASED ENVIRONMENTAL PROGRAMS AND RECOMMENDATIONS FOR IMPROVED EFFECTIVENESS 1, 4 (2005), available at http://www.epa.gov/perfrac/downloads/ECOS_Report_Final_01-13-05.pdf ("The fundamental goal of Performance Track and other state innovations programs is to achieve better environmental results.").

ing relationships between businesses, communities, and government agencies, and diffusing a more responsible ethos throughout corporate culture.⁴ In creating its Performance Track program, for example, EPA aspired to “transform” its relationships with industry, making them more “collaborative, cooperative, and focused on results.”⁵ Performance Track seeks to do so by moving beyond the traditional wielding of negative sanctions against rule-breakers, and instead by having EPA recognize and reward businesses that engage in responsible environmental conduct.⁶ In addition, by requiring that Performance Track members engage in community outreach, EPA hopes to spur increased confidence and trust by local organizations and citizens in their business neighbors.⁷ EPA Administrator Stephen Johnson has declared that Performance Track’s positive approach is “fundamentally strengthening the relationship between business and government.”⁸

State governments have similarly established environmental leadership programs with social goals in mind. The Environmental Council of the States (ECOS) has declared that these programs can help “foster greater collaboration between environmental regulatory agencies and high-performing companies.”⁹ At two recent workshops about ELPs organized by EPA, ECOS, and the Multi-State Working Group on Environmental Performance (MSWG), participants spoke extensively about the importance of improving relationships, enhancing trust, and changing culture.¹⁰ In interviews we conducted with state ELP manag-

ers, officials voiced support for the goals of improving relationships and changing the culture of businesses and government agencies.¹¹

Businesses also give priority to ELPs’ social goals. According to membership surveys that EPA conducted in 2004 and 2006, participants in Performance Track reported that social effects provide much of the value they receive from participation. In response to the 2004 survey, for example, members reported that their most important reason for joining and staying in the program was to enjoy a “collaborative/amicable relationship with EPA [and s]tates.”¹² The 2006 survey showed that most respondents believed that Performance Track participation had contributed to a “culture of continuous environmental improvement” and had improved their facilities’ “relationship with EPA.”¹³ Of course, the primacy that ELP members place on social goals should not be surprising, given that ELPs are voluntary programs. Businesses that choose to join are those willing to reach out to the government and those that find beneficial the opportunities the programs afford for engaging with representatives of agencies and community organizations.¹⁴

In bringing facility managers, community residents, environmental advocacy organizations, and government environmental agencies into closer contact through meetings and other mechanisms for information sharing, ELPs clearly aim to transform traditional adversarial relationships into more cooperative modes of interacting. The implications of such a shift, if achieved, could be substantial, as new perspectives and best practices could begin to permeate facility walls and transform both business and government agencies.¹⁵ It is also conceivable, though, that ELPs could negatively affect relationships and the regulatory culture. This could happen if ELPs raise expectations for cooperation that are not met, or if repeated interactions provide greater opportunities for disagreement or miscommunication, with the possible unintended result of increasing levels of mistrust among stakeholders. Especially if agencies promise benefits to participants that they fail to deliver, or if facilities fail to live up to the commitments they make, relationships could be damaged rather than strengthened. It is also conceivable that most programs’ commitments not to subject members to routine regulatory inspections could raise public suspicions and exacerbate public mistrust of both agencies and facilities. For example, when EPA moved to bolster benefits for Performance Track participants in 2005, environmental groups complained that the Agency

4. See, e.g., DANIEL J. FIORINO, *THE NEW ENVIRONMENTAL REGULATION 206* (2006) (stating that “a principal justification for . . . beyond-compliance programs like EPA’s Performance Track” is “to build a foundation for increased trust among the actors in the regulatory system”). For a discussion of how environmental management systems—a core requirement for membership in almost any ELP—can promote social goals, see William R. Moomaw, *Expanding the Concept of Environmental Management Systems to Meet Multiple Social Goals*, in *REGULATING FROM THE INSIDE: CAN ENVIRONMENTAL MANAGEMENT SYSTEMS ACHIEVE POLICY GOALS?* 126-45 (Cary Coglianese & Jennifer Nash eds., 2001).

5. U.S. EPA, *BUILDING ON THE FOUNDATION: PERFORMANCE TRACK SECOND ANNUAL PROGRESS REPORT 3* (2004), available at http://www.epa.gov/perfrac/PT_2nd_progress_rpt_FINAL.pdf. See also U.S. EPA, *BUILDING TRUST WITH PERFORMANCE*, available at http://www.epa.gov/perfrac/members/downloads/PTwaterincentives_reducedmonitoringpaper.pdf (noting that Performance Track “was designed in part to foster greater cooperation between regulated facilities and their state and federal regulators”).

6. See, e.g., Daniel J. Fiorino, *Performance Track Places Trust in the Carrot Over the Stick*, 10 *ENVTL. QUALITY MGMT.* 9 (2001).

7. As an employee of a Performance Track member has asserted, the program “establishes a climate of respect and trust with the community. It keeps the people who live near you comfortable that you are not polluting.” U.S. EPA, *LEADING CHANGE: PERFORMANCE TRACK FOURTH ANNUAL PROGRESS REPORT 22* (2006), available at http://www.epa.gov/perfrac/downloads/PT_4th_Progress_Report.pdf.

8. U.S. EPA, *Performance Track Celebrates Five Years of Environmental Leadership*, <http://www.epa.gov/perfrac/5thAnniversary.htm> (quoting Administrator Johnson).

9. ECOS, *supra* note 3, at 7.

10. Angela Vituli & Eric Ruder, *Summary of the May 8th Dialogue on Performance-Based Environmental Programs* (2007) (unpublished manuscript on file with the authors); Hetal Jain, MSWG’s “International Dialogue on Ecological Policy”: Notes on Dialogue on Defining, Measuring, and Communicating Results of Performance-Based Programs (2007) (unpublished manuscript on file with the authors); Eric Ruder, *State and Federal Performance-Based Environmental Programs: Assessing the Potential for Evaluation—State-*

ment of Research Questions and Program Goals (2007) (unpublished manuscript on file with the authors).

11. Borck et al., *supra* note 1.

12. Abt Associates Inc., *Results of 2004 Performance Track Customer Satisfaction Survey*, http://www.epa.gov/performance-track/members/news/mar05/survey_report.pdf (last visited Feb. 1, 2008).

13. U.S. EPA, *2006 PERFORMANCE TRACK SURVEY: SUPPLEMENTAL QUESTIONS, PRELIMINARY RESULTS* (2006).

14. See Jennifer A. Howard-Grenville et al., *Constructing the License to Operate: Internal Factors and Their Influence on Corporate Environmental Decisions*, 30 *LAW & POL’Y* 73 (2008) (reporting evidence that participants in Performance Track are more extroverted than non-participants).

15. Andrew A. King, *The Role of Management Systems in Stakeholder Partnerships*, in *LEVERAGING THE PRIVATE SECTOR: MANAGEMENT-BASED STRATEGIES FOR IMPROVING ENVIRONMENTAL PERFORMANCE 228-45* (Cary Coglianese & Jennifer Nash eds., 2006); IAN AYRES & JOHN BRAITHWAITE, *RESPONSIVE REGULATION: TRANSCENDING THE DEREGULATION DEBATE* (1995).

was improperly trading away needed regulations and enforcement authority to induce facilities to join.¹⁶

The possibility that ELPs could either succeed or fail to improve regulatory relationships and culture motivates the question we address in this Article: How can agencies (or anyone) measure and evaluate ELPs' social contributions? Agencies have put a great deal of work into developing environmental performance metrics for ELPs, collecting environmental performance data from member facilities and compiling data into summary reports purporting to show these programs' direct contributions to environmental protection.¹⁷ Much less attention has focused on establishing social impact measures and collecting social performance data.¹⁸ Therefore, this Article outlines what would be needed to demonstrate with confidence that ELPs achieve their social goals successfully. While we focus on the social goals of improving relationships and regulatory culture, the basic issues we raise apply equally to evaluating the contributions of these programs to achieving any other goal, whether it be achieving traditional environmental improvements, integrating environmental concerns into business decisionmaking, or realizing other policy objectives.

I. Measurement Issues: Defining and Operationalizing Social Goals

An initial hurdle in evaluating ELPs' progress toward social goals is to clarify what is meant by a "social goal" and develop appropriate performance measures. While environmental goals are themselves not always easy to define, for the most part they are based on the consequences of economic inputs and can be operationalized using outputs that can be isolated, measured, and tracked, such as emissions or energy usage. But what, exactly, does it mean to "improve relationships" among facilities, agencies, and surrounding communities? Or to "change culture" or "enhance trust"? Investigators cannot directly observe these phenomena—no "culture-o-meter" or "trust-o-meter" exists. In these cases, investigators must identify proxy variables that are correlated or associated with the underlying social variables of concern.

Proxy variables can fall into one of two categories: (1) revealed proxies; and (2) expressed proxies. Revealed proxies are measures of actual observable behavior consistent with the underlying social variables of concern. Revealed proxies are closely related to the concept of revealed preferences in economics: they are measures of real-world behavior that reflect the preferences and attitudes of the actors.¹⁹ Examples of revealed proxies for improving relationships might be the number of complaints community residents make to or about an industrial facility or the number of lawsuits filed by groups against polluters.

In contrast, expressed proxies are measures of actors' stated opinions. Expressed proxies are closely related to the concept of stated preferences in economics: they are measures of actors' preferences determined by their words, not their actions. An example of a stated proxy for improving relationships might be the responses to a survey question asking environmental or community organizations how much they trust industry or government regulators.²⁰

The problems with revealed proxies are that they are limited in availability and may be more closely related to factors other than the underlying variable of concern. The range of observable behavior is naturally limited. Businesses, regulators, and environmental and community groups engage in only a few observable actions plausibly related to an underlying social effect. And when they do, these actions are necessarily indirect measures of the underlying social variables of concern. Investigators must be aware of the possibility that the variation in a specific revealed proxy is related less to the underlying social effect than to some other factor or influence. For example, the number of lawsuits filed by environmental groups against polluters could be explained more (or entirely) by changes in groups' finances than by changes in underlying levels of trust or improvements in relationships. Investigators do not need to abandon such revealed proxies. Rather, they can use statistical techniques to account for these alternative explanations. But the issue certainly complicates efforts to identify the effects of ELPs.

Unlike revealed proxies, which are indirect and limited in availability, expressed proxies are readily obtainable and can be quite direct: investigators can ask directly about any underlying social variable of concern. But they simply cannot be sure that the responses they receive are accurate reflections of the true feelings or views of the respondents. Opinions obtained through surveys or interviews are prone to numerous biases that have been well documented in various literatures.²¹ For example, respondents can be swayed by the range of options in survey questions or by the first option provided in an interview. Respondents might respond strategically to a survey question to influence a perceived outcome. Respondents can also be influenced by the amount and nature of background information provided in a survey or interview or by external factors, such as the weather on the day they answer the survey or conduct the interview, that are entirely irrelevant to the questions being asked. Researchers are actively developing methods to elicit more truthful (or less biased) responses to survey or interview questions, and techniques have improved immensely from the earliest days of survey research.²² But implementing these improved methods to collect expressed proxies is both time-consuming and expensive.

Another, more fundamental challenge with both revealed and expressed proxies is that being that they are proxies, the

16. Letter from John Walke, National Resources Defense Council, to U.S. EPA Docket ID OA-2005-0003 (Nov. 3, 2005); Letter from Eric Schaeffer et al., President, Environmental Integrity Project, to Stephen Johnson, EPA Administrator (Jan. 25, 2006), available at http://www.environmentalintegrity.org/pubs/performance_track_letter_jan06.pdf.

17. Borck et al., *supra* note 1.

18. *Id.*

19. CHARLES KOLSTAD, ENVIRONMENTAL ECONOMICS 297 (2000); U.S. EPA, *Guidelines for Preparing Economics Analyses* §7.5, [http://yosemite1.epa.gov/ee/epa/ee/mfile.nsf/vwAN/EE-0228C-07.pdf/\\$File/EE-0228C-07.pdf](http://yosemite1.epa.gov/ee/epa/ee/mfile.nsf/vwAN/EE-0228C-07.pdf/$File/EE-0228C-07.pdf).

20. KOLSTAD, *supra* note 19, at 356-64; U.S. EPA, *supra* note 19.

21. Cary Coglianese, *Is Satisfaction Success?: Evaluating Public Participation in Regulatory Policymaking*, in THE PROMISE AND PERFORMANCE OF ENVIRONMENTAL CONFLICT RESOLUTION 69-86 (Rosemary O'Leary & Lisa B. Bingham eds., 2003); TOM TIETENBERG, ENVIRONMENTAL AND NATURAL RESOURCE ECONOMICS 38-40 (7th ed. 2006).

22. KOLSTAD, *supra* note 19, at 356-64; TIETENBERG, *supra* note 21, at 38-40; Paul Portney, *The Contingent Valuation Debate: Why Economists Should Care*, J. ECON. PERSP., Fall 1994, at 3.

direction of any correlation between the proxy and the underlying variable of concern can be unclear. Is the proxy negatively or positively correlated with the underlying variable? For example, an improved relationship might mean one with less conflict, perhaps measured by the number of complaints community residents make to or about an industrial facility. Or an improved relationship might mean one with more interaction, in which case an increase in the number of complaints might signal first steps in developing relationships that will eventually overcome the problems or misperceptions that underlie the complaints. Similarly, researchers might measure the quality of communication by counting the number of hours facility managers spend in meetings with agencies and other stakeholders. But then again, lengthy meetings could also indicate that facility and agency managers have reached an impasse and are no longer listening to each other. Finally, trust might be shown through managers' willingness to disclose large quantities of information about their environmental performance to community residents. On the other hand, being on the receiving end of a large "data dump" could feel like bombardment and perhaps only reinforce feelings of mistrust. In each of these cases, investigators can observe plausible proxies for the underlying social variables of concern, but it may not always be clear whether changes in the level of the proxy indicate improvement or deterioration in the underlying social goal.

The desire to understand whether an ELP leads to some improvement gives rise to an added measurement challenge—namely, that the proxies gathered to assess progress toward social goals should be collected not only from or about participants in the ELP but also from an appropriate sample of non-participants.²³ Some revealed or expressed proxies may be readily obtainable from businesses, regulators, and environmental and community groups both within and outside a program. For example, investigators can presumably obtain a count of the number of lawsuits filed by community or environmental groups against polluters, whether or not the groups and polluters are involved in the ELP. On the other hand, some proxies may be more difficult to obtain from non-participants. For example, managers of facilities participating in an ELP might readily respond to a program-sponsored survey asking them about trust in government regulators because of its obvious relevance, but non-participating facilities may be less likely to respond to a request to complete such a survey. Obtaining data from both types of facilities, however, is essential to draw proper inferences about any program's impact.

II. Inference Issues: Determining the True Effects of an ELP

Determining how well an ELP achieves its goals involves more than just measuring progress toward the goals, even if appropriate proxy measures can be identified and collected. Investigators must use techniques to assess whether any progress they observe is actually the result of the program itself, or whether some other factor explains the results. They should be particularly aware of two confounding effects: (1) the "Hawthorne effect"; and (2) omitted variables bias.

The Hawthorne effect draws its name from a study of worker productivity conducted in the early part of the 20th century at the Western Electric Company's facility in Hawthorne, Illinois.²⁴ The effect refers to the potential for study subjects to act or respond differently just because they know they are being studied. In the Hawthorne experiments, researchers varied working conditions (such as lighting, schedules, and so forth) in an effort to determine how these conditions affected worker productivity. They found that over time, the productivity of the experimental group always increased—regardless of the changes made to workplace conditions.²⁵ The Hawthorne effect reminds researchers that the very fact that members of a treatment group know they are part of an experiment and are being observed may influence how they perform or respond. The potential for the Hawthorne effect is undoubtedly inherent in any voluntary program like ELPs. Researchers need to be mindful of the possibility that any effects that appear to have come from participating in *the* program under study may simply have come about because participating facilities knew they were in *a* program and were being observed.

The problem of omitted variables bias occurs when one or more factors left out of an analysis—usually because it is difficult or impossible to observe them—have an important effect on the observed responses or outcomes.²⁶ In such cases, the risk is that the effect of an omitted variable will be incorrectly ascribed to the variable measured and included in the analysis. For example, when investigators analyze the effects of participating in an ELP on some proxy for trust of stakeholders, they may leave one or more important variables out of their analysis, perhaps inadvertently, but more likely because it is difficult to gather data on all the variables that might influence the proxy. One such variable could be the degree of top-level management support for environmental activities. Not only might a facility's degree of top-level management support influence its level of trust of other stakeholders as measured after it joined an ELP, but top-level management might also influence whether a facility joins an ELP in the first place.²⁷ If this support is not accounted for in a study of the social effects of ELPs, any higher level of trust observed among participants could appear to have been caused by participation in the ELP itself when, in fact, some (if not all) of the difference was due to the preexisting degree of top-level management support for the environment among participants. An analysis of trust will not reveal the true effect of participation in the program if the omitted variable is never taken into account.

One way to attack the omitted variables bias problem and isolate the true effect of an ELP is to find a suitable proxy variable for the omitted variable. These proxies are slightly

23. JAMES H. STOCK & MARK W. WATSON, INTRODUCTION TO ECONOMETRICS §13 (2d ed. 2007).

24. *Id.*; F.J. ROETHLISBERGER & WILLIAM J. DICKSON, MANAGEMENT AND THE WORKER (1939).

25. ROETHLISBERGER & DICKSON, *supra* note 24.

26. STOCK & WATSON, *supra* note 23, §6; JEFFREY M. WOOLDRIDGE, INTRODUCTORY ECONOMETRICS: A MODERN APPROACH §3.4 (2d ed. 2003).

27. Other research, in fact, shows that the degree of top-level management support is strongly associated with participation in voluntary environmental programs. See Jonathan C. Borck et al., *Why Do They Join? An Exploration of Business Participation in Voluntary Environmental Programs*, in BEYOND COMPLIANCE: BUSINESS DECISION MAKING AND THE U.S. EPA'S PERFORMANCE TRACK PROGRAM (Regulatory Policy Program Report No. RPP-10, 2006).

different from the proxies discussed in the section above. There, we discussed types of proxies for the underlying social effect of interest or the outcome of concern. Here, we seek proxies for unobservable variables that influence the outcome of concern. But the ideas and the challenges are the same in both cases. For example, a proxy variable for the omitted variable of facilities' preexisting levels of top-level management support could take the form of responses to survey questions that had been administered before the ELP was established. Under certain statistical conditions, this proxy variable can be used to control for the unobserved variable and help isolate the true effect of the ELP.²⁸

Another way to tackle the omitted variables bias problem in a statistical setting is the so-called instrumental variables technique.²⁹ Unlike a proxy variable, which is a variable correlated with the unobservable omitted variable, an instrumental variable is a variable correlated with participation in the ELP. Under certain conditions, the instrumental variable technique can be used to isolate the true effect of the program.³⁰ The method of instrumental variables is well known but challenging to implement, as it requires some known randomness in at least one factor affecting the voluntary decision to participate in a program like an ELP.³¹

A more straightforward way to address the omitted variables bias problem is to use the "differences-in-differences" method. Differences-in-differences is one of a class of statistical approaches that use data collected over multiple time periods.³² The differences-in-differences technique requires investigators to collect data on participants and non-participants in a program in two time periods: (1) before the program; and (2) after the program. The method assumes that participants would change over time the same way the non-participants did if they never joined the program. If so, then any additional change in the outcome variable or effect of concern among the participants can be inferred to have been due to the influence of the program itself.³³ For example, suppose that investigators observed that the reported level of trust among a sample of managers from participating facilities increased by a certain amount after the facilities joined an ELP. Moreover, suppose the investigators observed that trust also increased among a sample of facilities that did not participate in the program over the same time period. The differences-in-differences technique identifies the effect of the ELP as the additional increase in trust among participants, not the total increase in trust.³⁴ The method provides not only a better estimate of the effect of the program but also a more

confident one. It helps to rule out alternative explanations, such as a preexisting but unobservable levels or propensities to trust.

Note that all the statistical techniques described above require gathering data about facilities or other actors that are not involved in a program. These non-participants serve as the control group; that is, they provide a basis for estimating what participants would have done in the absence of the program (the "counterfactual").³⁵ Without a properly chosen sample of non-participants against which to compare the behavior and evolution of participants, investigators cannot confidently estimate how much progress an ELP is making toward its goals—social or otherwise.

The data needs are even more extensive to implement the powerful differences-in-differences technique. Investigators must collect relevant data from: (1) the participants before they joined the program; (2) the participants after they joined the program; (3) a sample of non-participants before the participants joined the program; and (4) a sample of non-participants after the participants joined the program. Investigators who only collect data from participants in an ELP after they have joined the program have only collected one of the four chunks of data required to implement the differences-in-differences approach and thereby gain a best estimate of the true effect of the ELP.

Of course, in mentioning these statistical techniques, we do not mean to imply that the social effects of ELPs can only be studied through large samples using advanced quantitative analytic tools. Important insights can also be obtained through in-depth study of smaller numbers of facilities or programs. Yet, the challenges we have discussed about measurement and inference still arise. Fortunately, qualitative or small-sample research can also be designed in ways that respond effectively to these challenges.³⁶

III. Linking Social Effects to Environmental Performance

Up to this point, we have been primarily concerned with social effects or outcomes themselves, as if these effects are intrinsically valuable. Perhaps they are, in that people may be happier to live and work in communities in which relationships among stakeholders are strong and cultures within businesses and government agencies are cooperative. However, there seems to be good reason to suppose that many policymakers and managers quite properly view social effects as simply a means to the larger end of improved environmental protection. After all, most agencies implementing ELPs are environmental protection agencies, not social capital-building agencies.

If social effects are important primarily because of their subsequent effects on environmental quality, for example, because firms that garner greater trust tend to be the ones that take greater strides to improve their environmental performance, then identifying and obtaining good measures of and inferences about social effects will not be enough. In addition to the measurement and inference challenges we have already discussed, researchers and policy analysts will also

28. WOOLDRIDGE, *supra* note 26, §9.2.

29. STOCK & WATSON, *supra* note 23, §12; WOOLDRIDGE, *supra* note 26, §15.

30. STOCK & WATSON, *supra* note 23, §12; WOOLDRIDGE, *supra* note 26, §15.

31. Lori Benneer & Cary Coglianese, *Measuring Progress: Program Evaluation of Environmental Policies*, ENVIRONMENT, Mar. 2005, at 22.

32. WOOLDRIDGE, *supra* note 26, §13; STOCK & WATSON, *supra* note 23, §13.

33. Cary Coglianese & Lori Benneer, *Program Evaluation of Environmental Policies: Toward Evidence-Based Decision Making*, in DECISION MAKING FOR THE ENVIRONMENT: SOCIAL AND BEHAVIORAL SCIENCE RESEARCH PRIORITIES (Gary D. Brewer & Paul C. Stern eds., 2005).

34. *Id.*

35. Cary Coglianese, *Empirical Analysis and Administrative Law*, 2002 U. ILL. L. REV. 1111-37.

36. GARY KING ET AL., *DESIGNING SOCIAL INQUIRY: SCIENTIFIC INFERENCE IN QUALITATIVE RESEARCH* (1994).

confront challenges in determining the association between social effects and environmental outcomes. Some of these challenges will be familiar. For example, to determine if any observed changes in social effects lead to changes in environmental outcomes, researchers will face measurement challenges, such as the need to identify and collect measures not only of social effects but also environmental outcomes. Unfortunately, not all environmental outcomes are regularly or reliably measured. Thus, investigators must use those measures that are available, such as EPA's toxic release inventory (TRI), as proxies for overall environmental performance. As with all proxies, investigators must be aware that the proxies may not be highly correlated with the underlying outcome of concern, in this case overall environmental performance. Reported TRI figures, for example, may provide at best only a partial indicator of facilities' aggregate levels of pollution and their overall environmental performance, especially when that performance is understood to include energy and water use, among other things.³⁷

Familiar issues of inference also complicate the linking of social outcomes to overall environmental performance. As is true for any research issue, correlation is not the same as causation.³⁸ Just because investigators observe that facilities with greater levels of various social indicators or variables also have superior environmental performance does not mean that the increase in levels of social indicators or variables led to the superior environmental performance. For one thing, superior environmental performance might be what leads to increases in levels of social variables, not the other way around. Cleaner firms may prompt, and generate for themselves, greater trust. Furthermore, an unobserved third variable—the familiar “omitted variable” described above—may be responsible for observed effects in both the social and environmental variables. In other words, there may be something else—perhaps effective managerial leadership—that promotes both greater trust and improved environmental performance. Researchers and policy decisionmakers must be particularly careful to consider, and try to rule out, alternative explanations and causal pathways before crediting improvements in multi-stakeholder relationships, increases in trust, or changes in organizational culture with any observed improvements in environmental performance.

37. For a discussion of the strengths and limitations of TRI data, see JAMES T. HAMILTON, *REGULATION THROUGH REVELATION: THE ORIGIN, POLITICS, AND IMPACTS OF THE TOXIC RELEASE INVENTORY PROGRAM* (2005) and Lori S. Bennear, *Strategic Response to Regulatory Thresholds: Evidence From the Massachusetts Toxics Use Reduction Act* (June 27, 2005) (unpublished working paper), available at <http://ssrn.com/abstract=776504>.

38. WOOLDRIDGE, *supra* note 26, §1.4.

39. See Coglianese, *supra* note 21; Bennear & Coglianese, *supra* note 31; and LAWRENCE B. MOHR, *IMPACT ANALYSIS FOR PROGRAM EVALUATION* (2d ed. 1995).

IV. Conclusion: A Path Forward

Evaluating the effects of any public program—not just ELPs—requires attention to the kind of issues we have discussed in this Article.³⁹ These issues also arise no matter what the goals of the program may be. For this reason, no one interested in ELPs should think that the challenges in evaluating these programs will necessarily be any easier when the goals are defined in social rather than environmental terms. Any well-executed and meaningful evaluation of ELPs—whether for their impact on social or environmental goals—will need to attend to these concerns. In other words, it will never suffice simply to poll program members to see if they are satisfied with the program or if they think it is having social or environmental effects.⁴⁰ These effects need to be demonstrated through careful empirical research that attends to the issues we have outlined here.

Even though such research can be a daunting task, it is not impossible. Our discussion of evaluation challenges suggests ways to design evaluations that will yield convincing results. For example, investigators should first take care to identify plausible proxies for the underlying social effects of concern. Identifying and collecting multiple proxies for a single social effect or outcome can increase the level of confidence that the proxies are measuring the right effect, especially when the direction of correlation between a single proxy variable and the underlying social effect or outcome is unclear.

If investigators are primarily interested in social effects as a means to improvements in environmental performance, they should take care to link the social effect to environmental performance explicitly. To improve the confidence in making inferences about the effects of ELPs, investigators should be sure to collect data both from participants in the program and an appropriate sample of non-participants. Since some of the most powerful inference techniques require data from before the participants joined the program, investigators may wish to focus evaluations on industrial sectors or groups of facilities that have not yet joined a particular program and follow them as they join. Taking these evaluation issues into account before establishing new ELPs could provide opportunities for collecting pre-program data that can be used to compare to post-program outcomes or responses.

Overall, investigators must be aware of and transparent about alternative explanations for any correlations they observe. Only by addressing the issues we have outlined in this Article can researchers and regulatory officials rule out alternative explanations for their results and thereby increase confidence in what they can conclude about ELPs' success in achieving their goals—social or otherwise.

40. Coglianese, *supra* note 21.