University of Pennsylvania
ScholarlyCommons

Master of Environmental Studies Capstone Projects | Department of Earth and Environmental Science

5-1-2011

Lessons Learned from EPA's Climate Leaders Program: An Evaluation to Fortify Voluntary Environmental Initiatives

Vivian Futran
University of Pennsylvania

Presented to the Faculties of the University of Pennsylvania in Partial Fulfillment of the Requirements for the Degree of Master of Environmental Studies 2011.

This paper is posted at ScholarlyCommons. http://repository.upenn.edu/mes_capstones/41
For more information, please contact repository@pobox.upenn.edu.

Lessons Learned from EPA's Climate Leaders Program: An Evaluation to Fortify Voluntary Environmental Initiatives

Abstract
A voluntary environmental program (VEP) called Climate Leaders was recently cancelled by the U.S. Environmental Protection Agency (EPA). To find out why and gather lessons learned, the program was examined using a three-pronged approach: 1) a meta-analysis of program evaluation theory literature, 2) a review of guides and external reports on Climate Leaders, and 3) interviews with former program participants and implementers. Findings reveal that environmental protection is best achieved by combining regulation with voluntary methods as they complement and buttress each other. Recommendations were compiled to help future VEPs minimize wasted resources and improve environmental conditions. These recommendations were vetted by industry, and ultimately contribute to a comprehensive "guiding framework" for the design, implementation, and evaluation of VEPs. Existing VEP models and evaluative tools still lack the perspective of several key disciplines, so additional program evaluations are necessary to capture all of the primary characteristics correlated with program success and complete the guiding framework.

Disciplines
Earth Sciences | Physical Sciences and Mathematics

Comments
Presented to the Faculties of the University of Pennsylvania in Partial Fulfillment of the Requirements for the Degree of Master of Environmental Studies 2011.

This thesis or dissertation is available at ScholarlyCommons: http://repository.upenn.edu/mes_capstones/41

LESSONS LEARNED FROM EPA'S CLIMATE LEADERS PROGRAM:

AN EVALUATION TO FORTIFY VOLUNTARY ENVIRONMENTAL INITIATIVES

Vivian Futran

Spring 2011

Leslie Jones, EPA Region 3

Elaine Wright, University of Pennsylvania

Special thanks to Michael Dunn and Leslie Jones at EPA's Region 3 Office

for generously volunteering their time and insight

ABSTRACT

LESSONS LEARNED FROM EPA'S CLIMATE LEADERS PROGRAM:
AN EVALUATION TO FORTIFY VOLUNTARY ENVIRONMENTAL INITIATIVES

Author: Vivian Futran

Advisor: Elaine Wright

A voluntary environmental program (VEP) called Climate Leaders was recently cancelled by the U.S. Environmental Protection Agency (EPA). To find out why and gather lessons learned, the program was examined using a three-pronged approach: 1) a meta-analysis of program evaluation theory literature, 2) a review of guides and external reports on Climate Leaders, and 3) interviews with former program participants and implementers. Findings reveal that environmental protection is best achieved by combining regulation with voluntary methods as they complement and buttress each other. Recommendations were compiled to help future VEPs minimize wasted resources and improve environmental conditions. These recommendations were vetted by industry, and ultimately contribute to a comprehensive "guiding framework" for the design, implementation, and evaluation of VEPs. Existing VEP models and evaluative tools still lack the perspective of several key disciplines, so additional program evaluations are necessary to capture all of the primary characteristics correlated with program success and complete the guiding framework.

TABLE OF CONTENTS

CHAPTER ONE Overview 3
Problem Statement 3
Need for Research 3
Purpose of Study 4
Research Questions 4
Definition of Terms 5

CHAPTER TWO Research Design and Methodology 7
Evaluation Literature Review 7
Climate Leaders Interviews 8
Interviewees 8
Interview Bias and Error 10
Interview Analysis 13
Instrumentation 13

CHAPTER THREE Findings in Literature 15
History of Environmental Programs in the US 15
Environmental Program Evaluation 16
Evolution of VEPs: Why Some Succeed and Some Do Not 17
Economic Considerations: Program Features,
Company Finance, Market Conditions 21
Climate Leaders Program Background 26
Climate Leaders Termination Letter 28
Why Was Climate Leaders Canceled and What Were the Effects 30
Existing Lessons from EPA VEPs 31

CHAPTER FOUR Discussion of Study 34
Format 34
Objectives 34

CHAPTER FIVE Recommendations 36
Results of Climate Leaders Evaluation 36
Key Lesson 41
Limitations 42
Future Research 43
Conclusion 44

APPENDICES 46
Appendix A 46
Appendix B 48
Appendix C 49

WORKS CITED 51

LIST OF TABLES

Table 1: Summary of Data Collection Methods

Table 2: Overview of the Evaluation Process

Table 3: Project Development/Evaluation Cycle

LIST OF CHARTS

Chart 1: Three-Pronged Approach to Program Evaluation

Chart 2: Three Steps to Climate Leaders Evaluation

Chart 3: Climate Leaders Evaluation Results

CHAPTER ONE
Overview

Problem Statement

The United States Environmental Protection Agency has carried out many federal voluntary environmental programs (VEPs). These are often canceled for reasons ranging from poor execution to a change in the administration's budget or political priorities. Even strong programs that are deficient in just one area or are unexpectedly terminated can hurt the reputation of VEPs, thus weakening an important tool for environmental protection. This leads to loss of resources invested by implementers and partners, highlighting the need for a tool to comprehensively asses all stages of a program's lifespan including design, implementation, and evaluation (Arora, 1995). Since voluntary programs represent a relatively new style of environmental governance, little progress has been made on VEPs in Environmental Program Evaluation research (Knaap, 1998).

Need for Research

The field of Environmental Program Evaluation is growing, but existing models for creating strong VEPs are not comprehensive. A guiding framework that captures the diverse program characteristics known to correlate with success is needed. It must cover the program from its initiation to its end, and accommodate the needs of government, business, and environment (Boyd, 2011). It must also be a synthesis of existing program design and assessment tools drawn from several disciplines. While completing a fully comprehensive framework is outside the scope of this project, the research and evaluation conducted here are steps in this direction. They should serve as a foundation for future research. Additional VEP evaluations from fields including public policy, political science, management, business, and economics will also be needed.

Purpose of Study

This study was intended to help compile characteristics for the ideal voluntary environmental program by evaluating a recently terminated voluntary environmental program called Climate Leaders. Successful strategies that could be applied to other ongoing and future programs were captured, as well as characteristics and trajectories to be avoided. In the short-term, these "lessons learned" will improve the performance of EPA's ongoing and future voluntary environmental programs. In the long-term, they will help advance a guiding framework for designing, steering and evaluating voluntary environmental programs. It might even prove relevant to VEP bundling schemes which group together voluntary environmental programs for easy access at the federal and regional level, like the Sustainability Partnership of EPA's Region 3 (Mid-Atlantic) Office in Philadelphia.

Research Questions

The research questions being investigated are:

1. How do you define success for the Climate Leaders program?
2. In what ways did Climate Leaders succeed?
3. How could Climate Leaders have been improved?
4. What circumstantial conditions (e.g. political, economic) are important to the success/failure of VEPs?
5. What inbuilt features (e.g. structural, procedural) are important to the success/failure of VEPs?
6. What lessons learned from Climate Leaders can be applied to ongoing EPA VEPs on a federal and regional level?
7. What improvements are not accounted for in existing frameworks that must be added for comprehensibility in evaluating program performance?

Definition of Terms

Program

Program here is defined as "a set of staff activities with a defined goal or purpose." The Government Performance and Results Act terms these as "activities" to achieve goals and objectives (Pumphrey, 1993).

Public Voluntary Program

A public voluntary program (henceforth referred to as a "voluntary program") is an initiative or event that seeks to achieve a goal by means of willful participation. It is a collaborative, non-regulatory approach most clearly characterized by what it is not: a mandatory program (Boyd & Manson, 2011). Voluntary programs are organized and implemented by an individual or group. As long as implementation faculties persist they can exist even without active participation, however they will not have a purpose without the involvement of participants. One subset within this category of program is specifically dedicated to environmental stewardship. "Voluntary environmental programs (VEPs) seek to improve the environment by encouraging, rather than mandating, businesses and other organizations to adopt environmentally protective measures" (Borck, 2009).

A voluntary program involves activity over a span of time, and stops existing when it ceases to be implemented. A voluntary *agreement* is fundamentally different because it is initiated by a one-time action: two (or more) parties sealing their commitment. This can be a handshake, signature, or anything else that constitutes a willful acceptance of mutually agreed upon terms. An agreement stops existing when it expires or when the parties involved revoke their commitment. A voluntary program can include a voluntary agreement by members, but a voluntary agreement does *not* imply the existence of a program.

Voluntary Environmental Program (VEP)

Voluntary environmental programs consist of programs, codes, agreements or commitments. They encourage private, public, or nonprofit organizations to voluntarily reduce their environmental impacts beyond the requirements established by environmental regulations (Darnall, 2005: Carmin et al., 2003).

VEP Guiding Framework

A comprehensive plan for the design and implementation of a voluntary program, tailored to address the special goals, needs, and scope of the environmental issue being addressed. It must guide program staff throughout the life of the program from conception to termination. It must also include all of the characteristics of performance and structure necessary for success and satisfaction by both implementers and participants.

Evaluation

The term evaluation describes different models or data collection strategies to gather information at different stages in the life of a project in a systematic investigation of worth or merit (Westat, 2002). The product is an assessment of the results, impact, or effects of a program or policy (Pumphrey, 1993). Evaluations have several distinguishing characteristics relating to focus, methodology, and function. An evaluation: 1) assesses the effectiveness of an ongoing program in achieving its objectives, 2) relies on the standards of project design to distinguish a program's effects from those of other forces, and 3) aims at program improvement through a modification of current operations (EPA Evaluation Glossary, 2010).

CHAPTER TWO
Research Design & Methodology

The primary objective of this research is to compose recommendations to improve future VEPs using the Climate Leaders Program as a case study. Ultimately this will contribute to a comprehensive guiding framework for voluntary environmental programs. There were three main steps involved in this undertaking. (See Appendix C, Chart 1 for the Three-Pronged Approach). The first was synthesizing and summarizing existing literature from evaluation theory. The second was reviewing publications on Climate Leaders to provide context. The third was assessing the Climate Leaders program through interviews. (See Appendix A, Table 1 for a Summary of Data Collection Methods).

Evaluation Literature Review

This study began with a meta-analysis of existing literature on program assessment standards. It focused on evaluative standards for programs that resemble Climate Leaders in goal and scope. The analysis included surveying existing evaluation methods from organizational literature, including the resources on the EPA Evaluation Support website (EPA, 2011). Next was researching scholarly journal articles (in the Journal of Environmental Economics and Management, the Annual Review of Environment & Resources, etc.) as well as publications and guidelines by EPA and other relevant organizations. By cross-referencing existing lists of criteria for successful environmental programs, especially those focused on voluntary ones, a list of VEP evaluation standards was created against which to compare the components and operations of Climate Leaders retrospectively.

Climate Leaders Interviews

Information about Climate Leaders was gathered in order to assess the program based on the evaluation literature reviewed. This entailed searching EPA websites, publications, articles, and conducting interviews with former EPA program implementers (EPA employees) and "partners" (participating organizations and businesses). (See Appendix C, Chart 2 for Three Steps to Climate Leaders Evaluation).

Interviewees

All interviewees were offered full anonymity (or alternatively, select attribution) to encourage candor. They were told that the information gathered would be presented in an aggregate form, without specific attribution or quotations unless explicit permission was given. Each was given background on the purpose of the study: to evaluate Climate Leaders for the improvement of ongoing and future EPA programs, and contribute to the development of a guiding framework for VEPs. Their incentive to participate was the opportunity to help improve VEPs from EPA and other organizations which would benefit the environment in the long-run. Interviewees were also offered a copy of the final project report.

Interviews with program implementers fell into two categories. First were EPA staff from headquarters in DC who had most control over the recently terminated Climate Leaders program, who provided insight into the intended structure and process behind the VEP. For additional perspective, interviews were also conducted with members of the EPA office in Philadelphia. This Regional office had decided independently to promote Climate Leaders at a more local level. They provided a viewpoint from the "inside," but theoretically with less ownership bias.

Interviews with program participants offered a complementary point of view, addressing how program intention was translated into action. Questions addressed what value was gained from involvement, if any. These two interview groups (former partners and EPA implementers) were chosen because they were most knowledgeable about the process and results of Climate Leaders. They invested more time in dealing directly with the VEP than did the designers or more senior agency Directors.

A total of thirteen formal interviews were conducted, each addressing all of the standardized interview questions. These played a fundamental role in this study because they helped define success for the Climate Leaders program. Of the formal interviews, four were with EPA employees, two of which had worked on the program full time. The other two had helped advertise and recruit for Climate Leaders through a Regional level voluntary program "bundle." The remaining nine interviewees were representatives from former Climate Leader partners, all of whom had engaged directly with the program as part of their company's "climate," "air," "environmental," or "sustainability" team. Of the partner interviewees, two were representatives of smaller businesses and seven were from large well-known companies. This category included one interviewee who had previously worked at EPA and helped design the program, providing a unique perspective that straddled both partner and implementer. (See Appendix C, Chart 3 for Formal Interviews Breakdown).

Other initial informational "interviews" that did not adhere to the same standard of consistency were included for context or background, but not in the interview analysis portion of this project. These will be referred to as "personal communications" from now on, and will not be included in further references to interviews.

Interview Questions

In each interview, eleven core questions were asked. These were posed to all interviewees, while follow-ups for clarifications varied depending on the thoroughness of the initial response. The interview questions for participants and EPA employees differed only slightly, and these variations are included with the questions in parenthesis where applicable. The core questions from the interview script can be found in the Instrumentation section below.

The interview questions were meant to find out how Climate Leaders functioned in practice, and went deeper than published descriptions of the program's theoretical design as envisioned by EPA. The interviewees illuminate how these programs actually performed, so that true faults—especially procedural ones—could come to light. Without these interviews, official analysis and published commentary would have to be taken at face value despite possible prejudice.

Interview Bias and Error

A major strength of this research was the objectivity of analysis; it was not authored by an EPA affiliate or former partner. However, while there was no prejudice on the part of the researcher, there were a few areas of potential error inherent in the study due to time and resource constraints.

First, constraints on time and finances limited the number of interviews to the point that statistical trends could not be extracted. It was not possible to interview a random sample of Climate Leader implementers and participants, as only those parties that responded to email solicitation were interviewed and included in analysis. Thus the sample was somewhat self-selected. This can lead to sampling error: the probability that if another sample of the same size were drawn, different results might be obtained (Sudman, 1976). It

was also not possible to conduct enough interviews to achieve statistical significance due to time and financial constraints.

Second, while it is true that larger samples would have reduced sampling error, sampling error is the smallest of the three components of error that affect the soundness of sample designs (Westat, 2002). Two other errors—sample bias (primarily due to loss of sample units) and response bias (responses or observations that do not reflect "true" behavior, characteristics or attitudes)—are more likely to jeopardize validity of findings (Westat, 2002).

There were two main forms of sample bias present. One was that the partners that accepted the invitation to be interviewed were mostly big, profitable organizations with a culture of prioritizing environmental performance. The respondent characteristics were also skewed towards those that had the resources dedicated to an environmental representative or team. Since these representatives are active in the larger environmental activities of industry and business, they had distributed their contact information at conferences and other environmental events making them easier to track down. A second form of sampling bias was due to the fact that the VEP partners and non-partners systematically differ given that participation in a public voluntary program is optional. More specifically, firms most likely to benefit from a VEP are thus more likely to sign up as partners. This probably lead to more progress than would have been possible if a random sample of the country's companies had been involved. "Large firms, greener firms, or firms with good management may have been more likely to participate for a variety of reasons. If these factors are independently related to outcomes (as they probably would be), the partner sample will be biased" (Boyd & Manson, 2011).

Finally, interviewees may have had their own reasons to be interviewed affecting their willingness to be involved in the study, and the content of their responses. For example, this project may have been seen by a company as an opportunity for positive press, or as an opportunity to get in the good graces of the EPA. This also extended to EPA employees who had the desired to represent their Agency well. These motivations may have lead to an overstatement of program benefits, thus skewing results (Boyd & Manson, 2011).

Researchers may not know all of the relevant characteristics that can distinguish between the treatment and comparison groups, however, and even if they could, some key factors (like a firm's motivation to innovate, or the presence of a dynamic leader to implement changes) may be impossible to measure. The effort to develop a complete set of control variables creates extensive data demands, particularly for a VEP that address more than one sector (Boyd & Manson, 2011). Future analysis should account for these biases, especially in a quantitative study.

Despite all of the constraints and potential distortion, it was worth sacrificing the randomness and pure impartiality of the sample to get as much information as possible from those groups available. As this was a qualitative investigation, it was most interested in locating information-rich cases for study in depth, not necessarily in statistically valid yield (Westat, 2002). Purposeful sampling was practiced instead of a random sample or a stratified sample of a project's participants in this evaluation. It focused on the higher tier achievers admitted to the program, on those participants who had a basis of comparison from having participated in similar programs, and those from a variety of business sizes.

Interview Analysis

Qualitative analysis methods were used in this research, so statistical significance and randomness were not as critical as they would have been in a quantitative study. Considering the time and resource constraints on the project, this design focused on revealing information about the successes and weaknesses of the program. In-depth descriptions were the goal, as opposed to a statistically significant number of totally standardized responses. Interviewees were given time to reflect and space to elaborate so that unique insights would arise from the natural train of thought of the interviewee. This organic exploration would likely have been impossible with shorter more pointed questions requiring responses that were easier to quantify.

The interview questions examined personal experiences with Climate Leaders, with the intention of eliciting as much candor from the interviewees as possible. The questions were designed to define success in the context of voluntary environmental programs, and probed for perceptions of Climate Leaders—successes and shortcomings—from different angles. In revealing what was encouraging or frustrating about the program, how each question was *interpreted* and what the interviewees *emphasized* was just as important as the content of the answer.

Instrumentation

The interview questions were drafted to reflect the literature review, and finalized according to the guidance of EPA professionals (Laskowski, 2011) who have experience with voluntary environmental programs at the federal and regional level. The questions were based on the expected categories in a VEP guiding framework, with an eye to validity and

reliability. The interviews were implemented with consent from EPA employees and program participants.

Interview Questions:

1. What was your role in Climate Leaders?
2. To your knowledge, how was the performance of this program tracked during its existence (According to what metrics and/or what indicators)?
3. According to what specific criterion would you personally have judged the success or failure of this program (if you were a manager in EPA)?
4. For Climate Leaders, how do you define success?
5. What was your overarching impression of the program: was it a success or a failure according your experience?
6. What qualities for success do you think this program exemplified, if any?
7. If you had to choose a single attribute, what was the main reason this program was a success/failure?
8. What specific qualities for success do you think this program lacked, if any? (i.e. How could this program have been improved?)
9. Do you think others with your role (implementer vs. partner, from a similar organization) had the same experience with this program?
10. Do you think others in a different role (implementer vs. partner, from a different kind of organization) had the same experience with this program?
11. Anything else to add?

CHAPTER THREE
Findings in Literature

History of Environmental Programs in the US

The US EPA was founded in 1970 and issued a Congressional mandate to ensure compliance with ecologically responsible practices (Darnall, 2005). The agency traditionally relied on regulatory programs to compel individuals and organizations (Borck, 2009), but struggled to monitor through inspections and audits against budget cuts and limited funding allocated by Congress. Critics of the regulatory style argued that this system was inefficient at achieving environmental improvements. They pointed to media divisions—air, water, waste, etc.—and the ensuing fragmentation of efforts (Darnall, 2005).

One way to address these concerns was through voluntary environmental programs. VEPs are collaborative arrangements between firms, regulators, or other third parties in which firms voluntarily commit to actions that improve the natural environment (Delmas, 2001). The appeal of voluntary environmental programs lies largely in their flexibility, and in their potential to mutually serve government, industry and environmental interests (Steelman, 2006). VEPs encourage rather than mandate businesses and other organizations to adopt environmentally protective measures (Borck, 2009). They rely on a "carrot" instead of a "stick" approach, spurring commitment to program goals by offering incentive and market-based initiatives as opposed to regulations. They utilize benefits to reward intentional and unintentional over-compliance which reduces enforcement needs. This allows for more holistic approaches with lower administrative costs than would otherwise be possible (Pizer, 2008).

Unintentional over-compliance has been recorded since the beginning of environmental regulation. However, in the late 1980's and into the 1990's there was a rise of

companies' *intentional* over-compliance with EPA regulations. From the business perspective, voluntary over-compliance yielded a variety of benefits and is a rational reaction to environmental regulation (Seema, 1996). For example, there is evidence that this environmental stewardship was engendered by increasing public awareness of environmental challenges (Arora, 1995). Companies became greener in an attempt to boost their reputation with consumers.

This increase in intentional over-compliance was accompanied by a corresponding rise in voluntary EPA programs (Borck, 2009). It coincided with the administrations of Presidents Reagan and Bush (41st) who tightened EPA's belt, and made it even more important for the Agency to cut costs without sacrificing environmental protection (Laskowski, 2010). By the late 1990s, Mazurek's (2002) estimates suggested that more than 13,000 companies were participating in VEPs, and this number has continued to grow (Darnall, 2005). EPA has since expanded its catalog to over 100 voluntary programs to buttress its command-and-control regulation and cut expenditures (Borck, 2009).

Environmental Program Evaluation

Not surprisingly, Environmental Program Evaluation is a growing sub-discipline (Knaap & Kim, 1998); evaluation of voluntary environmental programs is a research niche within this arena being gradually addressed by academia, particularly public policy and related fields (Arora, 1995). Since conventional principles and approaches to program evaluation were developed in a system predominated by command-and-control regulation (Arora, 1995), these must now be tailored to the newer wave of voluntary environmental programs. To this end, a selection of studies from the emerging field of voluntary program research was synthesized. There are two important points to emphasize before delving into

this existing literature: 1) the aim of this project is to contribute to an *evaluative* tool, not a measurement tool, and 2) the research mentions other forms of voluntary agreements in general, but focuses on the characteristics of voluntary *programs* specifically.

Evolution of VEPs: Why Some Succeed and Some Do Not

Daley (2007) used a political science lens to observe US public policy shift from a regulatory to a more "Voluntary Remediation Programs" approach. He analyzed the trend specifically in relation to the environment. He determined that state policy adoption was a response to interest group pressure, and had a "ripple effect" on surrounding states. This sub-national shift affected the business mindset in many formerly polluting/wasteful industries, which created a more welcoming atmosphere in which EPA could develop voluntary programs at the federal level (Daley, 2007).

Following a similar ripple effect logic, Arora and Cason (1995) reported that the characteristics of the *target businesses* determined participation—and hence the viability—of a voluntary program. They assessed the ability of a voluntary approach to supplement more traditional regulatory approaches in these target industries. The researchers used EPA's 33/50 Program as a case study to distill out the factors correlated with participation in voluntary programs: technologies, production process, size, 33/50 chemical release, financial health, etc. Based on formulas, they concluded that voluntary programs like 33/50 were effective (Arora, 1995). They believed that garnering private sector participation in VEPs by mirroring a competitive market could lead to significant environmental improvement at a national level. They recommended that for best results, EPA encourage intra-industry competition in environmental quality by offering substantial public recognition of achievement (Arora, 1995).

Delmas (2010) demonstrated that substantive cooperative strategies are more likely to be pursued by firms that join a VEP close to its initiation, while symbolic cooperation is more likely behavior by late joiners. Although VEP early joiners reduced their emissions more than nonparticipants, the study results show no significant difference overall between participants and nonparticipants in the reduction of their emissions (Delmas, 2010).

One reason this occured was put forth by King and Lenox (2000) as well as Rivera, de Leon and Koerber (2006). They indicated that without sanctions, independent oversight, and standards, VEPs were not effective in promoting improved corporate environmental performance (Khanna, 1999). This analysis applied to programs like Climate VISION, Climate Challenge and Climate Leaders, in which the expectations of industry were met through flexible and cheaper environmental protection requirements, whereas expectations of the environmentalists and other stakeholders were neglected (Levy, 1997). If VEPs were used to serve some interests (e.g. businesses) to the exclusion of others (e.g. environment and community), the authors determined that they were likely to lose their value as potentially useful regulatory instruments, thereby impoverishing the already limited governing tool box (Steelman, 2006).

Lyon and Maxwell (2002) also presented an empirical analysis of the efficacy of voluntary programs. They went a step further, creating a construct to organize the large variety of voluntary programs and broke them down into three broad types—unilateral, public, and negotiated— with corresponding examples. They emphasized factors that determined corporate willingness and ability to undertake voluntary action, and summarized the downsides to voluntary programs including social welfare-enhancements and social

welfare reductions, regulation weakening, regulatory monitoring reduction, etc. (Lyon, 2002).

Borck and Coglianese (2009) also described three different structures for VEP's: unilateral, bilateral, and public. They reviewed research literature on voluntary environmental programs to establish whether VEP's delivered real improvements, or just took credit for existing public momentum. Their focus was on the ways VEP's operated, citing this as the major factor in their success or failure (Borck, 2009). Their research suggested that "VEP's with more moderate standards—but with strong enforcement of them—will have the greatest overall effectiveness." The authors emphasized participation and environmental improvement as the two main variables in a simple "effectiveness equation" (Borck, 2009). By comparing VEP's to clubs with membership, standards and side-goals, the researchers extended this metaphor to illuminate some of the advantages (e.g. collective benefits) and challenges (e.g. defining environmental goals) as well (Borck, 2009).

Delmas and Terlaak (2001) dove deeper into the costs and benefits of voluntary initiatives when in a competitive context. They described the origins and workings of many programs in the environmental arena, and provided a construct for sorting such programs based on these costs and benefits. They compared two main kinds of voluntary agreements—negotiated agreements and public voluntary programs—with a focus on organizational design. The authors also investigated programs that had been bundled into larger comprehensive voluntary initiatives. Finally, they compared US voluntary programs to those in Europe to extract international best practices (Delmas, 2001).

Darnall (2005) also introduced VEP categories, although his focus on the design mechanisms rather than the outcome of the program. The author's findings suggested that

there are four distinct types of programs with varying degrees of rigor: 1) information, assistance and awareness, 2) environmental pledge, 3) voluntary reporting, and 4) performance monitoring programs. The author found that information for differentiating among program types was limited, which introduced a potentially problematic phenomenon: less rigorous VEPs could falsely signal that their administrative, environmental performance and conformance requirements were just as strong as programs with more robust designs. Moreover, the lack of monitoring and sanctions in less rigorous programs created opportunities for participants to free-ride and receive benefits without satisfying VEP requirements. Darnall warned that unless some means of distinguishing among program types is implemented, these issues could threaten the long term viability of VEPs as a tool for environmental protection, and the credibility of market mechanisms more broadly (2005).

Segerson and Miceli (1998) presented another model assessing effectiveness of voluntary schemes, but differentiated it by making it predictive. It was meant to determine if an initiative was likely to lead to efficient environmental protection *before* the initiative had been implemented. This was a very simple model helpful in determining not only the effectiveness of the voluntary agreement predicted, but also whether a voluntary agreement was even the likely outcome of a given interaction between a company and a regulating body. "Polluters" were grouped depending on if they reacted to a "stick" or a "carrot" alterative program. Results depended on various factors: allocation of bargaining power, magnitude of background threat, social cost of funding (Segerson, 1998). This model applied just as well to programs as to other forms of "agreements."

Pizer, Mogenstern, and Shih (2008) kept a narrow focus, looking at voluntary climate programs in the US. They investigated the participation of manufacturing firms in EPA's

Climate Wise program and the US Department of Energy's Voluntary Reporting of greenhouse gas program. The authors analyzed whether the results of either program had a real environmental impact, for how long, and why. By the metrics used in this study, the programs had modest—albeit statistically significant—impacts on fuel costs and electricity costs respectively, and only for a short time. The authors explain, however, that their methodology was limited by the youth and complexity of the field (Pizer, 2008).

Delmas and Keller (2005) also dove more deeply into a narrow topic: the Free Rider problem, a potentially major threat to the voluntary programs of interest. They investigated the WasteWise program to determine what factors (original motivation to enter the voluntary program, characteristics of the organization, etc.) encouraged or discouraged a particular behavior (Delmas, 2005).

Economic Considerations: Program Features, Company Finance, Market Conditions

There are three different dimensions on which economic factors weighed into the performance and results of a voluntary environmental program. The first was on the program level. Alberini and Segerson (2002) measured voluntary initiatives against traditional mandatory approaches (regulation or tax), taking into account market effects. They tracked the evolution of voluntary approaches to environmental protection through an economic lens by breaking the many voluntary agreements down into three broad categories: industry self-regulation, bilateral agreement, and government program. They laid out conditions for success in each category, using environmental effectiveness (participation/incentive and pollution abatement) and relative efficiency (cost and environmental quality) as measures. They identified voluntary program features key to efficiency: polluter choice, flexibility and

increased communication between industry and regulator. Conversely, a lack of enforcement and abatement above the norm were potential disadvantages (Alberini, 2002).

The second level on which economic factors weighed into the success of a VEP was through company features. Work by Khanna and Damon (1999) attempted to determine what factors encouraged VEP membership on this level. They concentrated on characteristics of partner firms, a level that they believed weighed heavily into whether a business would participate in a program (Khanna, 1999). These characteristics included a company's past environmental penalties, predictions about future regulation, leadership status in an industry's environmental standard-setting, and calculations of potential abatement and administrative costs (Futran, 2010). Companies are profit maximizing institutions, so determining the financial impact of an investment like program partnership before approving the expenditures is critical to their survival. It was important that firms be able to design and execute projects with financial benefits, or it would be less likely that the required investments would be made (Zatz, 1999).

There were several routes by which businesses benefited economically from participating in VEPs. One was by leveraging the environmental movement to grow consumer demand. Volunteer environmental programs provided a venue through which firms could compete to be leaders in environmental quality. Companies that publicized their efforts to be "greener" earned a reputation for environmental responsibility which is highly valued today by a public harboring greater awareness of environmental challenges (Videras, 2000). This title was especially valuable in industries that have high advertising expenditures and high contact with their consumers (Arora, 1996). In fact, as public recognition was established as a significant incentive for firm participation in EPA programs, the Agency

began publicly celebrating of achievers to motivate membership in its voluntary environmental programs. In this way, environmentalism was in line with the ultimate business goal of maximizing profit (Arora & Cason, 1995). Since consumers were willing to pay a premium for products and services linked to environmental stewardship, companies got a demand-side boost as a reward for participation in EPA programs (Arora & Cason, 1996). This ran contrary to the popular belief that environmental management hurt the bottom line (Klassen, 1996).

Another way firms reaped the benefits of joining VEPs was through savings. Even when the immediate impact of program participation was negative due to increased expenditures, evidence showed that investors expected such firms to be more profitable in the long run (Khanna, 1999). There was evidence linking strong environmental management to improved financial performance (as measured by the stock market) through cost cutting and market gains (Klassen, 1996). Money thrown away on inefficient energy use, water leaks, missed opportunities for recycling, over-packaging, pollution penalties, etc. could have been saved by exploiting the guidance of EPA program implementers. Also, market gains came from improved trading futures, which start with a consumer's willingness to pay even higher prices to a firm that generates less environmental harm (Arora & Cason, 1996).

It was clear that environmental stewardship could increase a firm's profitability (Klassen, 1996). Logically, it would follow that voluntary environmental programs which increased demand and decreased costs for participating firms would have greater company membership, increased longevity, and thus a greater impact on environmental protection than those that did not. For example, participants in EPA's 33/50 program calculated a net economic benefit from involvement, boosting the VEP's membership and contributing to its

success (Futran, 2010). These observations lent support to the claim that VEPs are more likely to succeed if they offered net savings as opposed to net loss for participant organizations. The bottom line is extremely important when deciding whether to get involved with EPA programs, but this is not a straightforward calculation. Many companies determine their success with a longer term measure than a quarterly or yearly statement. To do so, they keep an eye on public opinion and stock markets. They participate in voluntary environmental programs because they realize that their investments in a VEP are covered by an eventual payback, and their company's value and net profit are enhanced over time. Much more extensive literature is required to prove the actual impact of a partner's economic costs and savings in the success of federal voluntary environmental programs. An important variable that must be controlled is the larger economic context: whether the country or the industry is in a boom or a recession. The significance of this variable is yet unknown, and with future research it may change the entire equation.

The third way that economic factors affected VEP performance was on the market level. Boyd's (2011) work took this perspective, looking beyond individual company features to assess VEPs within the context of the larger economic environment. He asserted that the most fertile ground for new voluntary environmental programs would be laid by market failure—a situation in which regular market forces are not producing efficient environmental outcomes. According to Boyd, greenhouse gas emissions represented such an opportunity: the environmental costs of carbon emissions were being externalized. The market was not setting greenhouse gas limits despite the health and long-term economic harm. A government program free of membership charges like Climate Leaders addressed this disparity. It was the equivalent of subsidizing the information, resources, tools, or

guidance for participating firms, large and small. Large firms found it easier to "diffuse the fruits of technical assistance within their organizations (e.g., to multiple manufacturing facilities)" (Boyd & Manson, 2011). This implied that a given technical assistance intervention could yield large social benefits when applied to larger firms. Small firms, in contrast could arguably innovate aggressively with no need for government intervention or VEP assistance. However, they were found particularly susceptible to "research and development" market failures making it difficult for them to create, protect, and enjoy the benefits of new intellectual property (Boyd, 2011). VEPs could provide small firms with this security. Boyd grouped successful VEPs into two categories based on these parameters: 1) "programs that provide technical assistance and rectify the failure to invest in research, development, and technological innovation," and 2) "programs that provide signaling features which address information imbalances associated with complex products and processes, specifically incomplete information that can result in decisions that lead to negative environmental impacts." He warned against environmental programs that promote behavior which would have occurred anyhow, as a result of existing market conditions.

Boyd's claim that government programs should address market failures also applied to government evaluations of such programs. Government interventions, whether EPA program implementation or the evaluations that follow, were found more difficult to justify if there were no underlying conditions that would normally be considered market failure. Program evaluations are not easy to appropriate for profit, so they were not incentivized by the market (Boyd, 2011). Boyd found environmental programs and evaluations needed, especially those sponsored by academia and government, when markets failed to produce them. They provided information for communicating to a variety of stakeholders, and gave

managers the data they need to report to decision makers about the outcomes of their investments (Westat, 2002). This idea of reporting on the outcomes of federal investments has grown in prominence over the last several years, in large part due to the establishment of the Government Performance and Results Act (GPRA). The GPRA requires federal agencies to report on the accomplishments of their funded efforts annually, including establishing broad goals or strategic outcomes, performance outcomes, and performance indicators against which progress will be assessed (Westat, 2002). The evaluation of Climate Leaders undertaken in this project provided information that the market had failed to supply, and facilitated accounting of EPA's effort in the program.

This literature review of environmental program evaluation, specifically regarding the evolution of environmental voluntary programs, provided many important insights into what makes VEPs successful. It provided a context against which to analyze the Climate Leaders program, a step that will contribute to a guiding framework for VEPs. This tool will improve the overall performance of VEPs by taking in the many lessons learned from work in the past and yielding a comprehensive guide. This will boost the effectiveness and efficiency of VEPs, and help environmental organizations and agencies avoid repeating serious mistakes like prematurely aborted programs and wasted resources. The evaluation of Climate Leaders was a building block for this guiding framework.

Climate Leaders Program Background

Climate Leaders was an EPA voluntary program that worked with companies to develop long-term comprehensive climate change strategies (Design Principles, 2005). It was started in 2002 to get companies' carbon management off the ground, offering an early standard for tracking and reducing greenhouse gas (GHG) emissions. (See Appendix B, Quote B). To fulfill their program commitment, partners had to complete a GHG inventory,

set a reduction goal and meet that goal (McCarthy, 2010). To do so, EPA guided them in setting corporate-wide GHG reduction targets, planning emissions reductions projects, and gathering emissions inventories to measure progress.

The Climate Leaders' "Greenhouse Gas Inventory Protocol" was based on an existing corporate greenhouse gas inventory protocol developed by the World Resources Institute and the World Business Council for Sustainable Development. Through collaboration between industry, government, and non-governmental organizations, the World Resources Institute and the World Business Council for Sustainable Development produced widely accepted accounting practices for measuring and reporting corporate greenhouse gas emissions. Their "GHG Protocol" Corporate Standard consisted of a corporate accounting and reporting standard and calculation tools (EPA Design Principles, 2005).

According to EPA implementers, the success of Climate Leaders relied on highly engaged partners moving through program at good speed, submitting data in timely fashion (Anonymous #1, 2011). EPA tracked the program's process by the number of 1) partners signed up, 2) inventories submitted, and 3) goals approved by EPA (Anonymous #2, 2011). By reporting inventory data to EPA, partners created a lasting record of their accomplishments and identified themselves as corporate environmental leaders. Climate Leaders could also help strategically position partners for unfolding climate change policy (EPA Design Principles, 2005). The hope was that this program would allow partners to broaden their emissions reduction plan and align their priorities in other areas to be more environmentally responsible (Anonymous #1, 2011).

Climate Leader partners committed to (EPA Design Principles, 2005):

- Developing a corporate-wide **GHG inventory** of all sources of the six major gases (CO_2, CH_4, N_2O, HFCs, PFCs, and SF_6) using the Climate Leaders GHG Inventory Protocol
- Setting an aggressive corporate-wide GHG **emissions reduction goal** to be achieved in the subsequent 5-10 years
- Developing a corporate GHG inventory **management plan**
- Annually **report inventory data** and document progress towards their reduction goal
- **Publicize** their participation, reduction pledge, and accomplishments achieved through the program

In return, EPA provided (EPA Design Principles, 2005):

- Recognition
  - Press events
  - Articles and public service announcements in business and trade publications
  - Speaking engagements at industry conferences
  - Case studies highlighting Partner achievements
- Technical Assistance
  - Developing a GHG inventory
  - Reviewing inventory management plan
  - Setting a GHG reduction goal
  - Exchanging with peers through Climate Leaders Partner meetings
- Credibility
  - A credible, transparent GHG reporting mechanism that will develop with the science
  - Assurance that Partners have created a high-quality GHG management process

Climate Leaders Termination Letter

In September of 2010, EPA announced that it would "phase down services the agency offers under its Climate Leaders program over the coming year and encourage participating companies to transition to state or non-governmental programs" (McCarthy, 2010). According to the termination letter sent to partners, the transition would "allow the agency to realign resources to better assist companies in learning from the emissions data collected

under the GHG Reporting Program" (McCarthy, 2010). In other words, EPA terminated the program because it wanted to put its time and money elsewhere.

The letter said that among EPA'S reasons for this decision were that other new regulatory and voluntary programs existed that address greenhouse gas emissions. These include EPA's Green Power Partnership, ENERGY STAR, and other members of its climate Protection Partnership. Finally, the letter pointed to the multiple states and non-governmental organizations (NGOs) offering climate programs that EPA claimed were now robust enough to serve companies' greenhouse gas management needs. For example, many large firms were already members of the Carbon Disclosure Project (CDP) and The Climate Registry (TCR), two alternative voluntary greenhouse gas reduction programs. What percent of former Climate Change partners will continue GHG reduction activities is yet to be seen. EPA promised to look for new ways to promote, support and recognize climate leadership. It even noted the US's first mandatory greenhouse gas reporting rule that took effect on January 1, 2010 as evidence of the importance of GHG reduction to the Agency (EPA Press Release, 2010).

According to interviewees in this study, the decision to end Climate Leaders was made without input from those at EPA actually managing the day-to-day activities of the program. Climate Leaders staff was reassigned to other programs that did not necessarily deal directly with greenhouse gas emissions (Anonymous #2). Since the program team was the bridge to partners, this meant no word was passed on to companies participating in the program until September 2010 when they received the termination letter. That was just one month before the annual meeting. Former partners were so surprised and outraged at the

untimely notification that they attested to keeping their plane tickets to the meeting just so they could "yell at EPA in person."

Why Was Climate Leaders Canceled and What Were the Effects

The end of Climate Leaders was met with frustration and confusion by many corporate partners that had invested in meeting program requirements. (See Appendix B, Quote D). EPA gave one explanation for its decision to end the program: a desire to re-allocate resources where they were more needed and could be more efficiently used. What went unspoken were the reasons for the underlying change in perspective and shift in priorities. From interviews, several insights came to light about the consequences of a leadership change at the top, and many ways in which the program termination could have been handled better.

Several anonymous sources reported their impression that with the change from Bush's administration to Obama's, there was a new emphasis on environmentalism and reinvigoration of EPA. Climate Leaders had been considered by some as too industry friendly, and they demanded a more equally beneficial symbiosis between the public and private sector (Dunn, 2011). To the new administration, addressing this complaint meant empowering EPA to use the "stick" of regulation instead of the "carrot" of voluntary programs. The subsequent dismantling of federal environmental voluntary programs was interpreted by many companies as a desire by the new administration to show a new "tougher" stance on environmental issues. Unfortunately, turning to a more traditional regulatory approach also meant sacrificing the many benefits of VEPs. These ranged from lower administrative costs than regulatory alternatives, to stronger relationships (and knowledge sharing) between industry and government. According to some interviewees,

Climate Leaders was much better than regulation at gather information about GHG emissions on behalf of government and industry, and with the right amount of effort. These former partners claimed that the new regulation will be more focused on catching offenders than working with companies to improve their practices. This shift will be difficult for large companies as they can be cumbersome and have trouble adjusting to the new rules. It will also be difficult for small companies since they have limited resources to help them adapt to such changes.

This study yielded several ways that program termination process could have been improved to ease this transition. One was by earlier and clearer communication to partners about decisions that would directly affect the private sector. A second was even more fundamental: incorporating an "exit strategy," or planned ending, into program design from the beginning. Climate Leaders was developed by EPA, an agency that mostly focuses on a regulatory approach that is usually long term. VEPs like Climate Leaders, however, are public-private partnerships; something altogether different from traditional mandatory rules and enforcement methods. Instead, VEPs are approached more like business ventures. In the business world, such undertakings would have a startup phase, a money-making phase, and then an exit strategy. EPA would do well to follow this model for VEPs. They should plan for an initial design and "gear up" phase, a performance metric-gathering phase, and an exit strategy (Dunn, 2011). This would give program implementers EPA the time and ability to adequately prepare for the end, and to help partners do the same.

Existing Lessons from EPA VEPs

EPA experience with other VEPs yielded valuable lessons which the Agency applied to Climate Leaders with varying degrees of success. First, EPA learned how important it was to protect and elevate its programs' "brand name." If a VEP is backed up by rigorous and

successful environmental stewardship, the public will have positive associations with the program's name. Company membership in that program will then carry more value, increasing participation. (See Appendix B, Quote A). One example of successful branding is ENERGY STAR, arguably EPA's most successful VEP. ENERGY STAR encourages energy efficient buildings and products by offering partners (participants) a label for performance. The program has developed a reputation for results among consumers, so membership became highly coveted by industry. Over time, it evolved an intricate hierarchy of merits that culminates in a yearly ENERGY STAR Awards night. Today, the ENERGY STAR label is recognized by 84% of Americans according to EPA statistics, and is a symbol for financial savings through energy efficiency. In the case of Climate Leaders, companies that wanted to turn a new "green leaf" for the benefit of the environment or to impress consumers were able to associate with EPA through their membership. Former partners reported that learning how to be a greener company in their manufacturing operations helped secure them more contracts and make their employees happier. They felt this offered them a leg above their competition (Anonymous #2, 2011). Climate Leaders did celebrate partner achievements and provided public recognition, which helped establish its brand power. However, it was not as successful as ENERGY STAR at marketing itself as a premium "brand name" program.

The brand name idea was leveraged even less by EPA's Performance Track. This flagship VEP encouraged general environmental stewardship, lumping together high and low-achieving participants. Unfortunately, this lack of merit distinction degraded the value of membership in the program and in 2010, the program was ultimately terminated. Performance Track did get many things right, like relying heavily on Regional-level

involvement. This is the second lesson that EPA should have learned and applied; Climate Leaders would have done well to branch out of headquarters in this way. Unfortunately, its staff was small and centralized at EPA headquarters with minimal outreach to Regional offices (Jones, 2011). Still, with only 4 full-time staff members handling Climate Leaders implementation at any given time, it was a remarkably effective and well-regarded initiative until it was ended in 2010 (Anonymous #2, 2011).

CHAPTER FOUR
Discussion of Study

Format

The primary research in this study was a "process" and "outcome" evaluation of EPA's Climate Leaders Program. (See Appendix A, Table 2 for Overview of the Evaluation Process). Progress and impact of the VEP were analyzed through a literature review and interviews. Recording and presenting results in a qualitative form allowed for richer descriptions about program activities, context, participant behaviors, and stakeholder experiences than would have a quantitative approach (Bond, 1997). The information harvested was in a format that was geared towards policy makers and program implementers, to help them inform ongoing and future EPA programs. It was analyzed and sorted into recommendations meant to help standardize program expectations and design for a more holistic approach to environmental stewardship on a national—someday international—scale.

Objectives

To contribute to the field of environmental program evaluations, the recommendations emanating from this study are meant to do two things. First and foremost, the recommendations advance a comprehensive guiding framework for VEPs that will require combining the efforts of program evaluators from many disciplines, assessing all stages of program lifecycle. Today's program evaluation measures the success of a program, communicates its strengths and advantages, identifies potential improvements, and ideally redirects resources to ensure effective use (Pumphry, 1993). It can facilitate program management by helping identify particularly successful aspects of projects and programs, so managers can choose to divert resources to these more productive activities (Bond, 1997). Evaluation may also identify external or exogenous factors that hinder program success. The

guiding framework being advanced by this study will go a step further than the traditional program evaluations. It will assist environmental organizations and agencies in creating more efficient and effective VEPs, and thus avoiding premature cancellation of voluntary programs and squandering of resources. On a larger scale, however, it will actually help enhance existing programs by identifying redundancies, fill in gaps, and increase the overall efficiency of agencies that will no longer be forced to evaluate their programs only retrospectively. Due to its collaborative nature, the finished guiding framework will have been be vetted by all stakeholders including industry, government, and environmentalists to add legitimacy. It will model all of the key features common to successful VEPs from inception to completion, and will serve as a foundation, manual and modifiable blueprint. The framework will ultimately contribute to a more holistic approach to VEPs by which agencies and organizations can improve design and process proactively. Such progress will move the entire field of environmental program evaluations forward into a new era.

Second, this study sheds light on the shortcomings of VEPs like Climate Leaders, and help ensure that those under construction are built more soundly. While observations were based on a single program, recommendations may also apply to other voluntary program "bundles" (like the Sustainability Partnership in Region 3) and similar voluntary schemes in other regions.

CHAPTER FIVE
Recommendations

Results of Climate Leaders Evaluation

The standardized interview used in this project was meant to determine whether an interviewee felt that Climate Leaders was successful overall, and why or why not. The questions were intended to draw out specific characteristics of a VEP that contributed to this impression, but the open-ended design of the questions allowed space for elaboration into unplanned areas of insight. The information from this analysis was originally expected to yield very specific information about the structure of a successful VEP:

- ideal size (number of participant entities from the public)
- scope (types of organizations participating)
- assistance/support system (provided to participants)
- time-line (of program overall, and for benchmark goals throughout)
- management style (used by EPA's implementation staff)
- incentive structure (for participant involvement).

However, the research and interviews did not produce such specific targeted answers. This is because traits like ideal scope and size depend upon the program's unique environmental arena and targets therein. It turned out that VEP structures varied depending on whether they are tackling environmental challenges of local, regional, national or even global scale. Instead, the interviews in this project yielded more basic information about the core qualities of a successful program. (See Appendix C, Chart 4 for Climate Leaders Evaluation Results).

These design, process, and implementation results revealed lessons and warnings for ongoing and future programs. Answers to research questions numbers one through three focus on Climate Leaders and other similar ongoing VEPs. These are more likely to have narrow applicability: federal VEPs addressing one specific environmental problem. The other answers (to questions three through seven below) contain recommendations that will be

helpful more broadly: to any organization running VEPs, with no specific scope and size. By comparing and contrasting the responses of partners and implementers, many common trends and sentiments were observed as well as a handful of differences. When compiled, they answered the main research questions posed in Chapter 1:

1. ***How do you define success for the Climate Leaders program?***

   Success for Climate Leaders was defined by its ability to satisfy both industry (partners) and government/environmentalists (EPA). From the perspective of the private sector, VEPs are considered successful if they offer the company net savings as opposed to net loss for participant organizations (Futran, 2010). For EPA, success means using taxpayers' money efficiently to protect the environment and human health. This ultimately requires achieving a deep enough bend in the greenhouse gas emissions trajectory to make a positive impact on climate change.

2. ***In what ways did Climate Leaders succeed?***

   From the partners' perspective, Climate Leaders was successful because they received:

   - free technical assistant
   - easy-to-use tools and structure to inventory GHG, set goals, and reach them
   - standardized protocol for calculating/managing GHG
   - simplified federal rules surrounding GHG emissions (and preparation for pending regulation)
   - a relationship with EPA
   - significant knowledge-sharing among partners
   - instant credibility associated with EPA programs, which made the VEP a better value proposition in making the case for involvement to higher management
   - benefits to most partner companies' bottom lines
   - public recognition for achievements
   - relatively short-term payback (financial and/or reputational) on investments

   From EPA's perspective, the success of Climate Leaders was reached because it:

   - did a lot on a small budget (high level program with small staff)
   - helped make climate change an issue that the business world understood and addressed

- high number of engaged participants representing a significant portion of US GHG emissions
- helped ensure emission reduction measures were reported accurately
- set program goals high enough to merit public respect, but achievable enough for companies to agree to public disclosure and inventory preparation
- encouraged ambitious partner goals and high partner investment in abatement efforts
- allowed EPA to engage climate change at some level of during an administration that was reticent to regulate industry
- contributed to significant GHG emissions reduction on a national scale

3. How could Climate Leaders have been improved?

- Progressively higher membership standards so that the program maintains credibility, and recognition is more meaningful
    - Moving from normalized goals to absolute goals sooner
    - Differentiation between high and low participant achievements
- A greater line of communication open for partner feedback
    - Meeting more often to bolster the government/industry relations
    - Faster response time from consultants
- Better outreach to recalcitrant companies
- More value added for complex manufacturing firms (as opposed to catering more to commercial firms)
- Greater EPA headquarter outreach to EPA regional offices
- More dynamic annual analysis to determine what collected metrics say about program progress and results
- Plan for the entire life of program mapped out up front and shared with partners, including milestones, ending, and replacement initiatives

4. What circumstantial conditions (e.g. political, economic) are important to the success/failure of VEPs?

- Evidence of a major/urgent environmental problem needing to be addressed (e.g. climate change)
- Likelihood of pending environmental regulation
    - VEP that addresses anticipated requirements
    - Relationship of VEP to agency creating pending regulation
- Administrators that support voluntary approaches
    - Sufficient federal funding for environmental initiatives
    - Sufficient staff dedicated to the VEP
- A market receptive to green companies, including both consumers and executives
- Unexpected termination by higher management

5. ***What inbuilt features (e.g. structural, procedural) are important to the success/failure of VEPs?***

- Strong leadership
  - Credibility of organization or agency implementing
  - Well-trained implementation staff
- Soundly structured program
  - Well-known "gold standard" protocol as foundation for program (e.g. GHG protocol in Climate Leaders)
  - Sufficient resources (staff, funds, expertise, time, etc.)
- Large/growing membership of participants representing major portion of problem (emission, pollution, etc.)
- Commitment to significant environmental goals
- Progress in meeting goals
  - Partner cooperation and knowledge sharing
- Public recognition and promotions
- Clearly articulated expectations
  - Evolution of goals to continue pushing progress
  - Exit strategy

6. ***What lessons learned from Climate Leaders can be applied to ongoing EPA VEPs at a federal and regional level?***

It is critical to collect data throughout the life of the program, and judge progress based on multiple criteria. These should focus on showing how EPA is managing its resources, and also whether partners are showing a return on investments (Dunn, 2011).

Data should also be analyzed for unexpected outcomes. "Interaction effects" can yield unwanted results, as was the case when EPA's ENERGY STAR began labeling energy efficient refrigerators: instead of reducing use, the VEP ended up increasing demand as consumers began purchasing new refrigerators without getting rid of their old ones. The effect was to unexpectedly *increase* overall national energy use instead of reducing consumption (Gillingham, 2006). In the case of Climate Leaders, the economic concern is that energy savings from cut emissions might result in a reallocation of expenditures in another area. Such unintended results can weaken the environmental contributions of an otherwise strong program.

Climate Leaders was *least* successful in one key area: program termination. The VEP's end was not pre-designed or well-managed by implementers. It was terminated from the top down, without adequate notice and input for implementers and safety nets for participants. As a result, partners viewed the termination as premature, unexpected, and in some cases infuriating. This affects confidence in EPA and VEPs in general, which could damage similar efforts in the future. (See Appendix B, Quote C).

7. What improvements are not accounted for in existing frameworks that must be added for comprehensibility in evaluating program performance?

There has been a rise in evaluations that are part of a program from the beginning called *formative* evaluations. These are not separate from a project or simply amended to it as an afterthought, but are integrated throughout. In the past, evaluations were often used to measure performance based only on information gathered in a final, *summative* act. Such retrospective evaluations were too late to improve the program being evaluated. In formative assessments, planning, evaluation, and implementation are all parts of a whole working together concert. Research has shown that by interacting, they yield the highest quality programs (Westat, 2002). (See Appendix A, Table 3 for example of formative Project Development and Evaluation Cycle).

In an ideal VEP, data analysis is built into the design periodically to examine absolute trends, and compare these to hypothetical scenarios in which the program had not existed. This data analysis would help program implementers establish sector benchmarks based on their unique characteristics. The individualized point of reference would set a foundation from which to negotiate program goals with each partner company. Such a data analysis and benchmarking process is already underway by industry, as described by Tonkonogy (2007).

Finally, two tests should be included in the framework to test VEPs during the design and implementation stages. First, a "Procedural Test" is needed to ensure inclusive and responsible participation in a VEP. Programs should include industry, government and environmental groups to build accountability for the actions suggested and taken. Second, a "Substantive Test" is needed to guarantee validity and appropriateness of concerns raised by participants (Steelman, 2006).

Key Lesson

There was one lesson for organizations and agencies running VEPs that stood out among the rest: the need for a combination of regulatory and voluntary approaches. In many cases, the discussion about federal environmental strategy has become a battle pitting the two approaches against each other. This argument is addressing a false question. Research shows that neither management style is superior, but rather that *each is weakened without the existence of the other* (Bennear, 2007). In other words, both regulatory and voluntary strategies play a necessary role in effective environmental protection on a national level.

On one hand, voluntary programs are less effective without the backdrop of mandatory regulation than when launched alongside harder laws (Khanna, 1999). That is because mandatory regulation—even pending regulation—increases the incentive for companies to participate in voluntary initiatives. Khanna and Damon (1999) explained that "participation in voluntary programs depends to a large extent on the existence of a regulatory framework that would impose penalties on firms that do not undertake proactive measures for self-regulation." Over a decade later, Delmas (2010) came to the same conclusion, citing the difficulty involved in using voluntary schemes to induce improved environmental outcomes where no sanctioning mechanisms exist.

On the other hand, voluntary programs offer the tools companies need to meet evolving laws and rules. Climate Leaders is a perfect example, offering free help for companies to get their emissions under control and avoid regulatory penalties. EPA would benefit from a balance of the voluntary "carrot" and regulatory "stick" approaches, rather than sacrificing one for the other. Instead of debating which policy instrument is better, efforts should be put towards finding the most effective and efficient combination of VEPs and regulation. Together however, they provide a proper balance of strength and flexibility.

Limitations

In this project, EPA's Climate Leader program was evaluated as a case study. Voluntary environmental programs literature was reviewed, evaluation theory was synthesized, and lessons were gathered for incorporation into a guiding framework for VEPs. The evaluation was used to filter out success standards for Climate Leaders specifically, and to contribute to a guiding framework for VEPs at large. This process was not without challenges, however.

This analysis focused mostly on one program, Climate Leaders, so the results were not necessarily definitive, comprehensive, nor generalizable to programs of different characteristics (size, scope, aim, region, etc.). The evaluation selectively incorporated attributes of Environmental Program Evaluation deemed most fitting to Climate Leaders, which may have limited the applicability of the resulting recommendations; they are most relevant to American VEPs with origins at the national level. Additional research is needed to confirm whether the trends are reliable, and whether they apply more widely.

It is important to mention other limitations inherent in these research methods. One that could not be overcome was that the limited literature on how to best design, implement,

and evaluate a VEP was very diverse across several dimensions. First, the surveyed literature organizes program categories across breakdowns of different characteristics, and does so to varying degrees of granulation. Second, studies attribute success or failure of a program to unique factors, including characteristics of target participants, operations of program, or incentives offered by the program. They also group those factors across varying dimensions. Third, regional level programs may be different from national programs in their partnership base or employee performance. It was impossible to reconcile these different evaluation strategies and variables perfectly, so at times a judgment call is made when deciding which to include.

Finally, even once all of the necessary information to build a comprehensive guiding framework for VEPs is compiled, the framework itself will be a challenge to draw up. This is because of the nature of program diversity; a tool to guide VEP design and implementation must be broad enough to encompass the many kinds in existence, but specific enough to capture the qualities that make them each distinct. Standardizing VEPs in the future will make it possible to label programs to distinguish among the different types (Darnall, 2005).

Future Research

Future work in this area should collect similar data on other voluntary environmental programs from EPA's headquarters and regional offices, as well as other agencies (e.g. the Department of Energy) and organizations (e.g. the World Resources Institute). Programs that were terminated would make informative case studies by providing lessons learned, while successful ongoing programs (or those terminated on schedule) would contribute best practices. The more program evaluation results compiled, the more robust the guiding framework for VEPs can become. At EPA, data collection should be standardized and

analyzed across programs to be consistent and provide office and agency-wide information about program organization, size, and achievements (Boyd & Manson, 2011). EPA offices should develop standardized demand measures that begin with the following (Boyd & Manson, 2011):

- participation rates (number of partners);
- duration and consistency of participation (retention of partners);
- public voluntary program resource usage rates (hotline calls, conference participation, and guidance downloads and requests); and
- Evidence of repeat demand

Climate Leaders was part of EPA's Region 3 office "bundle" called the Sustainability Partnership, so once a guiding framework for VEPs is relatively robust, future research should test it for applicability to "bundled" voluntary programs: initiatives which offer guidance and access to multiple VEPs addressing different environmental issues all at one source. In addition to Region 3's Sustainability Partnership, these might include Region 7's (Kansas City) 2010 partnership bundles, and Region 2's (New York City) Green Team. Eventually this research can be expanded internationally to test and tweak the limits of the framework's applicability.

Conclusion

A comprehensive guiding framework for voluntary environmental programs will serve to arm agencies like EPA against blotched programs and squandered resources. It will allow future voluntary environmental programs to be designed with consistency, executed according to industry-vetted practices, and terminated responsibly. When enough—eventually all—VEPs are assessed at every stage, they will be at peak effectiveness. However, to reach this point, the guiding framework for VEPs must be rounded out by additional evaluations. This study provided one such evaluation of the Climate Leaders

program, and recommendations for its improvement. Similar evaluations and perspectives from other fields of social science will help create a robust and inclusive tool. Once complete, the guiding framework will help fine-tune VEPs according to standards vetted through evaluations, insuring they are designed with the highest likelihood of success. This will be a major challenge to academia, but will be a worthwhile endeavor. Such efficiency and effectiveness is vital in addressing issues that can be critically time-sensitive including air and water pollution, biodiversity decline, climate change, water quality and allocation, resource and waste management, and energy efficiency.

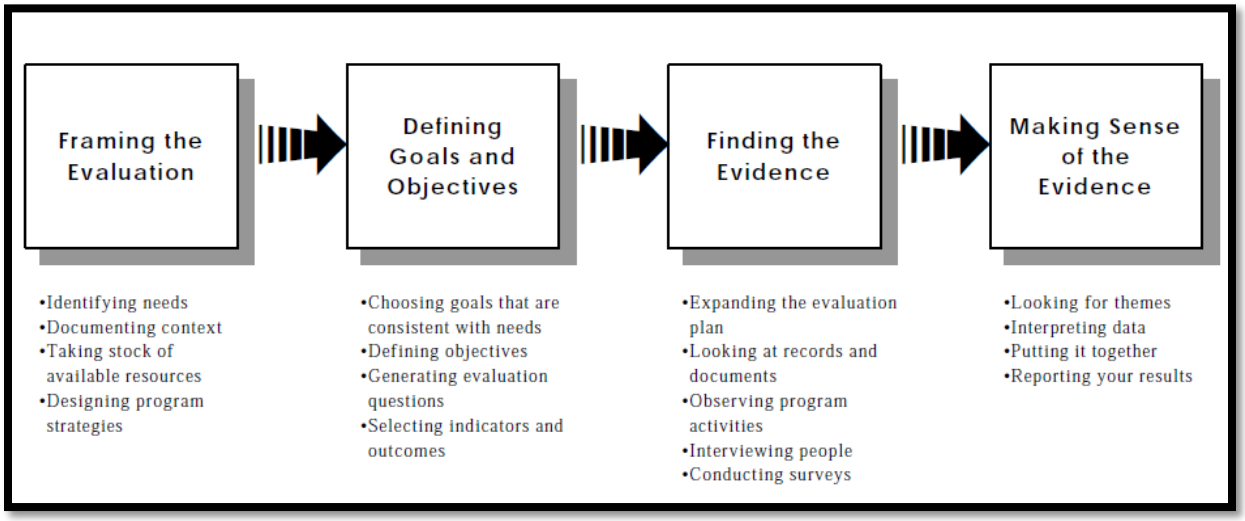
APPENDICES

Appendix A

Table 1: Summary of Data Collection Methods

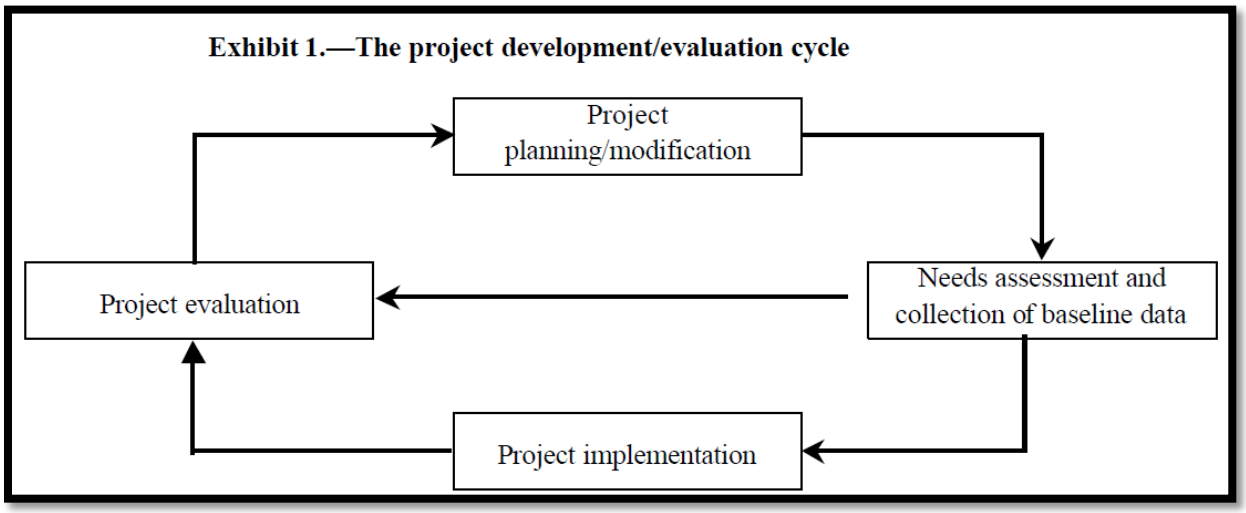
Research Question	Data Collection Methods	
	Literature Review	Interview
How do you define success for the Climate Leaders program?		X
In what ways did Climate Leaders succeed?	X	X
How could Climate Leaders have been improved?		X
What circumstantial conditions (e.g. economical, political) are important to the success/failure of VEPs?	X	X
What inbuilt features are important to the success/failure of VEPs? (e.g. structurally/procedurally)?	X	X
What lessons learned from Climate Leaders can be applied to ongoing EPA VEPs on a federal and regional level?		X
What improvements are not accounted for in existing frameworks that must be added for comprehensibility in evaluating program performance?	X	X

Table 2: Overview of the Evaluation Process



(Bond, 1997)

Table 3: Project Development/Evaluation Cycle



(Westat, 2002)

Appendix B

Sample Quotes by Interviewees

A. "EPA's blessing gave a program like Climate Leaders the status of a 'name-brand.' EPA didn't realize value attributed by companies to EPA's processes— inventories, numbers and goals. In some cases, that name-brand was premier benefit."

–Former Climate Leaders partner company representative

B. "EPA programs are held to the highest standard because they force participants to actually work for their partnership credentials. Currently, there are too many different organizations and non-profits pulling companies in so many directions. It would be nice if Climate Leaders provided the undisputed structure and calculations to address greenhouse gas emissions so we could focus our efforts. If EPA was the leadership standard, that would be success. Climate Leaders was an absolute success until it was terminated."

–Former Climate Leaders partner company representative

C. "I've heard a lot of people say, especially around October 'funeral meeting,' that this is last time they will sign up for an EPA program. These were people who had invested personal credibility, telling their company to get involved despite the work it would take. They had a stake in the program when, all of the sudden, EPA cut their legs out. They ended up thinking to themselves, 'Fool me once. I'm not signing up for one of these things again." There was more damage done in this way—the way the program ending was handled—than EPA may have realized."

–Former Climate Leaders partner company representative

D. "I don't think EPA saw how highly valued certain aspects of Climate Leaders were. This should be cautionary tale for EPA: don't start a program if you are not prepared to articulate that is a 'sun-setting' initiative with specific objectives that will come to an end, or have a plan to keep a program alive to ensure whoever signs up and helps make it successful doesn't feel abandoned."

–Former Climate Leaders partner company representative

E. "I saw work of Climate Leaders not as focus on game changing so much, but on systematic process of getting companies to be aware of their emissions and institute highly integrated plans to slowly and surely reduce them over time. It was an 'evolutionary' versus 'revolutionary' type of program."

–Former EPA Climate Leaders promoter

Appendix C

Chart 1: Three-Pronged Approach to Program Evaluation

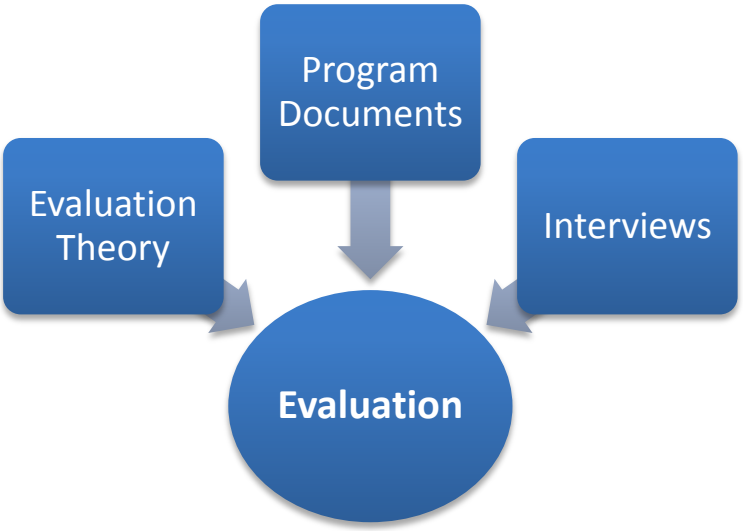


Chart 2: Three Steps to Climate Leaders Evaluation

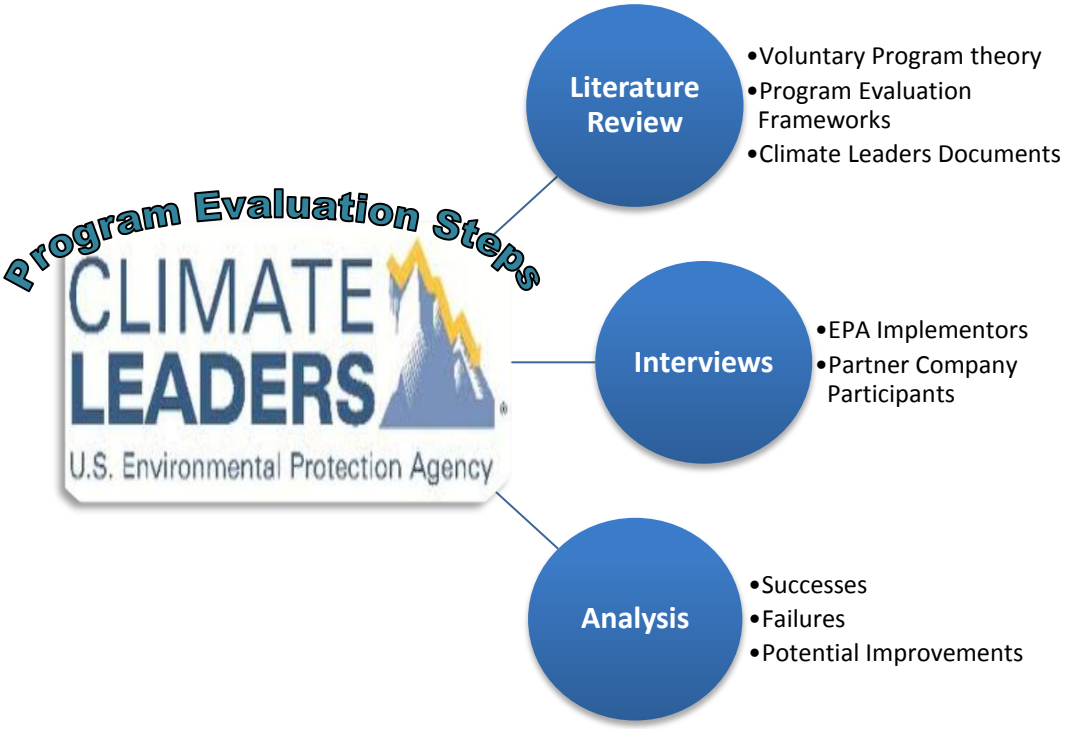


Chart 3: Formal Interviews Breakdown

Chart 4: Climate Leaders Evaluation Results

Program Strengths	Program Weaknesses
Well organized: clear expectations and instructions	Lacked predetermined termination date
Aggressive goals and rigorous requirements	Allowed excessive flexibility and lag-time on early goals
Committed and skilled staff	Consultants not intimately familiar with industries
Free data help, standardized tools, and framework	No external auditing services offered, only EPA audits
Free technical expertise	Assistance too individualized and inefficient overall
Public recognition of partner success	More differentiation among partner achievements
High standard for membership	Goals did not evolve in stride with growing program
High partner traction/numbers	Partners were mostly self-selected high achievers
Strong partner commitment/engagement	Goal-setting minimums not raised high enough
Industry sharing outside of competitive arena	No replacement program to ensure efforts continued
Feeling of cooperation/"family" among partners	Insufficient number of meetings and conference calls
Gave GHG management in business more value	Unsustainable pace/expansion and understaffed
Partners more prepared for GHG regulation	Ended before US GHG regulation went into effect
National-scale emissions awareness and reduction	Sudden end undermined individual/company investment
Showed how much VEPs could accomplish	Method and timing of termination reduced faith in VEPs
Work was transparent and robust	Analysis/utilization of data/results insufficient

WORKS CITED

Alberini, A., & Segerson, K. (2002). Assessing voluntary programs to improve environmental quality. *Environmental and Resource Economics*, 22(1-2), 157.

Anonymous #1. (2011), Personal communication, Environmental Protection Agency, 2 March.

Anonymous #2. (2011), Personal communication, Environmental Protection Agency, 21 March.

Arora, S., & Cason, T. N. (1995). An experiment in voluntary environmental regulation: Participation in EPA′s 33/50 program. *Journal of Environmental Economics and Management*, 28(3), 271-286.

Arora, S., & Cason, T. N. (1996). Why do firms volunteer to exceed environmental regulations? Understanding participation in EPA′s 33/50 program. *Land Economics*, 72(4), 413.

Bennear, L.S. & Stavins, R.N. (2007), Secon-best theory and the use of multiple policy instruments. *Environmental Resource Economics*, 37(1), 111-129.

Bond, S.L., Boyd, S.E., & Rapp, K.A. (1997). Taking stock: A practical guide to evaluating your own programs. NC: Horizon Research, Inc.

Borck, J. C. & Coglianese, C. (2009). Voluntary environmental programs: Assessing their effectiveness. *Annual Review of Environment & Resources*, 34(1), 305-324.

Boyd, J. & Manson, C. (2011). Attributing benefits to voluntary programs in EPA's Office of Resource Conservation and Recovery: Challenges and options. Resources for the Future, March, RFF DP 11-09.

Boyd, J. (2011), EPA Seminar, Assessing and Justifying Voluntary Programs, 28 March.

Daley, D. M. (2007). Voluntary approaches to environmental problems: Exploring the rise of nontraditional public policy. *Policy Studies Journal*, 35(2), 165-180.

Darnall, N. & Carmin J. (2005). Greener and cleaner? The signaling accuracy of U.S. voluntary environmental programs, *Policy Sciences*, 38(2-3), 71.

Delmas, M.A. & Montes-Sancho, M.J. (2010). Voluntary agreements to improve environmental quality: Symbolic and substantive cooperation, *Strategic Management Journal*, 31(6), 575.

Delmas, M. A., & Terlaak, A. K. (2001). A framework for analyzing environmental voluntary agreements. *California Management Review*, 43(3), 44-63.

Delmas, M., & Keller, A. (2005). Free riding in voluntary environmental programs: The case of the U.S. EPA WasteWise program. *Policy Sciences*, 38(2-3), 91.

Dunn, M. (2011), Personal communication, Environmental Protection Agency, 15 February.

Environmental Protection Agency (EPA). (2005). Design Principles: Climate Leaders Green House Gas Inventory Protocol.

Environmental Protection Agency (EPA). (2010). Press Release: EPA to Transition Climate Leaders Program.

Environmental Protection Agency (EPA). (2010). Program Evaluation Glossary. Retrieved 7 Nov 2010, 2010, from http://www.epa.gov/evaluate/glossary/m-esd.htm.

Futran, V.J. (2010). Voluntary environmental program success correlated with savings: Economic considerations and implications for firms in EPA's 33-50 Program (Working Paper), University of Pennsylvania.

Gillingham, K., Newell, R., Palmer, K. (2006). Energy efficient policies: A retrospective examination. *Annu. Rev. Environ. Resour.*, 31, 161-96.

Hassel, S., Clancy, N., Burger, N., Nelson, C., Rudavsky, R., Olmstead, O. (2010). An Assessment of the U.S. Environmental Protection Agency's National Environmental Performance Track Program, RAND Corporation.

Jones, L. (2011), Personal communication, Environmental Protection Agency, 31 March.

Khanna M. & Damon L.A. (1999). EPA's Voluntary 33/50 Program: Impact on toxic releases and economic performance of firms, *Journal of Environmental Economics and Management*, 37, 1-25.

Klassen, R.D., & McLaughlin, C.P. (1996). The impact of environmental management on firm performance. *Management Science*. 42 (8), 1199.

Knaap, J. G., & Kim, T. J. (Eds.). (1998). Environmental program evaluation: A primer. Urbana: University of Illinois Press.

Laskowski, S.L. (2010), Personal communication, 1 December.

Lyon, T. P., & Maxwell, J. W. (2002). Economic institutions and environmental policy. Ashgate Publishing Ltd.

McCarthy, G. (2010). Climate Leaders Termination Letter, Environmental Protection Agency (EPA), Office of Air and Radiation, September.

Pizer, W. A., Mogenstern, R., & Shih, J. (2008). Evaluating voluntary climate programs in the United States. Resources for the Future, July, RFF DP 08-13.

Pumphrey, B. (1993). Planning and implementing program evaluations within the Office of Solid Waste: Program Evaluation Network, Environmental Protection Agency's Government Performance and Results Act (GPRA), United States Senate, 32.

Seema, A. & Timothy C. N. (1996). Why do firms volunteer to exceed environmental regulations? *Land Economics*, 72 (4), 413.

Segerson, K., & Miceli, T. J. (1998). Voluntary environmental agreements: Good or bad news for environmental protection? *Journal of Environmental Economics and Management*, 36(2), 109-130.

Steelman, T.A. & Rivera, J. (2006). Voluntary environmental programs in the United States: Whose interests. *Organization & Environment*, 19(4), 505.

Tonkonogy, B., Sullivan, J., & Norris, G.A. (2007). Evaluating corporate climate performance: A model for benchmarking GHG reductions. *American Council for an Energy Efficient Economy*, Summer Study Paper.

Videras, J. & Alberini, A. (2000). The appeal of voluntary environmental programs: Which firms participate and why, *Contemporary Economic Policy*, 18 (4), 449.

Westat, J.F. (2002). The 2002 user-friendly handbook for project evaluation, National Science Foundation (NSF), Directorate for Education and Human Resources: Division of Research, Evaluation and Communication, January.

Zatz, M., Harbour S. (1999). The United States Environmental Protection Agency's 33/50 Program: The Anatomy of a Successful Voluntary Pollution Reduction Program, *Journal of Cleaner Production*, 7 (1999), 17–26.