DISCUSSION PAPER

July 2008 ■ RFF DP 08-13

Evaluating Voluntary Climate Programs in the United States

William A. Pizer, Richard Morgenstern, and Jhih-Shyang Shih

1616 P St. NW
Washington, DC 20036
202-328-5000 www.rff.org

RESOURCES FOR THE FUTURE

Evaluating Voluntary Climate Programs in the United States

William A. Pizer, Richard Morgenstern, and Jhih-Shyang Shih

Abstract

Despite the growing importance of voluntary programs as tools for environmental management, they have been subject to quite limited evaluation. Program evaluation in the absence of randomized experiments is difficult because the decision to participate may not be random and, in particular, may be correlated with the outcomes. The present study is designed to overcome these problems by gauging he environmental effectiveness of two voluntary climate change programs—the U.S. Environmental Protection Agency's Climate Wise program and the U.S. Department of Energy's Voluntary Reporting of Greenhouse Gases Program, or 1605(b)—with particular attention to the participation decision and how various assumptions affect estimates of program outcomes. For both programs, the analysis focuses on manufacturing firms and uses confidential census data to create a comparison group and to measure outcomes (expenditures on fuel and electricity).

Overall, we find that that the effects from Climate Wise and 1605(b) on fuel and electricity expenditures are no more than 10 percent and probably less than 5 percent. Virtually no evidence suggests a statistically significant effect of either Climate Wise or 1605(b) on fuel costs. Some evidence suggests that participation in Climate Wise led to a slight (3–5 percent) increase in electricity costs that vanished after two years. Stronger evidence suggests that participation in 1605(b) led to a slight (4–8 percent) decrease in electricity costs that persisted for at least three years.

Key Words: voluntary, regulation, energy, climate change

JEL Classification Numbers: Q2, Q4

© 2008 Resources for the Future. All rights reserved. No portion of this paper may be reproduced without permission of the authors.

Discussion papers are research materials circulated by their authors for purposes of information and discussion. They have not necessarily undergone formal peer review.

Contents

Introduction **1**

Background **3**

Evaluating Voluntary Programs 4

The Climate Wise Program 8

The 1605(b) Program 10

Data **11**

The Climate Wise Program 11

The 1605(b) Program 12

Models and Econometric Method **13**

Results **18**

Conclusions **21**

References **24**

Tables **27**

Evaluating Voluntary Climate Programs in the United States

William A. Pizer, Richard Morgenstern, and Jhih-Shyang Shih*

Introduction

Voluntary programs have been a key part of U.S. climate change policy since the early 1990s. Such programs formed the centerpiece of President Clinton's 1993 Climate Change Action Plan, which included Energy Star, Rebuild America, Green Lights, Motor Challenge, the Voluntary Reporting of Greenhouse Gases Program (required under Section 1605(b) of the Energy Policy Act of 1992, or EPAct), and Climate Wise. A number of these programs were initiated in the George W. Bush administration. More recently, voluntary programs have figured prominently in President Bush's 2002 climate change policy announcement, referencing agreements with the semiconductor and aluminum industries and leading to the creation of the Climate Leaders and Climate Vision programs (White House 2002, 2005).

A 2005 survey identified 87 voluntary programs administered by the U.S. Environmental Protection Agency (EPA), up from 54 in 1999 and 28 in 1996 (U.S. EPA 2005). In fiscal year 2006, voluntary programs comprised 1.6 percent of EPA's operating budget. Dozens more programs are operated by the U.S. Department of Energy (DOE) and other federal agencies and at the state level. Although many of these programs focus on climate change and energy, others cover waste, water, toxics, and agriculture. Voluntary programs have figured prominently in the national climate policies of other countries as well; for example, they continue to play a leading role in Japan.

The explosive growth in voluntary environmental programs may reflect changing societal attitudes about the environment and a growing optimism regarding the possibility of enhanced cooperation between government and business. It may also reflect widespread frustration with the long and expensive battles often associated with new environmental regulations. In most cases, voluntary programs are being used to control pollutants that have not yet been regulated and for which legislative authority may be difficult to obtain. In contrast to market-based approaches to environmental management, in which the conceptual roots are largely academic,

* Authors are Senior Fellow, Senior Fellow, and Fellow, respectively, at Resources for the Future.

voluntary programs have emerged as a pragmatic response to the need for more flexible ways to protect the environment.

The prominence of voluntary programs begs the question of whether these programs are effective in achieving their goals. For example, following the 2002 announcement by President Bush of voluntary efforts to achieve an 18 percent improvement in greenhouse gas (GHG) intensity by 2012, recent data indicating that we are on track to meet that 18 percent goal have been cited as evidence of the voluntary program's success (White House 2005). But are the data really evidence of the voluntary program's success, or do they simply reflect other coincident events?

The key issue we consider in this study is whether these programs actually work as advertised. That is, do voluntary programs deliver the promise of significant environmental gains without the burdens associated with mandatory regulation? Do they improve environmental and conservation outcomes relative to a realistic baseline, or do they pave the way for other actions that do so? Quantitatively, how large are the likely gains? Can such approaches serve as a substitute for mandatory requirements, or should only modest gains be expected from these efforts?

We address these questions by conducting a detailed analysis of two early voluntary programs: EPA's Climate Wise program and DOE's Voluntary Reporting of Greenhouse Gases Program, or 1605(b). Although they are not the most recent voluntary initiatives and did not benefit from potential improvements in voluntary program design that occurred over the past decade, the relatively long histories of these programs make them particularly amenable to statistical analysis. Whereas 1605(b) is principally a reporting program, Climate Wise offered more "carrots" in the form of technical assistance and greater public acknowledgement. Climate Wise also imposed more "sticks" in the sense of an expectation that participating firms will make larger emissions reductions. Climate Wise applied entirely to firms other than electric utilities, whereas 1605(b) is open to a broader range of entities. However, because of the nature of our matching sample, described below, we limit our analysis of both Climate Wise and 1605(b) to firms in the manufacturing sector.

We focus on the environmental effectiveness of these programs, with particular emphasis on the participation decision and how various assumptions affect estimates of program outcomes. We consider two alternative approaches to evaluating outcomes, while attempting to control for self-selection in joining the voluntary programs. In one approach, derived from Heckman and Hotz (1985), we consider a model in which program participation depends on both observed and

unobservable variables that may be correlated with the outcome. In the other approach, propensity score matching, we match participants to appropriate nonparticipants and consider pairwise differences. As part of our effort to develop a credible baseline, we are fortunate to have access to confidential plant-level data files for the manufacturing sector collected by the U.S. Census Bureau.

Understanding the true effectiveness of these programs is important. Proponents and opponents of the trend toward voluntary approaches are increasingly at odds, sometimes drawing opposite conclusions about the same program. Proponents, typically on the side of industry, see voluntary programs as a more practical, flexible approach to achieving environmental gains. Opponents, including some environmental advocates, often see voluntary programs as an obstacle to more stringent, mandatory programs. This polarization may be partly a consequence of poor information. Although intuition and anecdotes may provide some reason for believing that a given program has or has not had a beneficial environmental impact, careful empirical analysis with peer review is much more convincing. The goal of this research is to help fill that void and inform decisionmakers.

Background

In principle, voluntary programs offer opportunities for firms to get hands-on experience with new types of environmental problems without the straightjacket of regulation and, in the process, to enhance their environmental reputation with government agencies, customers, investors, communities, employees, and other firms. In some cases, a firm's participation may represent an effort to shape future regulations or to stave off mandatory requirements altogether. Some or all of these benefits may be reflected in the firm's bottom line over the short or long term.

Voluntary programs also provide opportunities for government agencies to gain experience with new problems and new industries. Most importantly, they may provide opportunities to achieve environmental improvements more quickly, and with lower administrative costs, than would otherwise be possible. Sometimes, such programs allow for the use of more holistic approaches than the media-specific, end-of-pipe focus of most existing regulations.

Notwithstanding the many potential benefits of voluntary approaches, the absence of deliberate price or regulatory signals to encourage fundamental changes in corporate or consumer actions or to stimulate demand for cleaner technologies is a clear limitation. In some

cases, the targets established for voluntary programs reflect only a business-as-usual scenario; this is referred to as "regulatory capture." Some voluntary programs may encourage free riding, wherein some firms avoid any effort while other, proactive, firms voluntarily address a problem and keep further regulation at bay. Taking this a step further, voluntary approaches may represent a shift in emphasis from the "worst" polluters to those most willing to abate emissions on their own initiative.

Because business is inherently dynamic, with firms constantly confronting new challenges, opportunities, and technologies, it is not sufficient to simply compare two distinct points in time to see whether firms' environmental performance has improved. Rather, environmental gains must be assessed with reference to a credible representation of what would have happened otherwise. One approach is to construct a business-as-usual forecast using the best available data. However, such an approach is limited by the large number of unpredictable influences on outcomes. An alternative is to compare participants to a suitably chosen group of nonparticipants. Still, biases may arise if participants and nonparticipants differ in some systematic way—for example, if participants are bigger, faster growing, or better managed. Unless the comparisons are carefully constructed, observed differences between participants and nonparticipants may reflect factors other than the effects of the program.

If it were possible, the most transparent way to measure the environmental performance of a voluntary program—or any program—would be to conduct a scientific experiment to see whether firms randomly assigned to the program exhibited outcomes that differed from those randomly assigned to a control group. Because the two groups would be otherwise identical (because of randomization), this would yield an unbiased estimate of the effect of the voluntary program on environmental performance. In reality, we rarely see such randomized experiments and are instead left with either forecast baselines or imperfect control groups, which provide only limited evidence on the environmental performance of participating firms compared to what realistically would have happened otherwise.

Evaluating Voluntary Programs

The literature on voluntary programs uses a variety of descriptors to identify particular designs: self-regulation, negotiated agreements, environmental covenants, business-led environmental strategies, and others. Nonetheless, a loose taxonomy has evolved, with three reasonably distinct bins based on how the parameters of the voluntary action are determined:

unilateral agreements by industrial firms, wherein commitments or reduction targets are chosen by firms or industry associations without significant government inputs; public voluntary programs, in which participating firms agree to protocols that have been developed by environmental agencies or other public bodies; and negotiated agreements, consisting of a target and timetable for attaining the agreed-upon environmental objectives, based on explicit negotiation between government authorities and a firm or industry group over specific terms. Climate Wise was a public voluntary program, as is 1605(b).

Economic analysis suggests that, because environmental mitigation typically is not costless and the benefits are not appropriable by the firm, profit-maximizing firms have little incentive to undertake such activities unless mandated by government to do so. It is not surprising, therefore, that by far the dominant issue in the academic literature on voluntary programs concerns the motivation for firms to participate in the programs. Extensive theoretical and some empirical work has focused on the importance of preempting regulatory threats; the potential to influence future regulations; the effects on stakeholder relations and the firm's public image; the importance (or unimportance) of technical assistance and financial incentives; the economic efficiency of the programs; the role of competitive pressures; and the potential to bring about savings in transaction or compliance costs. Several studies have shown that public recognition of participation in a voluntary program is a key motivation for firms.[1]

The largely theoretical work on environmental performance suggests that participation in voluntary programs does not guarantee an improvement in actual performance. Although it may encourage the exchange of information about best practices, a key result may be the provision of insurance to firms against stakeholder pressure. Thus, by implication, it might be argued that participation in voluntary programs may actually *reduce* incentives to cut emissions if it successfully staves off stakeholder pressure for more stringent actions. Theoretical studies have shown that improvements in actual environmental performance depend on the extent to which voluntary programs lead to lower abatement costs than mandatory regulation; the likelihood that regulation will be imposed if the program is not effective; the extent to which the regulator is willing to subsidize pollution reduction; the willingness of consumers to pay for green products; and other factors.

[1] For reviews of this literature see Khanna (2001) and Lyon and Maxwell (2002); see also Arora and Cason (1995, 1996), Celdren et. al. (1996), and Khanna and Damon (1999).

Some voluntary programs focus on the adoption of particular technologies (e.g., Green Lights, now part of the Energy Star Program), whereas others focus directly on environmental performance (e.g., 33/50, Climate Wise, 1605(b), or various audit-based programs). Technology programs can be difficult to evaluate because of the general absence of comprehensive databases on the performance of facilities that have not adopted the particular technologies.[2]

The empirical evidence is more extensive when we look at programs focused explicitly on environmental performance as opposed to technology adoption, particularly with regard to toxics, for which extensive analyses using Toxics Release Inventory data have been conducted. An in-depth analysis of EPA's 33/50 program, which aimed to reduce certain high-priority toxic chemicals by 33 percent below 1988 levels by 1992 and by 50 percent by 1995, by Khanna and Damon (1999) is probably the gold standard in the field. Like the present study, Khanna and Damon (1999) jointly model the decision to participate in the program and the actual outcomes. They first recognize that a firm's decision about the quantity of covered releases to emit will probably depend on both its participation in 33/50 and such factors as stakeholder pressure, output levels, and others. They then allow for the participation decision to depend on these same variables and to be correlated with the volume of releases. Using publicly available firm-level data, they found a statistically significant impact of the program on toxic releases and on firms' returns on investment and long-run profitability. Khanna and Damon (1999) hypothesize that the incentives for participation arise from three sources: program features, the threat of mandatory environmental regulations, and firm-specific characteristics.

Focusing on the period 1988–1995, Sam and Innes (2005) also found that participation in 33/50 lowered releases of the covered chemicals, particularly in 1992. Further, they found that participation in 33/50 was associated with a significant decline in EPA inspection rates of participating firms for the years 1993–1995. Gamper-Rabindran (2006) found that, although the

[2] A number of these programs have been subject to at least some evaluation. For example, DeCanio (1998) finds that the energy-efficiency investments carried out under the Green Lights program yielded annual real rates of return averaging 45 percent. DeCanio and Watkins (1998) find that specific characteristics of firms affect their decisions to join Green Lights and commit to a program of investments in lighting efficiency. Dowd et al. (2001) cite specific product purchase decisions that are influenced by Energy Star. After reviewing the evidence on Green Lights and Energy Star, Howarth et al. (2000) concluded that "voluntary agreements between government agencies and private sector firms can … lead to improvements in both technical efficiency of energy use and the economic efficiency of resource allocation". Unfortunately, none of these studies was able to distinguish between the improvements attributable to the voluntary programs and those changes that probably would have taken place even without the programs.

effects varied by industry, in the case of the largest participating industry—the chemical industry—the finding by Khanna and Damon (1999) that 33/50 reduced toxic releases is actually reversed when the analysis excludes two ozone-depleting chemicals whose phase-out was mandated by the Clean Air Act.

King and Lenox (2000) analyzed the environmental impact of firms participating in Responsible Care, an industry-sponsored effort to cut toxic releases distinct from the government-sponsored 33/50. Using pooled and panel data for the period 1991–1996, they find, surprisingly, that participants were reducing their releases more slowly than nonparticipants. Their fixed-effect model shows that Responsible Care had an insignificant effect on environmental performance. That is, despite the improved performance of the chemical industry over the studied period, the rate of improvement was not greater than in preprogram years.

A paper by Dasgupta et al. (1997) considered the adoption of ISO 14001 management practices by Mexican firms. They found a significant improvement in the self-reported compliance status of participating firms. They also found that explicit environmental training programs for nonenvironmental workers led to an improvement in the compliance status of the firms.

Turning to energy and climate change, an analysis of the effects of the DOE's Climate Challenge Program on carbon dioxide (CO_2) emissions focused on the largest 50 electric utilities east of the Rocky Mountains from 1995 to 1997 (Welch et al. 2000). Despite a number of intriguing results about the motivation of firms to participate in Climate Challenge, the study finds that adoption of the program seemed to have no significant effect on emissions. The authors hypothesize that the poor program performance is associated with the lack of at least a tacit regulatory threat of the type present in 33/50.

In a recent paper, Lyon and Kim (2007) examined the performance of electric utilities participating in the 1605(b) program. They use a two-stage model to account for both participation and environmental outcomes. The authors find that participants tend to be larger, with higher and more rapidly increasing emissions than nonparticipants. However, they also find that participation had no measurable effect on a firm's carbon intensity. They conclude that participation may be a form of greenwash, that is, an attempt to appear more environmentally friendly than is really the case.

Overall, the literature is characterized by a paucity of empirical studies on the actual environmental performance of voluntary programs and, equally important, an almost exclusive focus on toxics, as opposed to energy- or climate-related programs. Energy issues differ from

toxics in many ways, including the extent to which financial incentives in the form of nonzero energy prices are already in place to reduce emissions. That is, the desire to increase profit through lower energy costs already encourages conservation and energy efficiency, whereas no such forces exist to reduce toxic emissions. Thus, the potential for voluntary programs to achieve further reductions in energy-related CO_2 emissions may be more limited than the potential associated with toxics. A key motivation for this study is to encourage the rigorous analysis of program results and to emphasize the rapidly growing interest in energy- and GHG-related programs.

The Climate Wise Program

Officially established by EPA in 1993, Climate Wise was a performance-based voluntary program focusing on the nonutility industrial sector to encourage the reduction of CO_2 and other GHGs via adoption of energy efficiency, renewable energy, and pollution prevention technologies. Climate Wise remained in operation until 1999–2000, when it was renamed as the Energy Star program and placed under the Agency's Energy Star umbrella. Unlike Green Lights or EPA's other technology-based programs, Climate Wise members had the flexibility to use whatever technologies or strategies they chose to reduce their emissions. At its peak, Climate Wise had enrolled more than 600 industrial firms covering several thousand facilities nationwide. Although EPA has developed estimates of the emissions reductions associated with Climate Wise, there has been little outside evaluation of the program.

As stated in the program's 1998 Progress Report (U.S. EPA 1998, 2), the four broad objectives of the Climate Wise Program were to:

- "Encourage the immediate reduction of greenhouse gas emissions in the industrial sector through a comprehensive set of cost-effective actions;
- "Change the way companies view and manage environmental performance by demonstrating the economic and productivity gains associated with 'lean and clean' manufacturing;
- "Foster innovation by allowing participants to identify the actions that make the most sense for their organizations; and
- "Develop productive and flexible partnerships within government and between government and industry."

Climate Wise consisted of three interrelated components. First, the pledge component asked firms to commit to taking cost-effective, voluntary actions to reduce GHG emissions. Second, tailored assistance efforts were designed to facilitate companies' emissions-reducing efforts via a clearinghouse, workshops, and seminars. Finally, communication activities provided public recognition for actual progress in reducing emissions.

To join Climate Wise, a firm had to develop a baseline estimate of its direct emissions of CO_2 (and other GHGs) for the year it joined the program or any prior year of its choice since 1990. A more detailed (and burdensome) emissions inventory was not required. In addition to establishing a baseline, a firm was required to identify specific actions it proposed to undertake to reduce its emissions and, for each action, to indicate whether it was a "new," "expanded," or "accelerated" initiative. To encourage consideration of substantial reductions, EPA provided a checklist of major actions to improve equipment and processes, including those involving boiler efficiency, air compressor systems, steam traps, and piping and heat-generating equipment. EPA also suggested fuel switching, best management practices, and the further integration of energy efficiency in new product design and manufacturing. Firms were strongly encouraged, albeit not required, to select at least some of their proposed actions from this list. The only formal requirement was for a firm to establish an emissions goal and to provide a progress report directly to EPA. Participants were also encouraged, but not required, to report their progress to the DOE through the 1605(b) registry program.

EPA provided several types of technical assistance to participating firms, including a guide to industrial energy efficiency, various background publications on energy efficiency and related issues and, most importantly, free phone consultation with in-house energy experts and consultants retained by the agency. Information about financial assistance to support emissions-reducing actions was also made available to participants, including via Small Business Administration guaranteed loans and low-interest buy-downs from state providers, utility programs, and others. Further, EPA held an annual event open to the public to recognize the performance of outstanding Climate Wise participants.

Although the Climate Wise program focused on energy efficiency and the reduction of CO_2 emissions, a number of firms proposed to reduce emissions of non-CO_2 GHGs as well. Reportedly, the most substantial reductions of the non-CO_2 gases were in the chemical industry, where relatively large amounts of nitrous oxide (N_2O) emissions are released in the manufacture of adipic acid. Significant amounts of methane were also included in the action plans of several firms, especially in the beer industry.

The 1605(b) Program

In contrast to Climate Wise, which was initiated entirely by EPA, section 1605(b) of the EPAct directed DOE, via the Energy Information Administration (EIA), to develop a program to document voluntary actions that reduce GHG emissions or remove GHGs from the atmosphere. The EPAct mandated that EIA issue guidelines for reporting, establish suitable procedures for measurement, ensure confidentiality of trade secrets and of commercial and financial information, and establish a publicly available database. It also mandated consultation with EPA. The first reports covered the year 1994.

Although it involves fewer programmatic activities than Climate Wise, the 1605(b) program does provide recognition for entities that reduce GHG emissions or sequester carbon voluntarily, and it identifies innovative and effective ways to reduce emissions. Most of the reporting entities are affiliated with one or more other voluntary programs sponsored by EPA or other government agencies.

Consistent with its legislative mandate, the 1605(b) program is extremely flexible. Voluntary participants can choose to report reductions at the firm or project level and can then define the reporting boundary relevant for either the firm or the project. Participants can select a "basic" reference case as any single year between 1987 and 1990, or an average of those years. Alternatively, reductions can be reported against a "modified" or hypothetical reference case, reflecting the emissions (or sequestration) that would have occurred in the absence of the project. Further, participants can measure their reductions in either absolute terms or on the basis of emissions intensity.

Since its inception in 1994, activities reported under 1605(b) have increased dramatically: the number of reporting entities has doubled from about 100 to more than 200 per year; the number of projects has more than tripled from about 600 to more than 2,000 per year; and reported reductions in direct emissions have more than quadrupled from 63 million metric tons in 1994 to 277 million metric tons in 2004, reflecting a 3.9 percent reduction from reported emissions in the base year. Overall, the electric power sector reported more entities, projects, and tons of emissions reduced than any other sector in the database. Effective June 1, 2006, the program was revised and the reporting flexibility was reduced somewhat. However, the analysis presented in this study is based on the data firms reported to the program during 1994–2000.

The 1605(b) Program differs from Climate Wise and most other voluntary programs initiated during the early 1990s in its diversity of project types, participation, and approaches. The 1605(b) database provides examples of the types of concrete actions that organizations

report to reduce GHG emissions. The EIA notes some of the most important benefits of the 1605(b) Program as follows (U.S. EIA 2002, 1–2):

- "The program has served to teach staff at many of the largest corporations in the United States how to estimate greenhouse gas emissions and has educated them on a range of possible measures to limit emissions.
- "The program has helped to provide concrete evidence for the evaluation of activities reported to the many government voluntary programs launched since 1993.
- "Reporters have been able to learn about innovative emission reduction activities from the experiences of their peers.
- "The program has created a "test" database of approaches to emission reductions that can be used to evaluate future policy instruments aimed at limiting emissions.
- "The program has helped to illuminate many of the poorly appreciated emissions accounting issues that must be addressed in designing any future approaches to emission limitations."

Data

For both Climate Wise and the 1605(b) Program, we combine participation data from the relevant government agencies with outcome data and control observations drawn from the U.S. Census Bureau's Longitudinal Research Database (LRD). We focus exclusively on the manufacturing sector.

The Climate Wise Program

For EPA's Climate Wise Program data, we obtained a list of voluntary program participants from EPA describing firms that enrolled in Climate Wise in each of its operational years, from 1994 to 2000. This list includes name, zip code, and enrollment date data for two different types of participants, those that joined at the corporate level and those that joined as individual plant participants. Complete data were available for a total of 671 participants. The left panel of Table 1 displays the distribution of both types of participants over time. As shown, the number of corporate participants reached a peak in 1996 and gradually dropped to zero in 2000. However, the number of plant participants continued to increase until 2000.

This information on program participation then was linked to detailed data from the LRD using name and, for plant participants, zip code information. We succeeded in linking a total of 377 out of 671 participants, including 228 corporate participants and 149 plant participants. To some extent, the failure to link some participants to the census data reflects the fact that the LRD includes only manufacturing establishments, whereas the Climate Wise program, despite its programmatic focus on manufacturing, also includes nonmanufacturing participants (e.g., municipalities and commercial buildings).

These 377 linked participants from the original Climate Wise list translate into 2,311 facilities because corporate participants could have multiple associated facilities. The results of the matching are displayed in the left panel of Table 2.

Summary statistics for the linked sample, as well as the entire LRD, including participating firms and nonparticipants, are given in the left panel of Table 3. Here we see the principal differences among participants and the broader universe of plants in the census data: the participants are considerably larger. Our participant sample is also a very small fraction of the plants in the LRD—roughly 1 percent. This suggests that the full census sample is unlikely to be an appropriate control group as a whole, and that a large number of plants are available from which to choose a more appropriate subgroup of controls.

Linking Climate Wise and census data has important consequences for our ability to evaluate the effect of program participation over longer horizons. Because we are attempting to study behavior two or three years after joining, we are forced to drop plants that enrolled in 2000 and 1999, respectively, because our census data are available only through 2001. As noted, corporate participants provide the overwhelming majority of participant observations because they match to multiple facilities. Given the steep drop-off in new corporate participants after 1998, we do not sacrifice many observations by considering two to three years from the enrollment date. However, trying to discern effects four years after enrollment, with only those participants that joined between 1994 and 1997, we would have noticeably fewer observations and noisier estimates. Thus, we do not attempt to examine effects more than three years after enrollment.

The 1605(b) Program

For DOE’s 1605(b) Program, we obtained a list of reporting entities, sectors, years reported, and form type used for the years between 1994 and 2001. The reporting entities are distributed among six sector categories: agriculture, alternative energy, electric power, industry,

N/A, and other. The electric power sector accounts for more than one-third of total reporting entities (130 out of 383). In this study, we focus on manufacturing participants, which account for only 18 percent of the all reporting entities. In the left panel of Table 4, we provide sector distribution information for all reporting entities.

Unlike EPA's Climate Wise program, DOE 1605(b) data do not include participants' enrollment dates. Thus, we use the first reporting year as the enrollment year and assume that the participants continued in the program after that, even though individual entities may not have continuous reporting years. The right panel of Table 1 displays the enrollment year information based on either firm or plant participation.

We also obtained a separate entity file from EIA that includes entity identification number, name, street, city, state, contact, Internet address, and sector information. Using this information, we were able to match participation data with census data. In the right panel of Table 4, we show the sector distribution for the 1605(b) Program and LRD matching results. For the industrial sector, the matching rate is about 77 percent, and for the sector classified as N/A, the matching rate is 36 percent. For other sectors, the matching rate is only 13 percent because most of the others are electricity- and energy-relevant entities that do not fall into the manufacturing category.

After excluding missing enrollment year and other criteria, we were able to link 83 out of 383 participants, including 67 corporate participants and 17 plant participants. Compared to Climate Wise, we have a much lower matching rate for the 1605(b) Program because it includes both manufacturing and nonmanufacturing sectors. In fact, more than 50 percent of the participants are in the electric power and alternative energy sectors, which are not included in the census data. These 83 linked participants from the original 1605(b) Program list correspond to 1,791 LRD facilities because corporate participants could have multiple facilities. The right panel of Table 2 summarizes the matching of the 1605(b) data to the LRD. The right panel of Table 3 provides summary statistics for the linked sample, as well as the entire LRD for the 1605(b) Program.

Models and Econometric Method

With the linked census data described in the preceding section, we have access to variables indicating energy expenditures (separately for fuels and electricity), size (measured by the total value of shipments), location, and industry for a large sample of manufacturing plants over a range of years from 1992 until 2000. We also have linked information on which plants

participated in each of the two programs and the year in which they first participated. We consider two alternative approaches for evaluating outcomes, attempting to control for selection (i.e., the participation decision) based on observables in one and unobservables in the other. In each case, we can imagine two outcomes, Y_i, for every observed plant, i: the value associated with participation, $Y_i(1)$, and the value associated with nonparticipation, $Y_i(0)$. Here, $Y_i(D_i)$ is the outcome associated with either treatment, $D_i = 1$, or nontreatment, $D_i = 0$, and is the cost of either fuels or electricity measured in natural logarithms. The ideal study would measure the treatment effect,

$$Y_i(1) - Y_i(0)$$

for each plant; that is, the percentage change in energy expenditures when a plant joins the program. The obvious problem is that for every plant, we observe either $Y_i(1)$ or $Y_i(0)$, but never both. The problem, viewed this way, is one of missing data and the fact that selection determines which data are observed and which are missing (that is, who participates).

The simplest solution, and the one appropriate for randomized experiments, is to assume that the missing observations are *missing at random* (Rubin 1974). Under this assumption, the selection mechanism determining which outcomes are observed is *ignorable*. Formally $D_i \perp Y_i(1), Y_i(0)$, we can measure the *average treatment effect* as

$$E\left[Y_i(1) - Y_i(0)\right] = \frac{\sum_{D_i=1} Y_i(1)}{\sum_{D_i=1} 1} - \frac{\sum_{D_i=0} Y_i(0)}{\sum_{D_i=0} 1}$$

that is, as the difference in average outcomes between participants and nonparticipants. Or we can estimate the treatment effect from a simple regression model

$$Y_i(D_i) = \beta_0 + \beta_1 D_i + u_i$$

where we assume that u_i is uncorrelated with D_i and the treatment effect is the estimated value of β_1. Of course, in reality, observations are unlikely to be missing at random; hence, we must use the following two approaches.

In our first approach, we follow Heckman and Hotz (1985) and consider a model in which program participation depends on both observed and unobservable variables that may be correlated with the outcome,

$$D_i \perp Y_i(1), Y_i(0) \mid X_i, u_i$$

where u_i is an unobserved variable. We build a structural model in which, even though selection, D_i, is dependent on an unobserved variable, we can still consistently estimate the treatment effect. Specifically, we assume an outcome model of the form

$$Y_i = \beta_0 + \beta_1 \cdot X_i + \beta_2 \cdot D_i + u_i \tag{1}$$

where we have allowed for covariates. We still must deal with the problem that Y_i and D_i are not independent, and are even conditioned on X_i. In particular, we assume that u_i and D_i are correlated, thus violating a key assumption for unbiased estimation in an OLS model (i.e., that the error must be *uncorrelated* with all of the right-hand side variables). The solution is to specify a model for participation, D_i, and thereby parameterize the correlation with u_i.

In particular, we specify a selection model

$$D_i^* = \delta \cdot Z_i + v_i$$

where $D_i = 1$ if $D_i^* > 0$ and $D_i = 0$ otherwise, and Z_i is a set of covariates with at least one additional covariate not included in X_i (referred to as "excluded variables"). This condition is necessary for identification, and intuitively reflects the presence of a variable that influences the decision to participate in the voluntary program; however, it does not directly influence the energy expenditure outcome. For example, if we find that individual programs were more aggressively marketed in some years than in others, we could create a variable indicating whether firms joined in particular years. This variable would be precisely the kind that would help identify participation without directly influencing the energy expendigure outcome.

If we assume that (u_i, v_i) are jointly normal, it can be shown that

$$E\left[v_i \middle| D_i, Z_i\right] = \lambda\left(D_i, Z_i\right) = \begin{cases} \dfrac{\phi\left(-\delta \cdot Z_i\right)}{1-\Phi\left(-\delta \cdot Z_i\right)} & D_i = 1 \\ \dfrac{-\phi\left(-\delta \cdot Z_i\right)}{\Phi\left(-\delta \cdot Z_i\right)} & D_i = 0 \end{cases}$$

If we then specify the outcome equation as

$$Y_i = \beta_0 + \beta_1 \cdot X_i + \beta_2 \cdot D_i + \alpha \cdot \lambda\left(D_i, Z_i\right) + \varepsilon_i \tag{2}$$

where $u_i = \alpha \cdot \lambda\left(D_i, Z_i\right) + \varepsilon_i$, the error, ε_i, is no longer correlated with D_i because $\alpha \cdot \lambda\left(D_i, Z_i\right)$ reflects the expectation of u_i given D_i. Note the rationale for the identifying assumption that at least one variable in Z_i must be excluded from X_i: except for the nonlinearity in the function λ, the right-hand side variables would be colinear if X_i included all of the variables in Z_i.

The dependent variable in this model is the change in the natural log of energy expenditures (fuel and electricity) over different time horizons after a given year when plants join the voluntary program. When we estimate this model, we include linear and quadratic values of our key variables as controls, X_i: logged and lagged value of shipments, electricity costs, and fuel costs. We also include the change in the logged value of shipments over the given time horizon as a control variable. Although this is arguably endogenous, we believe that controlling for growth is critical: we observed that faster-growing plants are more likely to join voluntary programs. It seems unlikely that this growth was *caused* by joining; therefore, we need to control for it.

We also include dummy variables for census region and for the two-digit industry classification. Our Z_i variables include two variables that we believe are likely to influence participation but not the outcome—distance to the nearest EPA regional office and local membership rates in a national environmental organization.[3]

Our second approach, propensity score matching, is based on work by Rosenbaum and Rubin (1983); it was more recently used by List et al. (2003) and Dehejia and Wahba (2002). This approach makes an alternative assumption that the participation decision is ignorable conditional only on observed covariates, or

$$D_i \perp Y_i(1), Y_i(0) \mid X_i$$

This condition could be met via a model such as (1), except that it requires a correct specification of the X_i dependence—otherwise the estimated effect of the program remains mingled with covariates. Instead, Rosenbaum and Rubin (1993), and others, match participants to appropriate nonparticipants and consider the pairwise differences. Although the Heckman–Hotz approach attempts to control for selection on additional, unobserved effects correlated with outcome, it requires both a correct specification and identification of one or more excluded variables. This approach does not control for such effects, but it relaxes the specification assumption and does not require excluded variables.

The general problem of creating a set of matched, nonparticipating observations is quite challenging because we would want to match the many observable variables—in our case, those describing location, industry, size, energy intensity, and growth. However, Rosenbaum and

[3] Many thanks to the National Wildlife Federation for supplying these data.

Rubin (1983) show that we need to match only the expected likelihood of participation. That is, we simplify the difficult problem of matching all of these different variables to a much simpler one of matching a summary variable describing the propensity to join the program. This greatly simplifies the creation of matched nonparticipant observations.

Our model of propensity score—the likelihood of joining the voluntary programs—is similar to our model of outcome in (2). It depends on linear and quadratic terms involving value of shipments, cost of fuels, and cost of electricity (all in logarithms), as well as dummy variables for census region and two-digit industry classification. As before, we also include a term for growth in the value of shipments over a given horizon, *h*, as this turns out to be an important determinant of participation. Because it seems unlikely that participation would cause growth, we take this as a proxy for expected growth over the given horizon. We use samples matched with different horizons to estimate program effects over similar horizons.

Because each of the voluntary programs lasted a number of years, we decided to address the decision to join in a duration model framework. That is, in each period, conditional on not having joined, there is a given probability of joining based on the noted covariates and time. This allows us to combine data across years in estimating our model.[4] We therefore estimate a Cox proportional hazard model of the form:

$$\begin{array}{c}\text{probability of joining in year } t \\ \text{(assuming plant } i \text{ has not yet joined)}\end{array} = h(t)\exp\begin{pmatrix} \beta_{size}\ln TVS_{i,t-1} + \beta_{elec}\ln EE_{i,t-1} + \beta_{fuels}\ln CF_{i,t-1} \\ +\left[\text{all quadratic combinations of size, elec, fuels}\right] \\ +\beta_{growth}\left(\ln TVS_{i,t+h} - \ln TVS_{i,t-1}\right) \\ +\sum_{\text{industries } j}\beta_j 1\left(M_i = j\right) + \sum_{\text{region } k}\beta_k 1\left(G_i = k\right) \end{pmatrix}$$

Once estimated, we predict hazard rates for each participating firm in the year it enrolls and match it to the nearest-valued nonparticipant (i.e., nonparticipant with the closest matching propensity score) in that year. We then examine the difference across each pair in the changes in fuel and electricity expenditures; this difference-in-differences forms the estimate of program effectiveness.

[4] Note that, although the participants have an obvious enrollment year associated with them, the nonparticipants do not. In other words, plants may participate in various years, but this is not the case for nonparticipation. Outside of a duration model, it may not be possible to combine the data. In the Heckman–Hotz approach, we estimate effects for different cohorts of participants separately for this reason.

Results

Tables 5 through 8 present results for the Heckman–Hotz approach for the 1605(b) and Climate Wise programs, showing both the cost of fuels and the cost of electricity as outcome variables. We report only the results for a two-year horizon, where the dependent variable measures the change between the year prior to enrollment by a group of participants (a cohort) and two years later; the results are broadly similar for one- and three-year horizons. We have reported the results with and without the selection correction term, $\lambda\left(D_i, Z_i\right)$ from (2). The results without the correction term (first column) reflect the simplest model, with the outcome depending only on covariates (value of shipments, cost of fuels, cost of electricity, growth in the value of shipments, and region and industry dummy variables) and the dummy variable indicating whether a firm joins the program in a given year.

Among the results for this simple model in the first column of each table, without any correction for possible selection bias, we generally estimate small, statistically insignificant effects, with changes in energy expenditures of less than 10 percent. The three exceptions are a statistically significant 9 percent decline in electricity costs among 1605(b) participants in the 1994 cohort, a 6 percent increase in electricity costs among Climate Wise participants in the same cohort, and a *55 percent increase* in fuel costs among Climate Wise participants in the 1999 cohort. The first two effects are not inconsistent with our observations below, that electricity expenditures might increase among Climate Wise participants if efforts to reduce direct emissions lead to more electricity use and higher indirect emissions. In addition, a positive electricity effect could reflect a combination of specification error and the fact that larger or faster-growing firms are more likely to participate in voluntary programs. The latter, a 55 percent increase in fuel expenditures among one cohort of Climate Wise participants, probably reflects an outlier in the rather small sample (96 participants) for that year and/or specification error.

The preceding results ignore the potential for selection bias, which is the main focus of this exercise. The second column of Tables 5 through 8 presents results where we first estimate the probability of selection, and then include the selection correction term, $\lambda\left(D_i, Z_i\right)$ from (2), in the original outcome regression. For simplicity, we have not reported the results of the first-stage regression; however, these results showed that the excluded variables are *almost never*

statistically significant.[5] Empirically, it is difficult to see a difference in the distribution of either variable among participants and nonparticipants, suggesting that this approach may be problematic given available data.

When we look at the results across the four program/outcome variable combinations, the results are indeed problematic. Five of 24 estimates are statistically significant, ranging from a –1.42 (0.71) estimated effect on electricity costs in the 1999 participant cohort of the Climate Wise program, to a +0.60 (0.09) estimated effect on electricity costs in the 1994 cohort of the same program (this would suggest an effect ranging from –76 percent to +82 percent across different years). We also see much larger standard errors on the estimates compared to the simple estimates in column one. This probably reflects multicollinearity between the correction term and the right hand-side variables, where it may be mostly the nonlinearity of the $\lambda\left(D_i, Z_i\right)$ function identifying the parameters rather than the excluded variables. Because such a divergent range of effects driven by the participation year seems implausible—and because of the problem with the excluded variables—we tend to distrust this approach.

Instead, we now turn to the results from the propensity score matching approach shown in Tables 9 through 12. As noted in the methods section, we estimate a duration model for whether or not facilities enroll, using a variety of specifications. These specifications differ based on whether dummy variables are included for industry and region and whether quadratic terms are included, as indicated in the top three rows of each table. For each specification, we consider effects over one, two, and three years; we pool across all cohorts of matched pairs, and report both the mean and the median differences in energy expenditures across pairs.

As with the simple model (column one of Tables 5 through 8), all of the estimates suggest effects of less than 10 percent. We focus our discussion on the median estimates in the bottom half of each table because they are more robust to outlying observations of paired differences. Only 4 of these 72 median estimates are larger than 5 percent in magnitude, suggesting that any effect is probably even smaller than 10 percent. In general, the estimated changes in electricity expenditures are more likely to be statistically significant (6 of 36 estimates) than are the estimated changes in fuel expenditures (1 of 36). Interestingly, the 1605(b) Program seems to have a negative effect on electricity expenditures of perhaps several

[5] Distance to EPA regional office is significant in the 1997 cohort. Note that we experimented with various specifications of the excluded variable with no appreciable difference. The reported specification is based on logged distance to EPA regional office and county environmental group membership as a share of total county population.

percentage points (Table 10, where 17 of 18 median estimates are negative), whereas Climate Wise appears, if anything, to have a slight positive effect on electricity expenditures (Table 12, where 14 of 18 median estimates are positive). The positive effect of Climate Wise is not present in our most general matching model (Table 12, where median estimates in column one are not significant); however, the negative effect of 1605(b) is present in this model (Table 10, where median estimates in column one are significant). Further, we found no evidence of persistence in the Climate Wise results: the effects at the three-year horizon are all lower than at two years (see bottom two rows of Table 12). Meanwhile, estimates of 1605(b) effects in four out of six models are greatest for the longest horizon (bottom row of Table 10).

In summary, we make several key observations based on the simple results of the Heckman–Hotz approach (column one in Tables 5 through 8) and the propensity score matching approach (Tables 9 through 12). As noted earlier, we tend to distrust the Heckman–Hotz results.

1. The effects of Climate Wise and 1605(b) on fuel and electricity expenditures are no more than 10 percent and probably less than 5 percent.
2. We found virtually no evidence of a statistically significant effect of either Climate Wise or 1605(b) on fuel costs.
3. Some statistically significant evidence suggests that participation in Climate Wise led to a slight (3–5 percent) increase in electricity costs that vanishes after two years.
4. Some statistically significant evidence suggests that participation in 1605(b) led to a slight (4–8 percent) decrease in electricity costs that persists for at least three years.

The transient, slight increase in electricity costs under Climate Wise is certainly unexpected. Two explanations come to mind. First, participating plants may have pursued direct emissions reductions that required increased electricity use. Ignoring the indirect emissions associated with electricity use, this technically reduces emissions as defined by the program goals, but with the unintended consequence of higher indirect emissions from electricity use. Lower direct emissions might not show up in the cost of fuel measure because fuel switching—for example, a shift to biomass or a shift from coal to gas—might reduce emissions without changing expenditures. Or plants may have pursued nonenergy-related emissions reductions—such as the reduction of N_2O emissions at chemical plants, methane emissions at refineries, or CO_2 process emissions at cement or other industrial sources—that is not reflected in a lower cost of fuels.

Alternatively, a positive effect on electricity expenditures may reflect a failure to adequately control for growth. Although we matched participants with nonparticipants based, in part, on growth in the value of shipments, the tendency for faster-growing firms to enroll remains troubling. For example, we have no way of knowing about the underlying prices and quantity changes: participants might experience changes in quantities, whereas nonparticipants matched from the LRD might experience changes in prices. We cannot tease out controls that have that same pattern because details on prices and quantities are not available. If the estimated electricity expenditure growth effect really reflects an underlying and uncorrected difference in growth between participants and controls then, presumably, fixing it would raise the growth rate of the control group and make the estimated program effect on electricity and fuel costs more negative.

Conclusions

Thus far, the rigorous assessment of the environmental performance of voluntary programs, especially climate-related programs, has been quite limited. The key challenge is to measure performance relative to a realistic baseline. In the present study, we examine both an EPA- and a DOE-sponsored program, relying on confidential plant-level data for the manufacturing sector collected by the U.S. Census Bureau through 2001 to develop such a baseline based on a comparable set of nonparticipant controls. We consider two alternative approaches to evaluating outcomes, attempting to control for selection in joining the programs based on both observable and unobservable characteristics. In the Heckman–Hotz approach, we consider a structural model in which program participation depends on both observed and unobservable variables that may be correlated with the outcome. In the propensity score matching approach, we match participants to appropriate nonparticipants based on observable characteristics only, and consider pairwise differences. The results are sobering.

In contrast to the claims of relatively large emissions reductions reported by the sponsoring agencies, our analysis suggests that more modest reductions are attributable to the programs studied. Overall, we find that that the effects of Climate Wise and 1605(b) on fuel and electricity expenditures were no more than 10 percent and probably less than 5 percent. We found no evidence of reductions in direct emissions from fossil fuels (as measured by energy expenditures) attributable to the voluntary programs; however, we found some statistically significant impacts on the use of electricity. In particular, some statistically significant evidence suggests that participation in 1605(b) leads to a slight decrease in electricity expenditures, on the order of 4–8 percent. This decrease persists for at least three years. Our results suggest that Climate Wise may be associated with a slight *increase* in electricity expenditures, although that

effect vanishes after two years. Given the limitations of the analysis, we tend to discount these findings and conclude, instead, that participation in Climate Wise probably has, at most, a negligible effect on emissions.

The findings of modest, albeit statistically significant, reductions in electricity expenditures for 1605(b) reporters may have implications for other government-sponsored voluntary programs as well. Recall that most of the entities reporting under 1605(b) are also affiliated with one or more other government-sponsored programs. Thus, the observed emissions reductions for 1605(b) reporters may reflect the influence not only of the 1605(b) Program itself, but also that of other programs. Although our separate assessment of Climate Wise suggests that participation in that program is probably not associated with significant emissions reductions, larger programs, such as EPA's Energy Star Program, may be more effective. Unfortunately, the EIA reporting form does not require disclosure of the name of any of the other individual programs in which a firm participates.

Methodologically, our study highlights the inevitable complexity of assessing voluntary programs. Our study reinforces the work of others in emphasizing the importance of distinguishing between the participation decision and the environmental outcomes achieved. Our work also points to the value of working with micro-level data, and the particular need to take special care in matching otherwise disparate samples to obtain a credible control group. This process is all the more difficult in our case, where the samples were not coded via a uniform system. In terms of estimation, we have applied two distinct methods to evaluating outcomes. One, based on the work of Heckman and Hotz (1985), assumes that program participation depends on both observed and unobservable variables that may be correlated with the outcome. The other, propensity score matching, based on the work of Rosenbaum and Rubin (1983), matches participants to appropriate nonparticipants and considers pairwise differences. Because the Heckman–Hotz approach requires both a correct specification and identification of one or more excluded variables, it is more demanding than the propensity score matching approach of Rosenbaum and Rubin (1983), which relaxes the specification assumption and does not require excluded variables (but does not allow for correlated, unobserved errors in the selection and outcome model). Because of the difficulty of identifying excluded variables in the Heckman–Hotz approach, our results from the propensity score matching approach seem more plausible; we believe that such an approach may have wider applicability in the future evaluation of voluntary programs.

Our findings of, at most, a small effect should not be very surprising. Energy-related GHG emissions are quite different from many other types of emissions, such as unpriced

industrial byproducts (e.g., toxics) with no near-term localized effects and whose existence was widely ignored until the 1980s and 1990s. With no practical opportunity for end-of-pipe abatement, reductions in energy-related GHG emissions often amount to reductions in energy use itself—something that has been picked over for some time. The underlying positive price on energy *always* creates an incentive to reduce energy use. The existence of such underlying incentives, in turn, implies a far greater challenge for government in designing effective voluntary programs for industry.

References

Arora, Seema, and Timothy Cason. 1995. An Experiment in Voluntary Environmental Regulation: Participation in EPA's 33/50 Program. *Journal of Environmental Economics and Management* 28: 271–286.

———. 1996. Why do Firms Volunteer to Exceed Environmental Regulations? Understanding Participation in EPA's 33/50 Program. *Land Economics* 72: 413–432.

Celdren, A., H. Clark, J. Hecht, E. Kanamaru, P. Orantes, and M. Santaello Garguno. 1996. The Participation Decision in a Voluntary Pollution Prevention Program: The USEPA 33/50 Program as a Case Study. In *Developing the Next Generation of the USEPA's 33/50 Program: A pollution Prevention Research Project*, edited by H. Clark. Durham, NC: Duke University Nicholas School of the Environment.

Dasgupta, Susmita, David Wheeler, and Hemamala Hettige. 1997. *What Improves Environmental Performance? Evidence from Mexican Industry.* World Bank Policy Working Paper No. 1877 (December). Washington, DC: World Bank.

DeCanio, S.J. 1998. The Efficiency Paradox: Bureaucratic and Organizational Barriers to Profitable Energy-Saving Investments. *Energy Policy* 26: 441–454.

Decanio, S.J., and William E. Watkins. 1998. Investment in Energy Efficiency: Do the Characteristics of Firms Matter? *Review of Economics and Statistics* 80: 95–107.

Dehejia, Rajeev H., and Sadek Wahba. 2002. Propensity Score-Matching Methods for Nonexperimental Causal Studies. *The Review of Economics and Statistics* 84(1):151.

Dowd, Jeff, Kenneth Friedman, and Gale A. Boyd. 2001. *How Well Do Voluntary Agreements and Programs Perform at Improving Industrial Energy Efficiency?* Washington, DC: Office of Policy, U.S. Department of Energy.

Gamper-Rabindran, Shanti. 2006. Did the EPA's Voluntary Industrial Toxics Program Reduce Plants' Emissions? A GSIS Analysis of Distrributional Impacts and a By-Media Analysis of Substitution. *Journal of Environmental Economics and Management* 52(1), 391–410.

Heckman, J.J. and V.J. Hotz. 1985. An Investigation of the Labour Market Earnings of Panamanian Males: Evaluating the Sources of Inequality. *Journal of Human Resources* 21(4): 507-542.

Howarth, R., B. Haddad, and B. Paton. 2000. The Economics of Energy Efficiency: Insights from Voluntary Programs. *Energy Policy* 28: 477–486.

Khanna, Madhu. 2001. Non-Mandatory Approaches to Environmental Protection. *Journal of Economic Surveys* 15(3): 291–324.

Khanna, Madhu, and Lisa A. Damon. 1999. EPA's Voluntary 33/50 Program: Impact on Toxic Releases and Economic Performance of Firms. *Journal of Environmental Economics and Management* 37(1): 1–25.

King, A. A., and M. J. Lenox. 2000. Does Membership Have its Privileges? Analyzing Who Benefits from Industry Self-Regulation. Working Paper. Leonard N. Stern School of Business, New York University, New York, November.

List, John A., D.L. Millimet, P.G. Fredriksson, W.W. McHone. 2003. Effects of Environmental Regulations on Manufacturing Plant Births: Evidence from a Propensity Score Matching Estimator. *Review of Economics and Statistics* 85(4):944–952.

Lyon, T., and J. Maxwell. 2002. Voluntary Approaches to Environmental Regulation: A Survey. In *Economic Institutions and Environmental Policy*, edited by M. Franzini, and A. Nicita. Hampshire, UK: Ashgate Publishing, 142–174.

Lyon, Thomas P. and Kim, Eun-Hee, "Greenhouse Gas Reductions or Greenwash?: The Doe's 1605b Program" (November 2006). Available at SSRN: http://ssrn.com/abstract=981730

Rosenbaum, P., and D. Rubin. 1983. The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika* 70:41–55.

Rubin, D. 1974. Estimating Causal Effects of Treatments. *Journal of Educational Psychology* 66: 688–701.

Sam, A.G., and R. Innes. 2005. Voluntary Pollution Reductions and the Enforcement of Environmental Law: An Empirical Study of the 33/50 Program. University of Arizona, working paper.

U.S. EIA (Energy Information Administration). 2002. *Voluntary Reporting of Greenhouse Gases 2000*. February. Washington, DC: U.S. EIA, Office of Integrated Analysis and Forecasting. www.eia.doe.gov/oiaf/1605/vrrpt/index.html (accessed May 15, 2008).

U.S. EPA (Environmental Protection Agency). 1998. *A Catalogue of the Agency's Partnership Programs*. EPA-100-B-97-003. Washington, DC: U.S. EPA, Office of the Administrator.

———. 2005. *Everyday Choices: Opportunities for Environmental Stewardship*. Technical Report by the EPA Environmental Stewardship Staff Committee for EPA Innovation Council. Washington, DC: U.S. EPA.

Welch, E., A. Mazur, and S. Bretschneider. 2000. Voluntary Behavior by Electric Utilities: Levels of Adoption and Contribution of the Climate Challenge Program to the Reduction of Carbon Dioxide. *Journal of Policy Analysis and Management* 19(3): 407–425.

White House. 2002. Global Climate Change Policy Book. http://www.whitehouse.gov/news/releases/2002/02/climatechange.html (accessed August 1, 2006).

White House. 2005. Global Climate Change Policy Book. http://www.whitehouse.gov/news/releases/2005/05/20050518-4.html (accessed August 1, 2006).

Tables

Table 1. Enrollment Data for Climate Wise and DOE 1605(b) Participants

	Climate Wise			1605(b) Program		
Enrollment year	Plant	Corporate	Subtotal	Plant	Corporate	Subtotal
1994	0	8	8	0	105	105
1995	7	30	37	0	37	37
1996	38	141	179	3	23	26
1997	37	101	138	2	15	17
1998	36	70	106	8	53	61
1999	72	17	89	2	33	35
2000	144	0	144	6	53	59
Total	304	367	671	21	319	340

Table 2. Matching of Climate Wise (CW) and 1605(b) to LRD

	CW List	LRD Plants	LRD plant–year observations (1992–2001)	1605(b)	LRD Plants	LRD plant–year observations (1992–2001)
Corporate participants with multiple plants	135	2,053	11,503	54	1,762	8,724
Corporate participants with a single plant	93	95	316	13	13	63
Plant-level participants	149	163	946	16	16	122
Total	377	2,311	12,765	83	1,791	8,909

Table 3. Sample Statistics for LRD and CW and 1605(b) Program Participants

Variable	Summary statistics	CW		1605(b)	
		Full LRD sample (1992–2001)	Program participants	Full LRD sample (1992–2001)	Program participants
ln(*TVS*) (total value of shipments)	Mean	7.61	10.87	7.80	10.99
	Standard deviation	2.30	1.81	2.34	2.17
	Plant–year observations	1,157,606	12,605	871,316	8,758
ln(*CF*) (cost of fuels)	Mean	2.54	5.31	2.69	5.36
	Standard deviation	2.12	2.23	2.18	2.32
	Plant–year observations	839,934	11,280	638,520	7,582
ln(*PE*) (purchased electricity)	Mean	3.17	6.31	3.34	6.22
	Standard deviation	2.21	1.83	2.25	2.17
	Plant–year observations	1,019,042	12,377	784,502	8,564
	Number of Plants	515,189	2,311	385,531	1,791

Table 4. The Sector Distribution for 1605(b) Reporting Entities and for Matched 1605(b) and LRD Data

Sector	Reporting counts	Matched counts
Agriculture	12	D*
Alternative energy	63	D*
Electric power	130	D*
Industry	69	53
N/A	94	34
Other	15	13
Total	383	100

Note: Items marked with D* are included in "Other."

Table 5. Effect of the 1605(b) Program on Natural Log of Fuel Expenditures after Two Years: Heckman–Hotz Approach

Cohort	W/o correction	With correction	Sample size	Number of Participants
1994	–0.05 (0.05)	0.00 (0.24)	14686	343
1995	–0.06 (0.08)	0.30 (0.36)	24369	193
1996	–0.06 (0.20)	–0.55 (0.49)	22480	28
1997	–0.14 (0.08)	–0.79 (0.44)	13146	192
1998	–0.03 (0.09)	–0.51 (0.37)	21107	164
1999	0.09 (0.11)	–0.05 (0.48)	17667	162

Table 6. Effect of the 1605(b) Program on Natural Log of Cost of Electricity after Two Years: Heckman–Hotz Approach

Cohort	W/o correction	With correction	Sample	Participants
1994	–0.09 (0.03)*	–0.86 (0.14)*	15319	343
1995	0.06 (0.06)	–0.71 (0.23)*	26123	193
1996	–0.17 (0.14)	–0.52 (0.35)	24089	28
1997	0.04 (0.05)	0.29 (0.24)	13754	192
1998	0.04 (0.06)	–0.25 (0.24)	22536	164
1999	0.05 (0.07)	0.29 (0.32)	18768	162

Table 7. Effect of EPA Climate Wise Program on Natural Log of Cost of Fuels after Two Years: Heckman–Hotz Approach

Cohort	W/o correction	With correction	Sample	Participants
1994	0.06 (0.03)	0.20 (0.14)	18788	809
1995	0.06 (0.06)	0.08 (0.20)	32768	335
1996	0.04 (0.05)	0.26 (0.33)	29111	656
1997	–0.04 (0.05)	–0.33 (0.29)	16706	835
1998	–0.04 (0.04)	–0.49 (0.29)	28658	1063
1999	0.55 (0.14)*	0.41 (0.96)	18702	96

Table 8. Effect of EPA Climate Wise Program on Logged Cost of Electricity after Two Years: Heckman–Hotz Approach

Cohort	W/o correction	With correction	Sample	Participants
1994	0.06 (0.02)*	0.60 (0.09)*	19627	809
1995	0.04 (0.04)	–0.16 (0.14)	34880	335
1996	0.02 (0.03)	0.36 (0.21)	31253	656
1997	–0.02 (0.03)	–0.29 (0.18)	17534	835
1998	0.01 (0.02)	–0.75 (0.16)*	30693	1063
1999	0.05 (0.12)	–1.42 (0.71)*	33971	96

Table 9. Effect of the 1605(b) Program on the Natural Log of Fuel Expenditures over Different Horizons: Propensity Score Matching Approach

Matching model (all models include logged value of shipments, cost of fuels, cost of electricity, and growth in shipments)							Matched sample
Industry	x	x				x	
Region	x	x			x		
Quadratic	x			x	x	x	
Mean							
1-year effect	0.02	0.03	0.04	0.07	0.04	0.04	547
	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	
2-year effect	–0.06	–0.04	0.02	–0.03	–0.11	0.01	349
	(0.06)	(0.06)	(0.06)	(0.07)	(0.06)	(0.06)	
3-year effect	–0.08	–0.01	–0.07	0.00	–0.09	–0.05	298
	(0.07)	(0.06)	(0.07)	(0.07)	(0.07)	(0.07)	
Median							
1-year effect	0.02	–0.01	0.03	0.01	–0.03	0.02	547
	(0.03)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)	
2-year effect	0.03	0.01	0.03	0.04	–0.02	0.03	349
	(0.03)	(0.05)	(0.04)	(0.05)	(0.04)	(0.05)	
3-year effect	–0.05	–0.01	–0.07	–0.02	–0.07*	–0.02	298
	(0.06)	(0.05)	(0.04)	(0.05)	(0.04)	(0.05)	

Note: Data are pooled across cohorts and the difference-in-difference is based on propensity score nearest neighbor matching.

Table 10. Effect of the 1605(b) Program on the Natural Log of Fuel Expenditures over Different Horizons

Model (all models include logged value of shipments, cost of fuels, cost of electricity, and growth in shipments)							Matched sample
Industry	x	x				x	
Region	x	x			x		
Quadratic	x			x	x	x	
Mean							
1-year effect	–0.04*	0.00	–0.01	–0.02	0.00	–0.04	581
	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)	
2-year effect	–0.03	–0.03	–0.10*	0.00	–0.08	–0.05	388
	(0.04)	(0.04)	(0.04)	(0.05)	(0.04)	(0.04)	
3-year effect	–0.07	–0.11*	–0.04	–0.03	–0.01	0.05	336
	(0.04)	(0.05)	(0.05)	(0.05)	(0.05)	(0.06)	
Median							
1-year effect	–0.04*	–0.01	–0.03	–0.03	–0.02	–0.03*	581
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	
2-year effect	–0.03	–0.04	–0.05	–0.01	–0.05	–0.03	388
	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	
3-year effect	–0.05*	–0.08*	–0.05	–0.04	–0.03	0.01	336
	(0.03)	(0.03)	(0.04)	(0.03)	(0.04)	(0.02)	

Note: Data are pooled across cohorts and the difference-in-difference is based on propensity score nearest neighbor matching.

Table 11. Effect of the EPA Climate Wise Program on the Natural Log of Fuel Expenditures over Different Horizons

Model (all models include logged value of shipments, cost of fuels, cost of electricity, and growth in shipments)							Matched sample
Industry	x	x				x	
Region	x	x			x		
Quadratic	x			x	x	x	
Mean							
1-year effect	–0.06	–0.03	–0.04	–0.05	–0.05	–0.02	949
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	
2-year effect	0.04	0.02	0.00	0.02	–0.02	–0.04	830
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	
3-year effect	–0.02	–0.06	–0.09	–0.07	–0.07	–0.10	764
	(0.04)	(0.05	(0.04)	(0.04)	(0.05)	(0.05)	
Median							
1-year effect	–0.01	0.00	–0.02	–0.02	–0.01	0.00	949
	(0.03)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	
2-year effect	0.03	0.03	0.02	0.01	–0.01	–0.03	830
	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	
3-year effect	–0.01	0.01	–0.10	0.01	–0.04	–0.04	764
	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	

Note: Data are pooled across cohorts and the difference-in-difference is based on propensity score nearest neighbor matching.

Table 12. Effect of EPA Climate Wise Program on the Natural Log of Fuel Expenditures over Different Horizons

Model (all models include logged value of shipments, cost of fuels, cost of electricity, and growth in shipments)							Matched sample
Industry	x	x				x	
Region	x	x			x		
Quadratic	x			x	x	x	
Mean							
1-year effect	0.05*	0.06*	0.06*	0.04	0.04	0.08*	1004
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	
2-year effect	0.04	0.05*	0.05	0.03	0.05*	0.01	888
	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	
3-year effect	–0.01	0.01	0.00	0.03	0.02	–0.02	837
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	
Median							
1-year effect	0.00	0.02	0.03	0.01	0.01	0.02	1004
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	
2-year effect	0.02	0.05*	0.03	0.03*	0.01	0.02	888
	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	
3-year effect	–0.01	–0.01	–0.01	0.02	0.00	–0.02	837
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	

Note: Data are pooled across cohorts and the difference-in-difference is based on propensity score nearest neighbor matching.