DISCUSSION PAPER

August 2007 ■ RFF DP 07-[##]

Evaluating Voluntary Climate Programs in the US

Richard Morgenstern, William A. Pizer, 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 US

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

Abstract

Voluntary programs are playing an increasingly important role in environmental management. Despite their growing importance, however, they have been subject to limited evaluation. As is well known, 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 research is designed to overcome these problems by measuring the environmental effectiveness of two voluntary climate change programs -- EPA's Climate Wise program,and DOE's Voluntary Reporting of Greenhouse Gases Program, 1605(b) -- with particular attention to the participation decision and how various assumptions affect estimates of program effects. For both programs, the analysis focuses on manufacturing firms and uses confidential Census data to create a comparison group as well as 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% and likely less than 5%. There is virtually no evidence of a statistically significant effect of either Climate Wise or 1605(b) on fuel costs. There is some statistically significant evidence that participation in Climate Wise led to a slight (3-5%) increase in electricity costs that vanishes after two years. There is also some statistically significant evidence that participation in 1605(b) led to a slight (4-8%) decrease in electricity costs that persists for at least three years.

Key Words:

JEL Classification Numbers:

Contents

Evaluating Voluntary Climate Programs in the US

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

Introduction

Voluntary programs have been a key part of U.S. climate change policy since the early 1990s. Such programs figured prominently in President Bush's 2002 climate change policy announcement, referencing recent agreements with the semi-conductor and aluminum industries and leading to the creation of the Climate Leaders and Climate Vision programs (White House 2002, 2005). They were also 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), and Climate Wise. In fact, a number of these programs were initiated in the George W. Bush Administration. A 2005 survey identified 87 voluntary programs at 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% of EPA's operating budget. Dozens more programs operate at the U.S. Department of Energy (DOE) and other federal agencies and at the state level. While 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, and continue to play a leading role in Japan.

Arguably, the explosive growth in voluntary environmental programs reflects changing societal attitudes about the environment and a growing optimism on the possibility of enhanced cooperation between government and business. It may also reflect the 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. Unlike market-based approaches to environmental management, where the conceptual roots are largely academic, voluntary programs have emerged as a pragmatic response to the need for more flexible ways to protect the environment.

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

The prominence of voluntary programs begs the obvious question of whether or not these programs are effective in achieving the stated goals. For example, following the 2002 announcement by President Bush of voluntary efforts to achieve an 18% improvement in greenhouse gas intensity, recent data indicating that we are track to meet that 18% goal has been cited as evidence of the voluntary program's success (White House 2006). But is the data really evidence of the voluntary program's success, or does it simply reflect other coincident events?

The key issue considered in this research is whether or not 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? 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? Unfortunately, the existing literature—which primarily emphasizes the motivation of firms to participate rather than the environmental accomplishments of the programs—provides only limited answers to these questions.

This research addresses these questions by conducting a detailed analysis of two early voluntary programs: the U.S. EPA's Climate Wise program, and the U.S. DOE's Voluntary Reporting of Greenhouse Gases Program, 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. While 1605(b) is strictly a reporting program, both programs emphasize flexibility for the participants. Arguably, Climate Wise offers more 'carrots' in the form of technical assistance and public acknowledgement. It also imposes more 'sticks' in the sense of at least an implied expectation that participating firms will make larger emission reductions. Climate Wise is oriented entirely to non-electric utility firms, 1605(b) is open to a broader range of entities. However, because of the nature of our matching sample, (described below), our analysis of both Climate Wise and 1605(b) is limited to firms in the manufacturing sector.

Our focus is on the environmental effectiveness of these programs, with particular emphasis on the participation decision and how various assumptions affect estimates of program effects. 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. We consider two alternative approaches to evaluating outcomes, attempting to control for self-selection in joining the voluntary programs. In one approach we consider a model where program participation depends on both observed and unobservable variables that may be correlated with the outcome. In the other approach, we match participants to appropriate non-participants and consider pairwise differences. The latter method is referred to as propensity score matching.

Understanding the true effectiveness of these programs is important. Protagonists and antagonists of the trend toward voluntary approaches are increasingly at odds, sometimes drawing opposite conclusions about the same program. Protagonists, typically on the side of industry, see voluntary programs as a more practical, flexible approach to regulation. Antagonists, 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. While 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.

Background

In principle, voluntary programs offer opportunities for business 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, customers, investors, communities, employees, and other firms. In some cases, the firms' 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 firms' 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 provide opportunities to achieve environmental improvements more quickly, and with lower administrative costs, than otherwise possible and, sometimes, via more holistic approaches than the media-specific, end-of-pipe focus of most existing legislation. In the view of some observers, by encouraging proactive approaches from industry, voluntary programs may help foster a common understanding of both environmental problems and the mutual responsibilities to address them.

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. The term "regulatory capture" applies when the targets established for the voluntary programs reflect only a business-as-usual scenario. Free riding, wherein some firms avoid any effort while other, proactive firms voluntarily address a problem and keep further regulation at bay, may be an issue with certain voluntary programs. Taking this a step further, voluntary approaches may represent a shift in emphasis from the “worst” polluters to those most willing to abate on their own initiative. Some, particularly in the environmental community, see voluntary programs as a distraction from the real work of taking mandatory action.

Since business is inherently dynamic, with firms constantly confronting new challenges, opportunities, and technologies, it is not sufficient to simply look at two distinct points in time to see if firms’ environmental performance has improved. Rather, environmental gains must be assessed with reference to a credible representation of what would have happened otherwise. Defining such a baseline is, of course, quite difficult to do. 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 non-participants. Still, biases may arise if participants and non-participants 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 non-participants may reflect factors other than the effects of the program.

If we imagined a laboratory setting, 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 different outcomes than those randomly assigned to a control group. Because the two groups would be otherwise identical (due to randomization), this would yield an unbiased estimate of the effect of the voluntary program on environmental performance. In real life, we rarely see such randomized experiments and are instead left with either forecast baselines or imperfect control groups. This provides 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 contains a variety of descriptors to identify particular mechanisms: 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 commitment are determined:

- *Unilateral agreements by industrial firms*. Business-led corporate programs fall under this heading, as do commitments or reduction targets chosen by firms or industry associations. Examples of such agreements include the Chemical Manufacturers Association's "Responsible Care" program for reducing chemical hazards and McDonald's replacement of its Styrofoam clamshell containers with paper packaging.
- *Public voluntary programs*, in which participating firms agree to protocols that have been developed by environmental agencies or other public bodies. Although the public agencies may promote the programs to industry, they generally do not negotiate over the specific terms. Eligibility criteria, rewards, obligations, and other elements are established by the public agencies. Examples of such programs in the United States include the 33/50 program, Climate Wise and 1605(b).
- *Negotiated agreements*, consisting of a target and timetable for attaining the agreed-upon environmental objectives, are created out of a negotiation between government authorities and a firm or industry group over specific terms. In some cases, participating firms receive relief from an otherwise burdensome tax, making the voluntary notion of the program somewhat hazy. Sometimes, firms are held liable for compliance on an individual basis while in others, such as Japan, industries generally are liable on a collective basis for the environmental performance stipulated in the agreements. The XL program is an example of a negotiated agreement.

Economic analysis suggests that since environmental mitigation typically is not costless and the benefits 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 as measured by the number of articles or books published on the subject, 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 firms' public image; the importance (or unimportance) of technical assistance and financial incentives to the firms' participation decision; 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 the importance of public recognition provided by participation in a voluntary program to be a key motivation for firms.[1]

While the literature on the motivation for firms to participate in voluntary programs is extensive, there are only a limited number of previous analyses of environmental performance. The largely theoretical work on the issue suggests that participation in voluntary programs does not guarantee an improvement in actual performance. While it may encourage the exchange of information about best practices, a key factor may be to provide 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 is successful in staving 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 relative to mandatory regulation; the likelihood that regulation will be imposed even 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.

In considering environmental performance of voluntary programs it is useful to distinguish between those programs that focus on the adoption of particular technologies (e.g., Green Lights, now part of the Energy Star Program) and those that focus directly on environmental performance (e.g., 33/50, Climate Wise, 1605(b), or various audit-based programs). In the former, success is measured as adoption of specific technologies. In the latter, it is measured as a reduction in emissions. In both cases, there is the need to define a baseline: Measured over the same period, how many firms (or households) would have installed the technologies, or how much would emissions have been reduced, even without the voluntary program?

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. Despite this limitation, a number of these programs have been subject to at least some evaluation. The Green Lights Program is an innovative, voluntary, pollution-prevention program sponsored by the U.S. EPA focused on the installation of energy-efficient lighting

[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).

where profitable and where lighting quality can be maintained or improved. DeCanio (1998) finds that the energy-efficiency investments carried out under this program yielded annual real rates of return averaging 45%. DeCanio and Watkins (1998) find that specific characteristics of firms affect their decision to join Green Lights and commit to a program of investments in lighting efficiency.

Energy Star is a voluntary labeling program designed to identify and promote energy-efficient products to reduce greenhouse gas emissions. Dowd et al. (2001) cite specific product-purchase decisions being influenced by Energy Star, including a number of favorable “soft” and “dynamic” effects associated with the program. 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 likely would have taken place even without the programs.

The empirical evidence is more extensive, though still mixed, when we look at programs focused explicitly on environmental performance as opposed to technology adoption, particularly with regard to toxics where there has been extensive analysis using TRI data. What is probably the gold standard in the field is an in-depth analysis of the 33/50 program by Khanna and Damon (1999), who jointly modeled the decision to participate in the program as well as the actual outcomes. They first recognize that a firm’s decision about the quantity of covered releases to emit will likely 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 both 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, as well as on firms’ return on investment and long-run profitability. Khanna and Damon 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 for the years 1993–1995. A study by Gamper-Rabindran (2006) found that while the effects varied by industry, in the case of the largest participating industry, namely the chemical industry, the positive results that 33/50 reduced toxic releases (reported by Khanna and Damon (1999)) are 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 that participants were reducing their releases more slowly than non-participants. 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 pre-program years and, most surprisingly, it was slower for participants in Responsible Care than for non-participants.

A paper by Dasgupta, Hettige and Wheeler (1997) focused on 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 non-environmental workers led to an improvement in the compliance status of the firms.

Turning to energy and climate change, an analysis of the U.S. Department of Energy's Climate Challenge Program on CO2 emissions focused on the largest 50 electric utilities east of the Rocky Mountains from 1995–1997 (Welch, Mazur, and Bretschneider 2000). Despite a number of intriguing results about the motivation of firms to participate in Climate Challenge, the authors find that adoption of the program seems to have no effect on emissions. In fact, those firms predicted to volunteer higher reduction levels were found to reduce their CO2 emissions less. The authors hypothesize that the poor program performance is associated with the lack of at least a tacit regulatory stick of the type present in 33/50. A recent paper examined the performance of electric utilities participating in the 1605(b) program (Lyon and Kim 2007). 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 non-participants. 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. As is well known, energy issues differ from toxics in many ways, including the extent to which financial incentives are already in place to reduce emissions. That is, market forces already encourage conservation and energy efficiency, whereas no such forces exist to reduce toxic emissions. Thus, the potential for voluntary programs to achieve reductions in energy-related carbon dioxide emissions may be more limited than the potential associated with toxics. A key motivation for this research is to increase the attention paid to the rigorous study of program results and to emphasize rapidly growing interest in energy- and greenhouse-gas-related programs.

The Climate Wise Program

Officially established by the U.S. Environmental Protection Agency (EPA) in 1993, Climate Wise is a voluntary program focusing on the non-utility industrial sector to encourage the reduction of carbon dioxide (CO2) and other greenhouse gases (GHGs) via adoption of energy efficiency, renewable energy and pollution prevention technologies. Climate Wise remained in operation until 1999-2000 when it was renamed and placed under the Agency's Energy Star umbrella. Unlike Green Lights or EPA's other technology-based programs which require the adoption of particular technologies, Climate Wise members had the flexibility to use whatever technologies or strategies they chose to reduce their emissions. The basic requirements of Climate Wise were that a participating firm develop baseline emission estimates of its GHGs, pledge forward looking emission reduction actions, and make periodic progress reports. As part of the program operations, Climate Wise provided public recognition and certain types of technical assistance to its members. At its peak, Climate Wise had enrolled more than 600 industrial firms covering several thousand facilities nationwide. More recently, Climate Leaders, a program noted above with some design features similar to those of Climate Wise, has been embraced by the Bush administration as a key element of its climate change initiative. Although EPA has developed estimates of the emission reductions associated with Climate Wise, there has been little outside evaluation of the program.

As stated in the Program's 1998 Progress Report the four broad objectives of the Climate Wise Program are 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 organization; and
- Develop productive and flexible partnerships within government and between government and industry. (EPA 1998 page 2)

Climate Wise consists of three interrelated components. First, the pledge component asks firms to commit to taking cost-effective, voluntary actions to reduce greenhouse gas emissions. Second, the tailored assistance efforts are designed to facilitate companies' emission reducing efforts via a clearinghouse, workshops, and seminars. Finally, communication activities provide public recognition for actual progress in reducing emissions.

To join Climate Wise, a firm has to develop a baseline estimate of its direct emissions of CO2 (and other greenhouse gases) for the year it joined the program or any prior year of its choice since 1990. Since an estimate of baseline emissions estimate does not involve the detailed accounting information required for a full emissions inventory, the burden on the firm was relatively modest.

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 this is 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. Also included were fuel switching and best management practices, as well as 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 for the year 2000, and to provide a progress report directly to EPA. Participants were also encouraged, but not required, to report their progress to the U.S. Department of Energy through the 1605(b) registry program.

EPA provided several types of technical assistance to participating firms, including a guide to industrial energy efficiency, various government publications on energy efficiency and related issues and, most importantly, free phone consultation with government and private sector energy experts retained as consultants 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, low interest buy-downs from state providers, utility programs, and others. Further, EPA set up an annual event open to the public to recognize the performance of outstanding Climate Wise participants.

Although the focus of the Climate Wise program is on energy efficiency and the reduction of CO_2 emissions, a number of firms did propose to reduce emissions of non-CO_2 greenhouse gases 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 were released in the manufacture of adipic acid. Significant amounts of methane (CH_4) were also included in the action plans of several firms, especially in the beer industry.

The 1605(b) Program

Unlike Climate Wise, which was initiated entirely by the EPA, section 1605(b) of the Energy Policy Act of 1992 (EPACT) directed the DOE to develop a program to document voluntary actions that reduce emissions of greenhouse gases or remove greenhouse gases from the atmosphere. The Voluntary Reporting Program was to be administered by the U.S. Energy Information Administration (EIA). The EPACT mandated that EIA issue guidelines for reporting, establish suitable procedures, ensure confidentiality of trade secrets, commercial and financial information, and establish a publicly available database. It also mandated consultation with the 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 greenhouse gas emissions or sequester carbon voluntarily, and it attempts to identify innovative and effective ways of reducing emissions. Most of the reporters to the Voluntary Reporting Program are affiliated with one or more EPA or other government-sponsored voluntary programs.

As originally developed, the 1605(b) program is extremely flexible. Both direct and indirect emissions, including sequestration, can be included. Voluntary reporters can define the boundary of the entity or the project, and can choose to report reductions at the entity or the project level. Reporters 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 what emissions (or sequestration) would have been in the absence of the project. Further, reporters 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 per year to more than 200 per year;

the number of projects has more than tripled from about 600 per year to more than 2000 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. 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 research 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 program's database offers abundant examples of the types of concrete actions that organizations report to reduce greenhouse gas emissions. The EIA notes some of the most important benefits of the 1605(b) Program as follows (EIA 2002 pp 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 Census data. As noted, we focus exclusively on the manufacturing sector.

The Climate Wise Program

For EPA's Climate Wise Program data, a list of voluntary program participants was obtained from EPA describing who joined Climate Wise in each of its operational years from 1994 to 2000. This list includes name, zip code and join date data for two different types of participants, those who joined at the corporate level and those who joined as individual plant participants. There were a total of 671 participants with complete data. 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 at the Census Bureau 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 participants to the Census data reflects the fact that Census data only includes manufacturing establishments, while the Climate Wise program includes both manufacturing and non manufacturing participants (e.g., municipalities, commercial buildings, etc. – despite its programmatic focus on manufacturing).

These 377 linked participants from the original Climate Wise list translate into 2311 facilities because corporate participants can have multiple associated facilities. The data are displayed in Table 2.

Summary statistics for the linked sample, as well as the entire Census database, are given in 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 Census database—roughly 1%. This suggests that the full Census sample is unlikely to be an appropriate control group as a whole, and that there are a large number of plants from which to choose a more appropriate sub-group of controls.

It is worth noting that the linking Climate Wise and Census data has important consequences for our ability to evaluate the effect of program participation over longer horizons. As we are attempting to study behavior 2 or 3 years after joining, we are forced to drop plants that joined in 2000 and 1999, respectively, because our Census data ends in 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 looking 2 to 3 years out. However, trying to discern effects 4 years after joining, with only participants who joined between 1994 and 1997, we have noticeably fewer observations and noisier estimates. Thus, we do not attempt to look at effects more than 3 years after participants join the program.

The 1605(b) Program

For DOE's Voluntary Reporting of Greenhouse Gases (1605(b)) Program, a list of reporting entities, sectors, years reported, and form type used was obtained 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. Most of the participants are in the energy relevant sectors. For example, the electric power sector accounts for more than one third of total reporting entities (130 out of 383). In this research, we are most interested in manufacturing participants which account for only 18% of the all reporting entities. In Table 4, we provide sector distribution information for all the reporting entities.

Unlike EPA's Climate Wise program, DOE 1605(b) data does not have join date information. However, as noted, the year and type of form reported for program participants are available in the database. Thus, we use the first reporting year as the join year and assume that the participants continue in the program after that, even though individual entities may not have continuous reporting years. Table 5 displays the join year information based on either firm or plant participation.

A separate entity file was also obtained from EIA. It contains 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 Table 6, we show the sector distribution for DOE 1605(b) and LRD matching results. For the industrial sector, the matching rate is about 77%. For sector classified as N/A, we were able to match 36% of them. For others, the matching rate is only 13%, because most of the others are electricity and energy relevant entities which do not fall into the manufacturing category. Due to the small number of observations, items marked with D* are included in the 'Other' category.

After excluding missing join year and others, we were able to link 83 out 383 participants, including 67 corporate participants and 17 plant participants. We have a much lower matching rate for the DOE 1605(b) program because it includes both manufacturing and non manufacturing sectors. In fact, more than 50% of the participants are in the electric power and alternative energy sectors, which are not in the Census data. These 83 linked participants from the original DOE 1605(b) list corresponds to 1791 LRD facilities because corporate participants can have multiple facilities. Table 7 summarizes the matching of the 1605(b) data to

Longtitudinal Research Database (LRD). Table 8 provides summary statistics for the linked sample, as well as the entire Census database for the DOE 1605(b) voluntary program.

Models and Econometric Method

With the linked Census data described in the preceding section, we have access to variables indicating energy expenditures (separately on 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 our two programs and what year they first participated. We now consider two alternative approaches to evaluating outcomes, attempting to control for selection 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 non-participation, $Y_i(0)$. Here, $Y_i(D_i)$ is the outcome associated with either treatment, $D_i = 1$, or non-treatment, $D_i = 0$, and is either the cost of fuels or electricity measured in natural logarithms. The ideal study would measure the treatment effect,

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

for each plant i, that is, the percent 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 selection process determining 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). Another way to say this is that the selection mechanism determining which outcomes are observed is *ignorable*. Under this assumption, 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, the average outcome among those participating minus the average outcome among those not-participating. Or from a simple regression model,

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

where by assumption u_i is uncorrelated with D_i and the treatment effect is the estimated value of β_1. Of course, in reality, missing at random is unlikely and hence we proceed to our two approaches.

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

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

with u_i being an unobserved variable. We build a structured model where, 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, even conditioning on X_i. In particular, we assume u_i and D_i to be correlated, thus violating a key assumption for unbiased estimation in an OLS model (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 but does not directly influence the emission outcome. For example, if we find that in some years individual programs were more aggressively marketed than 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 but not directly influence the emissions outcome.

If we assume (u_i, v_i) are jointly normal, it is easy to show 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 intuition for the identifying assumption that there must be at least one variable in Z_i excluded from X_i: Except for the non-linearity in the function λ, the right-hand side variables would be co-linear if X_i included all the variables in Z_i.

Our dependent variable in this model is the change in logged 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 logged value of shipments over the given time horizon as a control variable. While this is arguably endogenous, we believe controlling for growth is critical: we observe that faster growing plants are more likely to join voluntary programs. It seems unlikely that this growth is *caused* by joining; therefore, we need to control for it.

We also include dummy variables for Census region and two-digit industry classification. Our Z_i variables include two variables we believe are likely to influence participation but not the outcome – distance to the nearest regional EPA office and local membership rates in a national environmental organization.[2] We discuss the results of this approach in the results section that follows.

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

$$D_i \perp Y_i\left(1\right), Y_i\left(0\right) \big| X_i$$

This could be accomplished 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 (and others) match participants to appropriate non-participants and consider the pairwise differences. While the Heckman and Hotz approach

---

[2] Many thanks to the National Wildlife Federation for supplying this data.

attempts to control for selection on additional, unobserved effects correlated with outcome, it requires both on a correct specification and identification of one or more excluded variables. This approach, while not controlling for such effects, relaxes the specification assumption and does not require excluded variables.

While the general problem of creating a set of matched, non-participating observations is quite challenging (there are many observable variables – in our case describing location, industry, size, energy intensity, and growth – that we would want to match), the important result based on Rosenbaum and Rubin is that we only need to match the expected likelihood of participation. That is, we simplify the difficult problem of matching all these different variables to a much simpler one of matching a summary variable describing the propensity to join the program. This approach is referred to as propensity score matching.

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 2-digit industry classification. As before, we also include a term for growth in value of shipments over a given horizon h, as this turns out to be an important determinant of participation. As it seems unusual to imagine participation causing growth, we take this as a proxy for expected growth over the given horizon. We use samples matched with different horizons h to estimate program effects over similar horizons.

Because each of the voluntary programs lasted a number of years, it seems more natural to think about 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.[3] We therefore estimate a Cox proportional hazard model of the form:

[3] Note that while the participants have an obvious join-year associated with them, non-participants do not. That is, there are different years when plants begin participation, but not when they begin non-participation. Outside of a duration model, it does not seem possible to combine the data. In the previous approach, we estimate effects for different cohorts of participants separately for this reason.

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

Once estimated, we predict hazard rates for participants in the year they join and match them to the nearest valued non-participant in that year. We then examine the difference in changes in cost of fuels and electricity across each pair; this difference-in-differences forms the estimate of our program effect.

Results

Table 9 through Table 12 present results for the first, Heckman-Hotz approach for the DOE 1605(b) and EPA Climate Wise programs, and both cost of fuels and cost of electricity as outcome variables, respectively. We report only the results for a "two-year" horizon; that is, the dependent variable measuring the change between the year before a group of participants join the program (cohort) and two-years later; the results are broadly similar at one- and three-year horizons. We have reported the results with and without the selection correction term, $\lambda(D_i, Z_i)$ in (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, and cost of electricity, growth in value of shipments, as well as region and industry dummies) 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 of less than 10%. The three exceptions are a statistically significant 9% decline in electricity costs among 1605(b) participants in the 1994 cohort, a 6% increase in electricity costs among Climate Wise participants in the same cohort, and a *55% 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 might increase in Climate Wise if efforts to reduce direct emissions lead to more electricity use and higher indirect emissions. Similarly, a positive electricity effect could reflect a combination of specification error and the fact that larger / faster growing firms tend to participate in voluntary programs. The latter, 55% effect likely 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 purpose of this exercise. The second column presents results where first we estimate the probability of selection, and then include the selection correction term, $\lambda\left(D_i, Z_i\right)$ in (2), in the original outcome regression. For simplicity, we have not reported the results of the first stage regression. However, an important observation in these results is that the excluded variables are *almost never* statistically significant.[4] Empirically, it is difficult to see a difference in the distribution of either variable among participants and non-participants, suggesting 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. We also see much larger standard errors on the estimates, compared to the simple estimates in column one. This reflects the likely multi-collinearity between the correction term and the right-hand-side variables, where it may mostly be the nonlinearity of the $\lambda\left(D_i, Z_i\right)$ function identifying the parameters rather than the excluded variables. In any case, 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 EPA 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% to +82% across different years). Given that such a divergent range driven by the participation year seems implausible, and the noted problem with the excluded variables, we tend to distrust this approach.

Instead, we now turn to the results from the propensity score matching approach in Table 13 through Table 16. As noted in the methods section, we estimate duration model for whether or not facilities join, using a variety of specifications. These specifications differ based on whether dummies are included for industry and region, and whether or not quadratic terms are included, as indicated in the top three rows of each table. For each specification, we consider effects over 1, 2, and 3 years; we pool across all cohorts of matched pairs, and report both the mean and median across pairs.

As with the simple model (column 1 of Table 9 through Table 12), all of the estimates suggest effects of less than 10%. We focus our discussion on the median estimates in the bottom

[4] 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.

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% in magnitude, suggesting any effect is probably even smaller than 10%. There is generally more statistical significance among the electricity cost estimates (6 of 36) versus fuel cost estimates (1 of 36). Interestingly, the 1605(b) program seems to have a negative effect of perhaps several percent (Table 14, where 17 of 18 median estimates are negative), while Climate Wise appears, if anything, to have a slight positive effect (Table 16, where 14 of 18 median estimates are positive). The positive effect in Climate Wise is not present in our most general matching model (Table 16, where median estimates in column 1 are not significant); the negative in 1605(b) is present (Table 14, where median estimates in column 1 are signficant). Further, there is no evidence of persistence in the Climate Wise results (effects at 3 year horizon are all lower than at 2 years; see bottom 2 rows of Table 16). Meanwhile 1605(b) estimates in 4 out of 6 models are largest for the longest horizon (bottom row of Table 14).

Putting this all together, we have several key observations based on the simple results (column 1 in Table 9 through Table 12) and propensity score matching approach (Table 13 through Table 16). As noted earlier, we tend to distrust the Heckman-Hotz results.

1. Voluntary program effects from Climate Wise and 1605(b) on fuel and electricity expenditures are no more than 10% and likely less than 5%.
2. There is virtually no evidence of a statistically significant effect of either Climate Wise or 1605(b) on fuel costs.
3. There is some statistically significant evidence that participation in Climate Wise led to a slight (3-5%) increase electricity costs that vanishes after two years.
4. There is some statistically significant evidence that participation in 1605(b) led to a slight (4-8%) decrease in electricity costs that persists for at least three years.

Among these results, the transient, slight increase in electricity costs under Climate Wise is certainly anomalous. Two explanations come to mind. First, participating plants may have pursued direct emission 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 be not show up in the cost of fuel measure because fuel switching among purchased fuels – for example, a shift to biomass or from coal to gas – might reduce emissions without changing expenditures. Or plants may have pursued non-energy-related emission reductions -- such as 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.

A second explanation for a positive effect on electricity is that it may reflect a failure to adequately control for growth. While we match, in part, on growth in the value of shipments, the tendency of faster growing firms to join remains troubling. For example, we have no way of knowing about the underlying prices and quantity changes – participants might experience changes in quantities while those matched from the Census database might experience changes in prices; we cannot tease out controls that have that same pattern because there is no available detail on prices and quantities. If the estimated electricity expenditure growth effect is really reflecting 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. The present research, which examines both a DOE- and an EPA-sponsored program, relies on confidential plant-level data for the manufacturing sector collected by the U.S. Census Bureau to develop such a baseline based on a comparable set of non-participant controls, focusing on activities through 2001 when available Census data ends. We consider two alternative approaches to evaluating outcomes, attempting to control for selection in joining the programs based on both observable as well as unobservable characteristics. In one approach we consider a structural model where program participation depends on both observed and unobservable variables that may be correlated with the outcome. In the other approach, we match participants to appropriate non-participants based on observable characteristics only, and consider pairwise differences – a method known as propensity score matching. The results are sobering.

In contrast to the claims of relatively large emission reductions reported by the sponsoring agencies, our analysis suggests more modest reductions are attributable to the programs studied. Overall, we find that that the effects from Climate Wise and 1605(b) on fuel and electricity expenditures are no more than 10% and likely less than 5%. There is no evidence of reductions in direct emissions from fossil fuels attributable to the voluntary programs; however, there is some statistically significant impacts on the use of electricity. In particular, there is some statistical evidence that participation in 1605(b) lead to a slight decrease in electricity expenditures, on the order of 4-8 percent. This decrease persists for at least three years. The statistically significant evidence on Climate Wise is that the program 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 in all likelihood, participation in Climate Wise 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 the EIA observation that most of the entities reporting under 1605(b) are also affiliated with one or more other government-sponsored programs. Thus, the observed emission reductions for the 1605(b) reporters may reflect the influence not only of the 1605(b) program itself but also that of other programs. While our separate assessment of Climate Wise suggests that participation in that program is not likely associated with significant emission reductions, other larger programs, e.g., 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 the firm participates.

Methodologically, our research highlights the inevitable complexity of assessing voluntary programs. Our research 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 non-participants and considers pairwise differences. Because the Heckman and Hotz approach requires both a correct specification and identification of one or more excluded variables, it is more demanding than the Rosenbaum and Rubin approach 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 our difficulty identifying excluded variables in the former method, our results seem more plausible with use of the latter approach, and we think such an approach may have wider application in the future evaluation of voluntary programs.

Overall, the evaluation of environmental programs seeks to determine what works and what does not. Our findings of at most a small effect should not be all that surprising. Energy-related greenhouse gas emissions are quite different than many other types of emissions, e.g., un-priced industrial byproducts such as toxics with no near-term localized effects whose existence was widely ignored until the 1980s and 1990s. With no practical opportunity for end-of-pipe abatement, reductions in energy-related greenhouse gas emissions often amount to reductions in energy use itself – something that has been picked over for some time. Given the underlying positive price on energy, there is *always* 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. 1996. Why do Firms Volunteer to Exceed Environmental Regulations? Understanding Participation in EPA's 33/50 Program. *Land Economics* 72: 413–432.

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

Bazelon, Coleman and Kent Smetters, "Discounting Inside the Beltway." Journal of Economic Perspectives, 1999, 13, 4: 213 - 228.

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).

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

Comment [MSOffice1]:

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 (11 pages).

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.* 2006.

Howarth, R., B. Haddad, and B. Paton. 2000. The Economics of Energy Efficiency: Insights from Voluntary Programs. *Energy Policy* 28: 477–846.

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. 2002. Does Membership Have its Privileges? Working paper.

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):944-52.

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.

Lyon, T., and Kim E. 2007 "Greenhouse Gas Reductions or Greenwsh?: The DOE's 1605b Program," University of Michigan, mimeo, April.

Comment [MSOffice2]: Page: 23
Need chapter page numbers.

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," Working paper, University of Arizona.

U.S. Energy Information Administration. 2002. *Voluntary Reporting of Greenhouse Gases 2000,* Office of Integrated Analysis and Forecasting, Washington, D.C. (February). www.eia.doe.gov/oiaf/1605/vrrpt/index.html.

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

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

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.2005/05/20050518-4.html [accessed August 1, 2006.], 2005

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

Tables

Table 1: Join Data for Climate Wise Participants

Join year	Corporate	Plant	Subtotal
1994	8	0	8
1995	30	7	37
1996	141	38	179
1997	101	37	138
1998	70	36	106
1999	17	72	89
2000	0	144	144
Subtotal	367	304	671

Table 2: Matching of Climate Wise (CW) to Longtitudinal Research Database (LRD)

	CW List	LRD Plants	LRD plant-year observations (1992-2001)
Corporate participants with multiple plants	135	2,053	11,503
Corporate participants with a single plant	93	95	316
Plant-level participants	149	163	946
Total	377	2,311	12,765

Table 3: Sample Statistics, LRD and Program Participants

Variable	Summary Statistics	Full LRD sample (1992-2001)	Program Participants
ln(*TVS*) (total value of shipments)	Mean	7.61	10.87
	Standard deviation	2.30	1.81
	Plant-year observations	1,157,606	12,605
ln(*CF*) (cost of fuels)	Mean	2.54	5.31
	Standard deviation	2.12	2.23
	Plant-year observations	839,934	11,280
ln(*PE*) (purchased electricity)	Mean	3.17	6.31
	Standard deviation	2.21	1.83
	Plant-year observations	1,019,042	12,377
	Number of Plants	515,189	2,311

Table 4: The Sector Distribution for DOE 1605(b) Reporting Entities

Sector	Counts
Agriculture	12
AlternativeEnergy	63
ElectricPower	130
Industry	69
N/A	94
Other	15
Total	383

Table 5: Join Year for DOE 1605(b) Participants

Join Year	Plant	Firm	Total
.	7	36	43
1994	0	105	105
1995	0	37	37
1996	3	23	26
1997	2	15	17
1998	8	53	61
1999	2	33	35
2000	6	53	59
Total	28	355	383

Table 6: The Sector Distribution for Matched DOE 1605(b) and LRD Data

Sector	Counts
Agriculture	D*
AlternativeEnergy	D*
ElectricPower	D*
Industry	53
N/A	34
Other	13
Total	100

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

Table 7: Matching of DOE 1605(b) to Longtitudinal Research Database (LRD)

	1605(b)	LRD Plants	LRD plant-year observations (1992-2001)
Corporate participants with multiple plants	54	1762	8724
Corporate participants with a single plant	13	13	63
Plant-Level participants	16	16	122
Total	83	1791	8909

Table 8: Sample Statistics, LRD and Program Participants

Varibale	Summary Statistics	Full LRD sample (1992-2001)	Program Participants
ln(TVS)	Mean	7.80	10.99
(total value of	Standard Deviations	2.34	2.17

Table 9: DOE 1605(b) program, effect of program on logged cost of fuels after 2 years, Heckman-Hotz approach

Cohort	w/o correction	with correction	sample	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 10: DOE 1605(b) program, effect of program on logged cost of electricity after 2 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 11: EPA Climate Wise program, effect of program on logged cost of fuels after 2 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 12: EPA Climate Wise program, effect of program on logged cost of electricity after 2 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 13: DOE 1605(b) program, effect of program on logged cost of fuels over different horizons and pooled across cohorts (difference-in-difference based on propensity score nearest neighbor matching)

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)	

Table 14: DOE 1605(b) program, effect of program on logged cost of electricity over different horizons and pooled across cohorts (difference-in-difference based on propensity score nearest neighbor 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	-0.04*	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)	

Table 15: EPA Climate Wise program, effect of program on logged cost of fuels over different horizons and pooled across cohorts (difference-in-difference based on propensity score nearest neighbor 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	-0.06	-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)	

Table 16: EPA Climate Wise program, effect of program on logged cost of electricity over different horizons and pooled across cohorts (difference-in-difference based on propensity score nearest neighbor 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	0.05*	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)	