THE EFFECTS OF FIRM AND FACILITY CHARACTERISTICS ON ENVIRONMENTAL COMPLIANCE

By

Marcella Recher

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Interdisciplinary Studies: Environmental Management

August, 2009

Nashville, Tennessee

Approved:

Professor Mark A. Cohen

Professor David S. Cordray

Professor James H. Clarke

Professor Georgine M. Pion

Professor Mark D. Abkowitz

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to all the people who helped and supported me throughout this endeavor. First of all, I want to express my thanks to Mark Cohen, my committee chair, for being my mentor throughout my time at Vanderbilt. I also want to thank the other members of my dissertation committee David Cordray, Georgine Pion, Mark Abkowitz and James Clarke for contributing their time and providing me thoughtful critiques about my research. I especially want to acknowledge David Cordray for all the hours he spent helping me learn the statistical techniques I needed to analyze my data.

In addition, I also want to express my appreciation to all the people that aided my data collection effort. In particular, I want to thank Yi Zhu at the Florida Department of Environmental Protection, Phillip Barnett at EPA Region 4, Errol Reksten at the Chattanooga-Hamilton County Air Pollution Control Bureau, and everyone else at the local air protection offices I contacted for sending me data. My thanks also to Jamie Jarrell at the Tennessee Department of Environmental Conservation, Betty Gatano at the North Carolina Department of Environment and Natural Resources Division of Air Quality, Jody Hamm and Karla Mew at the South Carolina Freedom of Information Center, and Deborah Parrish and her colleagues at the Memphis-Shelby County Health Department Pollution Control Section who helped facilitate access to the many files I requested. Special mention should also be given to David Konisky, now at the University of Missouri, Department of Political Science who provided the 1st quarter, fiscal year 2000 data on fines.

I am also grateful to all the people who took the time to listen to my ideas, answer questions, and provide me helpful suggestions that allowed me to refine my research. This included all of my Vanderbilt Center for Environmental Management Studies colleagues, Julie Aslinger at the Tennessee Department of Environmental Conservation, Professor Michael P. Vandenbergh at Vanderbilt University Law School, Tim Bent at Bridgestone Americas and Jason Stansbury. I would also like to express my thanks to Anj (Judy) Shaw for providing me project management help and Dr. Robert Cohen who helped me through the last phase of this process. I found it challenging to complete my dissertation away from the academic community at Vanderbilt. I really appreciated the time Dr. Cohen took to review and discuss my numerous revisions. Dr. Cohen’s editorial advice was essential to the completion of this dissertation.

I am also greatly indebted to my friends and family who motivated me to make it to the finish line. This includes Yusong Li, Erin O’Brien, Dr. Susan Raines, Sue Cohen, Paula Guarino, Sha Sha, Gabi Dobrot, Emma Schlesinger-Dawley, Maggie Reynolds, all my colleagues at EPI and all the other amazing people in my life. All of you kept my sprits up on the days I was tired and irritable from putting in such long hours in front of the computer! Special thanks also to my mom and dad who came to cheer me on at my defense and my kitty Bear who kept me company while I typed.

Financially, I would not have been able to pursue my Ph.D. without the generous fellowship I received from Bridgestone Americas. I want to thank Bridgestone Americas for making it possible for me to obtain my Ph.D. I also want to highlight the fact that this research is supported, in part, by a grant from the Vanderbilt University Graduate School.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AFS- Air Facility System

CAA- Clean Air Act

CWA- Clean Water Act

EPA- Environmental Protection Agency

EH&S- Environmental Health & Safety

ECHO- Enforcement and Compliance History On-line

FCE- Full Compliance Evaluation

HAP- Hazardous Air Pollutant

HPV- High Priority Violation

LON- Letter of Non-compliance

NESHAP- National Emissions Standards for Hazardous Air Pollutants

NOV- Notice of Violation

NOx- Nitrogen Oxide

NBRM- Negative Binomial Regression Model

NSPS- New Source Performance Standards

PCE- Partial Compliance Evaluation

PTE- Potential to Emit

PRM- Poisson Regression Model

PSD- Prevention of Significant Deterioration

RCRA- Resource Conservation and Recovery Act

SIC- Standard Industry Classification Code

TDEC- Tennessee Department of Environmental Conservation

TPY- Tons Per Year

TRI- Toxic Release Inventory

VOC- Volatile Organic Compounds

ZINB- Zero Inflated Negative Binomial Regression

ZIP- Zero Inflated Poisson Regression

CHAPTER I

INTRODUCTION

In an effort to ensure environmental quality in the United States, numerous laws have been passed, such as the Clean Air Act, which impose restrictions on manufacturing plants. Government environmental protection agencies are responsible for making sure manufacturing facilities comply with these laws. These agencies want high compliance rates (ideally 100%); however, they have limited monitoring and enforcement budgets to accomplish this aim. Given these budgetary concerns, government authorities and policymakers have expressed interest in identifying characteristics that could help determine which plants should be targeted for compliance assistance and/or additional monitoring and enforcement (OECD, 2004, p. 7; Clark, 2004, p. 20).

To date, there has been little direct research that examines the relationship between firm and facility characteristics and environmental compliance rates. There is, however, research found in the area of economics which could help identify those types of facilities that would be good candidates for additional governmental assistance and oversight (Cohen, 1999, p. 48). The Becker optimal penalty model, which is the foundation of the deterrence literature, suggests that the higher a firm's compliance costs and the lower the expected penalty, the less incentive firms have to comply with government regulation (Becker, 1968, p. 177). It has often been noted in the deterrence literature that "government monitoring activities are often quite limited" and "fines are low" (Cohen, 1999, pp. 46-49). Findings of limited government monitoring and low penalties might imply that high facility compliance costs are influential factors in determining compliance rates. Note that compliance costs do not simply include the physical costs of building and operating pollution control equipment. Included also is anything that makes compliance more difficult or less likely, such as the cost of acquiring information, or the cost of monitoring and controlling managers and employees. Understanding the nature of these additional costs may provide guidance as to which types of firm and facility characteristics affect a plant's ability to comply with environmental regulations.

The focus of this paper is to generate theory-based and empirical evidence that could help pinpoint which types of facilities face higher environmental compliance costs. The next part of this chapter reviews the literature, first outlining the Becker optimal penalty model and associated empirical research that looks at the effect of government monitoring and enforcement on environmental compliance. The next section of the literature review explores how agency costs and capacity costs could affect a plant's ability to comply with environmental regulations. Several hypotheses are then presented to test whether certain types of plants are confronted with higher agency or capacity costs. The methods and results sections follow in chapters two and three and examine evidence that plants with certain characteristics, such as large plants and plants that are part of multi-facility firms, are subject to higher compliance costs. Controlling for government monitoring and enforcement, these plants are shown to be more likely to be out of compliance. In summary, it is suggested that information from this study could assist with the development of future theoretical models of environmental deterrence and aid policymakers in developing more effective environmental compliance strategies.

Becker Model

The Becker optimal penalty model provides a framework to help understand the costs that different types of facilities face when complying with environmental laws. This is a formal model that operationalizes deterrence theory, which suggests that human behavior can be influenced by the use of incentives (Becker, 1968). Becker's model uses economic theory to explain why individuals commit criminal offenses. Using an economic approach, Becker theorizes that a person will commit an offense "if the expected utility to him exceeds the utility he could get by using his time and other resources at other activities" (p. 177). He represents this as:

$$Oj = Oj\ (p_j, f_j, u_j)$$

Oj= the number of offenses he would commit during a particular period
p_j= his probability of conviction per offense
f_j= his punishment per offense
u_j= a portmanteau variable representing all these other influences. Becker notes that the supply of offenses also depends upon "other variables, such as the income available to him in legal and other illegal activities, the frequency of nuisance arrests, and his willingness to commit an illegal act." This suggests, for example, that if there is less income available from legal activities, the number of offenses he commits would go up (Becker, 1968, p.177, equation 12).

Becker's model ultimately distills an individual's decision to commit a crime to a calculation weighing the benefits of the crime against the probability of getting caught and severity and likelihood of punishment if caught (Becker, 1968, p. 177). Although Becker's optimal penalty model is based on an individual committing an act of street crime, Becker envisioned his ideas to apply to a broad range of offenses (p. 170). Becker's insight has been used to help explain why firms comply with government regulations (Cohen, 1999). As applied to firms, this model would lead us to believe that firms, acting rationally to maximize their profits, will comply if the benefits of compliance outweigh the costs (Cohen, 1986, p. 167; Gray & Shadbegian, 2005, p.

242).[1] This means that the higher a firm's compliance costs and the lower the expected penalty, the less incentive firms have to comply with government regulation. Thus, understanding which facilities have high compliance costs could help determine which plants will have more violations, especially if actual government penalties are low, as is often the case (Cohen, 1999, pp. 46-49).

Most of the research in the area of environmental compliance has focused on understanding the effect of government monitoring and enforcement on facility compliance. Few studies have looked at the effect of compliance costs on facility violation rates. In other words, few have looked at what should be included in variable "u_j = a portmanteau variable representing all these other influences" (Becker, 1968, p. 177, equation 12). The next section will review empirical work that looks at the effect of government monitoring and enforcement on environmental compliance rates, followed by an examination of theory and empirical work that explores the relationship between compliance costs and violations rates.

Deterrent Effect of Government Monitoring and Enforcement on Environmental Compliance

Several economics studies covering different media such as oil spills, air pollution, water pollution, and hazardous waste management, have tested the effect of government monitoring and enforcement on compliance rates. The results of these studies help to understand the practical application of the Becker model framework in the area of environmental compliance.

[1] Based on Becker's model, an easy way to reduce government monitoring and enforcement costs would be to increase the penalty to the point where no one would risk violating the law; thus little monitoring and enforcement would be required. This would mean that the few violators that were caught would be severely punished. Cohen points out that such severe penalties are not realistic for several reasons: the limited wealth of the offending firm could lead to a firm becoming insolvent, legislation that imposes a limit on penalties due to social norms of fairness, and risk averse firms may choose not to operate if there were extreme negative payoffs. As a result of these constraints, government monitoring and enforcement expenditures are necessary (Cohen, 1999 sec. 3.2-3.3).

Oil Spills

Epple and Visscher (1984) studied whether enforcement policies of the Water Quality Improvement Act of 1970 reduced the frequency of oil spills by tanker ships and barges. The law made polluters responsible for clean up of discharges of oil and other hazardous substances into U.S. inland and coastal waterways (p. 30). Polluters could also be subject to civil penalties for "intentional spills resulting from negligence" (p. 30). Using data from the U.S. Coast Guard Pollution Incident Reporting System spanning from 1973-1977, they estimated, based on their results, that a 10% increase in Coast Guard Monitoring yields a 3.1% reduction in spill volume (p. 40).

Grau and Groves (1997) also looked at the effectiveness of U.S. Coast Guard enforcement from January 1984 to December 1987. They obtained the data from the U.S. Coast Guard Pollution Incident Reporting System and U.S. Coast Guard Quarterly Report (p. 320). Variables were created for the number of monitored transfer operations and fines (p. 322). It was found that Coast Guard monitoring decreased both the frequency of oil spills and the spill size; however, expected fines did not have an effect on either the frequency or size of oil spills as expected (pp. 315, 355). A later study by Weber and Crew (2000) that used U. S. Coast Guard data from 1992-1996 observed that "enforcement offices with the lowest fines exhibited the greatest amount of oil spilled" (p. 165). They also found that "a 10-day improvement in the speed with which penalties are assessed will reduce oil spillage by 0.6% and selected increases in the severity of punishment for oil spills can reduce spillage by 10%" (p. 161). These studies provide some evidence that monitoring affects spill rates, but mixed evidence that fines have the intended effect.

Water Pollution

Magat and Viscusi (1990) looked at the effect of inspections and associated enforcement on water pollution discharge levels of firms in the pulp and paper industries. They had quarterly plant data from 1982-1985 that they obtained from EPA's Permit Compliance System (pp. 339, 342. Inspections were used as a proxy for both monitoring and associated enforcement actions (p. 339). They found that inspections and associated enforcement actions had a strong effect on compliance with "both pollution levels and rates of compliance with permit levels" (p. 359). Inspections also induced plants "to report more regularly their discharge levels to the EPA or designated state enforcement agency" (pp. 354-355, 359).

A later study by Laplante and Rilstone (1996), using data from the Quebec Ministry of the Environment spanning from 1985 to 1990, looked at the impact of monitoring strategies (inspections) on pulp and paper plants compliance with water pollution levels in Quebec (pp. 19, 23). Water pollution levels were determined by biological oxygen demand (BOD) and total suspended solids (TSS) levels (p. 21). Their findings showed that both inspections and threat of inspections had an impact on reducing emissions and inducing more frequent self-reporting (p. 35).[2]

Air Pollution and Hazardous Waste

Gray and Deily (1996) studied the compliance rates of steel plants with air pollution permits using compliance and enforcement data from EPA databases for 1980-1989 (p. 105). The authors defined compliance as a violation in any of the four quarters in a particular year (p.

[2] Self-reporting involved submitting reports to the government regulatory body. Plants had to submit a monthly discharge report to the Quebec Ministry of Environment. (Laplante & Rilstone, 1996, p.21)

98). One of their key findings was that at the plant level, greater enforcement, defined as inspections, letters, phone calls and enforcement orders, led to greater compliance (pp. 96, 110).

A study which focused on the pulp and paper industry by Nadeau (1997) looked at the effectiveness of monitoring and enforcement activities in reducing the length of time pulp and paper plants spent in violation of air quality regulations. Monitoring was defined as inspections and tests, while enforcement was defined as administrative orders, legal actions, and penalties (p. 57). Nadeau used quarterly EPA data, from EPA's Compliance Data System, which spanned from 1979-1989 (p. 67). His sample consisted of 175 companies. Nadeau found that a "10% increase in EPA monitoring activity leads to a 0.6-4.2% reduction in violation time" (p. 54). Additionally, he noted that "the same increase in enforcement activity results in a 4-4.7% reduction in violation time" (p. 54). Similar to Nadeau, Gray and Shadbegian (2005) found that greater enforcement with air pollution regulations led to increased compliance by pulp and paper mills (p. 238). Enforcement included inspections and non-inspection actions such as notices of violation, penalties and phone calls (p. 248). Gray and Shadbegian, used data from EPA's Compliance Data System and confidential census data to study 116 pulp and paper mills from 1979-1990 (p. 240).

A later study by Gray and Shadbegian (2007) that used EPA and census data from the 1990's for 521 manufacturing plants located 50 miles from St. Louis, Cincinnati, and Charlotte, North Carolina found that inspections and regulatory actions had a positive, but non-significant impact on facility compliance with air pollution regulations (p. 73).

Looking at hazardous waste, Stafford (2002), found that violations of RCRA decreased when the penalties were increased in 1991 (p. 290). Stafford used data from EPA's Resource Conservation and Recovery Act (RCRA) Information System and EPA's Biennial Reporting

System. She notes that this was consistent with Becker's work, but that the decrease was smaller than expected compared to the increase in penalties (p. 307).

Implications

These studies provide some evidence that firms do respond positively to government monitoring and enforcement as envisioned by Becker's model. Along with government monitoring and enforcement, Becker's model suggests that firm's compliance costs are also an integral factor in determining violation rates (Becker, 1968). This implies that given the same level of government enforcement and monitoring, facilities with higher compliance costs will have lower compliance rates. The next section reviews empirical studies that look at compliance costs and violations rates in conjunction with several theories of firm behavior in order to try to determine which types of plants may have higher compliance costs.

Agency Costs and Firm Environmental Compliance Rates

Cohen (1999) points out that most of the literature in environmental economics assumes that decisions made by management are implemented as directed (pp. 70-76). In reality, corporate environments are complex and managers face challenges when trying to convince employees to act according to company policies (Cohen, 1999, sec. 3.10). As a result, firms often incur "agency costs" to monitor employees. This section reviews "agency theory" and explores whether certain types of firms/facilities face higher agency costs than others. Large firms/facilities, firms with distant or multiple facilities, and facilities that are part of publicly owned companies are highlighted.

Agency Theory

According to agency theory, costs stem from the relationship between a principal (firm owner/manager) that delegates work and an agent (employee), to which the work is delegated. A problem can occur if the principal and agent have different interests. For example, an agent earning a fixed wage may not have the incentive to work as hard as possible to increase the firm's profit, whereas the goal of the principal is to maximize the firm's profit. Since employees have private information about their level of effort, it may be difficult for principals to verify the actions of agents. Thus, there is an "information asymmetry," whereby the agent has more knowledge of their level of effort than the principal. If the principal wants to obtain more information about the agent's performance, the principal needs to monitor the agent's performance to prevent, for example, shirking. The time and money the principal incurs monitoring the agent's performance (as well as any residual loss from not being able to monitor performance) is considered an "agency cost" (Cohen, 1999, sec. 3.10).

Agency costs may, therefore, be relevant to firm environmental compliance. For example, employees may be instructed to comply with environmental laws as a matter of company policy, but may also be provided bonuses to meet production goals. Under such an incentive structure, it would not be surprising to find that employees whose bonuses are based on the profitability of the firm, for example, would choose to invest the bulk of their resources to ensure that production targets are met and not enough to ensure environmental compliance. In this example, agents are shirking their responsibility to comply with the law in order to earn the bonus (Cohen, 1999, Sec. 310). Even absent bonuses of this nature, employees may shirk their responsibility of performing adequate monitoring and maintenance of pollution control equipment to prevent a non-compliance event (Cohen, 1999). It is believed that the principal-agent problem is more likely to be associated with certain firm and facility characteristics such as firm and facility size, discussed next.

Agency Theory and Firm/Facility Size

When agency theory is applied to firm and facility size, it is suggested that the larger the firm (measured by the number of employees), the more costs the firm incurs to monitor employee performance. Thus, there are higher costs required to comply with environmental regulations. As noted earlier, the Becker model as applied to firms predicts that the higher a firm's compliance costs and the lower the expected penalty, the less incentive firms have to comply with government regulation (Cohen 1986, p. 167). The implication is that given the same level of government monitoring and enforcement, larger firms and/or facilities (in terms of number of employees) would be less likely to meet environmental compliance standards than smaller firms and facilities. This same reasoning could also be applied to multi-plant firms. It may be more difficult for top management to oversee operations at multiple facilities; therefore, it could be the case that multi-plant operations would have higher agency costs and lower compliance rates compared to single plant firms.

There is some empirical data that provides support for the outcomes suggested by agency theory with regards to facility size. Several studies, across different disciplines, found that facility size had an effect on facility environmental performance. Gray and Deily (1996), studying steel plants from 1980-89, found that larger facilities were less likely to be in compliance (p. 106).[3] Similarly, Gray and Shadbegian (2005), studying pulp and paper mills over the period of 1979-1990, found that larger plants were less likely to be in compliance with

[3] Facility size was defined by manufacturing capacity (p. 99), which could act as a proxy for employee size.

air pollution regulations (p. 240).[4] Likewise, Gray and Shadbegian (2007) that looked at 1997 air compliance data for 521 manufacturing plants located 50 miles from St. Louis, Cincinnati, and Charlotte, North Carolina found that larger plants had lower compliance rates (p. 82).[5] Outside the area of regulatory compliance, Grant, Bergesen and Jones (2002), who analyzed TRI emissions of chemical facilities, found that 'large chemical plants emit toxins at significantly higher rates than do small plants' (p. 389).[6] Grant, Bergesen and Jones used number of employees as a proxy for plant size (pp. 396, 389). They defined the rate of emissions as "log ratio of annual pounds of chemicals released on site to annual pounds of chemicals used and stored on site" (p. 396).

Several studies also found that firms with multiple facilities had worse environmental performance than single facility firms. Gray and Deily (1996) studied forty-one steel making plants and found that firms with multiple steel plants were less likely to be in compliance than single plant firms (p. 389). Grant, Bergesen and Jones (2002), looked at TRI emissions of chemical facilities and found that 'large chemical plants emit toxins at a significantly higher rates ….especially if they are embedded in a wider corporate structure" (p. 389). Similarly, King and Shaver (2001), also using TRI emissions data, found that chemical and petroleum firms generate more waste if 'they operate multiple facilities across multiple jurisdictions in the United States' (p. 1069). To determine waste generation rates, King and Shaver created a relative measure that "compared waste generation rates of the facility to other facilities in the industry" (p. 1076).

---

[4] Facility size was defined as the "log of the plants real value of shipments" which could act as a proxy for employee size (Gray & Shadbegian, 2005, p. 251).
[5] Plant size was defined a "log of real shipments in 1997" (Gray & Shadbegian, 2007, p. 73).
[6] TRI refers to the Toxic Release Inventory. The Emergency Planning and Community Right-to-Know Act (EPCRA) require certain industrial facilities to report releases of particular toxic chemicals to the EPA. The data is available to the public through EPA's Toxic Release Inventory Program. It should be noted that TRI emissions are self- reported emissions; facilities are only required to provide estimated releases. Additionally, the toxic chemicals required to be reported are not necessarily "permitted" emissions, which would mean that stationary sources do not have to meet a prescribed emissions limit for the chemical. See United States Code, Title 42, Public Health and Welfare, Chapter 133, Pollution Prevention for more information and http://www.epa.gov/TRI/index.htm

On the other hand, there is some empirical evidence that challenges the outcomes suggested by agency theory. Gray and Shadbegian's 2005 paper, that looked at 116 pulp and paper mills over the period of 1979-1990, found that firm level characteristics such as firm size (log of firm employment) and firm ownership of multiple paper mills had no significant effect on compliance with air pollution regulations (pp. 247, 250, 258). A later study by Gray and Shadbegian (2007) found that larger facilities had lower air pollutant emissions (pp. 75-76). [7] Air pollutant emissions were defined as TRI emissions and the conventional pollutants $PM_{2.5}$ and SO_2 (p. 74). The 2007 study also noted that single plant firms had lower compliance levels with air pollution regulations (p. 82). Additionally, Grant, Bergesen and Jones (2002) found that TRI log emission rates were "not significantly higher when plants belong to a large firm" (p. 400).[8]

These contradictory outcomes may be explained by alternative theories. For example, an economies of scale argument would propose that larger companies should be better environmental performers, since there are fixed costs associated with learning the applicable regulations and implementing compliance programs. These costs could be spread across large firm operations and shared across multiple facilities. Additionally, one could expect that multi-plant firms may be more willing to comply with environmental regulations to create a positive externality, i.e. a "reputation benefit" for the other facilities (Gray & Deily, 1996, p.108). This desire would produce an opposite result than the one anticipated by agency theory.

There is some empirical evidence that lends support for agency theory, as discussed above. Conflicting empirical results, however, suggest that alternative explanations such as economies of scale and 'reputation effects' may outweigh any agency costs associated with large

[7] Gray and Shadbegian (2007) note that "since emissions are calculated relative to plant size, and only plants with relatively large emissions are included in the EPA database" the results for plant size "can hardly be treated as evidence of economies of scale in controlling emissions" (pp. 75-76).
[8] Firm size was determined by taking the log of the number of employees (p. 396).

firms/facilities and multi-facility firms. Given the mixed empirical results, it is difficult to determine which effect dominates. For that reason, additional empirical analysis that tests agency theory and firm and facility size would be beneficial.

Agency Theory and Distant Headquarters

In addition to firm/facility size, the principal-agency problem could also extend to the distance that a facility is from firm headquarters. It could be argued that the more distant a facility is from their parent headquarters, the costlier it is for owner/managers to monitor the plant's activities. This added cost could mean that facilities that are farther away from parent headquarters would have higher violations rates than those close to the headquarters.

There is only very limited research that looks at the effect of distant headquarters on facility behavior. A sociology study by Grant, Jones and Trautner (2004) looked at TRI emission rates[9] and found that absentee managed plants (i.e. plants with out-of-state headquarters) emit more toxins on average; however, when they accounted for the amount of chemicals they had on-site, along with other factors, they found that plants with out-of-state headquarters did not perform any differently than other plants (p. 203). Similar studies looking at the effect of distance headquarters on plant's compliance with permit requirements could not be found. Given the lack of empirical data on the effect of having distant headquarters on firm environmental compliance, empirical testing would be beneficial to determine if distance from headquarters impacts facility environmental compliance.

[9] The Emergency Planning and Community Right-to-Know Act (EPCRA) require certain industrial facilities to report releases of particular toxic chemicals to the EPA. The data is available to the public through EPA's Toxic Release Inventory Program. It should be noted that TRI emissions are self- reported emissions; facilities are only required to provide estimated releases. Additionally, the toxic chemicals required to be reported are not necessarily permitted emissions, which would mean that stationary sources do not have to meet a prescribed emissions limit for the chemical. See United States Code, Title 42, Public Health and Welfare, Chapter 133, Pollution Prevention for more information and http://www.epa.gov/TRI/index.htm

Agency Theory and Corporate Ownership

Agency theory may also be applicable to the number of owners a company has. Publicly traded firms can have many shareholder/owners. The more shareholder/owners a firm has the more difficult it could be for them to monitor activities of the agents – including managers of the firm (Alexander & Cohen 1996, p. 422; Alexander & Cohen, 1999). In contrast, private firms tend to have a few long term owners, so theoretically those owners are more likely to monitor the activities of employees. In many private companies this is especially true since managers are often owners.

It may also be harder to align incentives of shareholder/owners with managers of public companies that are under short term pressures to meet quarterly earnings targets.[10] Those managers may be less likely to have the incentive to make long term investments in areas like pollution prevention. Private firms, on the other hand, traditionally have more long term investors that may be interested in making the kind long term investment needed to be in compliance.

Empirical testing would be beneficial to determine whether parent company ownership affects compliance at the facility level. While obtaining information on employee incentives is outside the scope of this study, it is still possible to test for agency costs.

In summation, an understanding of agency theory could be advanced by gaining insight into the effects of parent company ownership on compliance, as well as the effect of firm/facility size, multi-facility ownership and facility distance from headquarters on violation rates. Additionally, knowledge about firm environmental compliance rates could be enhanced by understanding which firm's face higher costs to develop the knowledge and expertise to comply with local regulations, as discussed in the next section.

[10] See Cohen (1999) for a discussion on aligning top management incentives with shareholders interests.

Costs to Develop Compliance Expertise

In addition to agency theory, another idea discussed in the literature that could affect an individual firm's ability to comply is whether the regulated parties have the capacity to "behave as intended" (Burby & Patterson, 1993, p. 755). In order for a firm to comply, firms need to sufficiently understand the complex environmental regulations that are applicable to them and have the technical expertise to meet the relevant requirements. Some facilities may face higher costs developing this expertise than others. Assuming firms are profit maximizers, Becker's optimal penalty model suggests that firms and/or facilities that have higher costs to develop this expertise would have less incentive to comply with regulations than firms and/or facilities who do not incur these costs, given the same level of government monitoring and enforcement Becker, 1968).

King & Shaver (2001) suggest that "foreign-owned firms have less expertise and information useful to meeting local requirements" (p. 1072). In the area of waste prevention, King & Shaver note that it requires 'highly tacit, conditional and local skills.' Furthermore, many studies they cite suggest that pollution prevention practices are not easily diffused across a firm with multiple locations because of differences in local conditions; therefore, they hypothesize that "foreign-owned firms would have less expertise and information useful to meeting local environmental requirements," thus causing them to generate more waste (p. 1072).

In the empirical part of their study, King and Shaver (2001) looked at whether there was a difference in waste generated by foreign vs. domestically owned firms. Foreign ownership was defined in terms of the ultimate global parent being from a nation other than the United States (p. 8). To determine waste generation by plants, they used information from the TRI database. TRI emissions are self reported emissions.[11] They limited the study to two industries: chemical (SIC 28) and petroleum (SIC 29). The study found that foreign-owned firms tend to generate and manage more hazardous waste than domestic firms in the same industry, due to being "unfamiliar with the business, cultural and legal environment" (pp. 1069, 1082, 1083). Additionally, they noted that "foreign firms expend considerable effort to compensate for this disadvantage by ameliorating more of the waste they generate" (pp. 1083).

King and Shaver's hypothesis may also be applicable to facility compliance with environmental regulations related to permitted emissions limits and other point source requirements. For example, foreign firms have been known to test pollution control equipment abroad, which does not always work as well in local conditions. Additionally, foreign-owned facilities may have less experience understanding and complying with domestic regulations and the domestic regulatory structure; they may, for example, be more accustomed to a more cooperative approach in their home country. This could imply that foreign firms, compared to domestic firms, may face higher costs developing the expertise to understand and comply with complex U.S. regulations. Assuming firms are profit maximizers, this could mean that foreign-owned firms would have more violations than domestically owned firms. Empirically testing this hypothesis would be beneficial to see if a similar result is found when testing facilities compliance with regulatory requirements.

[11] The Emergency Planning and Community Right-to-Know Act (EPCRA) require certain industrial facilities to report releases of particular toxic chemicals to the EPA. The data is available to the public through EPA's Toxic Release Inventory Program. It should be noted that TRI emissions are self- reported emissions; facilities are only required to provide estimated releases. Additionally, the toxic chemicals required to be reported are not necessarily "permitted" emissions, which would mean that stationary sources do not have to meet a prescribed emissions limit for the chemical. See United States Code, Title 42, Public Health and Welfare, Chapter 133, Pollution Prevention for more information and http://www.epa.gov/TRI/index.htm

Research Questions

As discussed above, agency theory suggests that certain observable firm and facility characteristics will result in higher violation rates. It is thought that larger firms and facilities, facilities that are part of firms that own multiple facilities and/or are distant from headquarters will have higher agency costs, because more resources are needed to monitor employee performance. It is also thought that facilities that are part of foreign companies may incur more costs complying with local rules than domestically owned facilities. Agency costs and the costs to develop expertise to meet local requirements could increase a facilities overall compliance costs. The Becker model tells us that given the same level of monitoring and enforcement, facilities with higher compliance costs may have lower compliance rates. It is within this context that the following hypotheses were tested:

Hypothesis #1

Large facilities will have more violations than small facilities.[12]

Hypothesis #2

Facilities that are part of larger firms will have more violations than facilities that are part of smaller firms.

Hypothesis #3

Facilities that are part of multi-facility firms will have more violations than single facility firms.

Hypothesis #4

Facilities whose company headquarters are far away will have more violations than facilities that are located closer to their company headquarters.

[12] Economies of scale suggests an opposite effect.

Hypothesis #5

Facilities that are part of publicly traded firms will have more violations than facilities that are part of private firms.

Hypothesis #6

Facilities that are part of foreign-owned firms will have more violations than facilities that are part of domestic firms.

CHAPTER II

RESEARCH METHODS

This chapter discusses the research methods used to test the hypotheses developed in Chapter I. The first part discusses the scope of the sample and is followed by an overview of the Clean Air Act and its application to manufacturing facilities. The third section elaborates on the procedures used to collect data to test the proposed hypotheses. The last two parts discuss variable construction and model selection.

Scope of Sample

Originally, the study design intended to include all facilities nationwide with Clean Air Act (CAA), Resource Conservation and Recovery Act (RCRA covers the handling and disposal of hazardous waste), and Clean Water Act (CWA) permits. However, due to data quality issues, the scope had to be substantially narrowed. Ultimately, due to the difficulty in obtaining accurate data, the sample was restricted to a small subset of chemical manufacturing facilities with major air permits, located in EPA Region 4.[13]

The study was limited to one industry, as was done with other economic studies in the area of environmental enforcement, due to the challenges associated with comparing dissimilar industries (See Nadeau, 1997; Laplante, 1996; Magat & Viscusi, 1990; Laplante & Rilstone, 1996; Gray & Deily, 1996). The chemical and allied product industries were chosen (SIC 28) because chemical facilities are highly regulated, and often have permits that include detailed compliance requirements. Additionally, this industry has a high level of foreign direct

[13] Any reference to EPA in this paper refers to the United States Environmental Protection Agency.

investment; thus, selecting this industry made it possible to adequately test all the proposed hypotheses (King & Shaver, 2001, p. 1074).

The study was further narrowed to include only facilities with a major operating air permit for several reasons. First, there were substantially less data available at both the state and federal level on minor permit holders, because the EPA does not generally require that states submit inspection or violation data on minor facilities (EPA, 2005b, ECHO. Detailed Facility Report: Data Dictionary, sec. Facility Characteristics, ¶ Facility Universe, CAA; EPA, 2005a, ECHO. About the Data. Data Completeness). This made it challenging to get reliable data on minor facilities. Secondly, at the time data for this study were collected (2005) it was difficult to identify current RCRA permit holders in EPA's database, given that they maintain a list of anyone who ever had a RCRA permit. [14] This meant that determining whether a facility was active for the period of the study would require going through individual company records at each state office.[15] This was not feasible given the scope of project. Lastly, because the study was limited to only include chemical facilities in EPA Region 4, there proved to be too few chemical companies that were CWA major permit holders to allow for a meaningful analysis of

[14] This was done so that the EPA could identify, if needed in the future, who was responsible for CERCLA violations at a particular location. According to EPA'S ECHO website data quality section: "Under the RCRA statute, no regulatory requirements mandate that sites that once handled hazardous waste notify that they have ceased waste management activities. As a result, the RCRA Info database (which feeds ECHO) contains a listing of all regulated sites that at one point managed hazardous waste. The database includes both active sites and those that are no longer managing hazardous waste and/or are permanently closed. Including all sites assists EPA and the public in determining prior uses of land. EPA and the states recently completed a workgroup process to develop a method for "inactivating" sites in RCRA Info" (EPA, 2008, ECHO, Known Data Problems, Sec. Hazardous Waste, General)

[15] Stafford (2002) cross checked RCRA permits with EPA' Biennial Reporting System. Facilities (large quantity generators are required to report) that reported data as required for the biennial report were assumed to be active. This may be a helpful approach to narrow down samples for future research. However, these reports are usually only available two years after the information is submitted to the EPA, making it less useful when looking at current data. Also, according to the EPA, "the 2001 National Biennial Report includes management and receipts data from both permitted treatment, storage and disposal facilities and generators that are not required to be permitted (e.g., those that recycle solvent hazardous waste generated on-site)" (EPA, September 2008, 2001 National Hazardous Waste Biennial Report, sec. The Data Presented in the 2001 National Biennial Report).

the proposed hypotheses on CWA violations (88 total). For these reasons, the hypotheses were only tested on one media- air permit violations.

The study only included facilities from several states in EPA Region 4 because of the difficulty in getting accurate data from all regions and states across the United States (see U.S. GAO, June 2006, p. 8 for a discussion about EPA data). Both EPA and state data were needed to create the dataset. The main EPA database used was the EPA Air Facility System (AFS) database, which is composed of information required to be reported by the states to EPA. The database contains permit information, as well as some inspection and enforcement figures on all facilities with major permits. At the time the study was formulated, the data quality information was reviewed on EPA's ECHO website (EPA, 2005c, ECHO, Known Data Problems, sec. Air).[16] After reviewing this information, EPA Region 4 was chosen because the EPA reported that the Region had no known data quality issues (EPA, 2005c, ECHO, Known Data Problems, sec. Air). Additionally, Region 4 included hundreds of permitted chemical and allied product facilities.

Once EPA Region 4 was selected, a request for records was sent to the Region 4 office. The office provided electronic files that included some but not all of the required data. The files included the following information about major air permit holders in the region: high priority violations (limited details were available), SIC category, and demographic information of areas surrounding the facilities. To get more detailed inspection and violation information a records request went out to state offices; however, only five out of the eight states in EPA Region 4 were able to respond to the request included: Tennessee, Georgia, North Carolina, South Carolina, and

[16] The fact there were no known data quality issues, did not mean there were not any data quality issues.

Florida. For the most part, the air offices[17] in Tennessee, Georgia, North Carolina and South Carolina all required that someone come to their offices to sift through their files and make photocopies of relevant documents. Florida air offices, unlike the others, were able to send electronic copies of all their information.

Several states did not respond or had difficultly responding to the records request; therefore, not all major air permit holders in EPA Region 4 with an SIC 28 were included in the sample. Kentucky air offices had trouble responding, in particular the Louisville office, because they were behind on enforcement actions and could not make public much of the enforcement information requested (D. Spillman, personnel communication, April 18, 2005).[18] This meant sixty-two cases from Kentucky were not included in the sample. Alabama's Department of Environmental Management information officer asked several times that the request be reconsidered since they had very limited staff available to assist with providing the needed files. According to the information officer, the information was only available in paper files and would require that a staff person be present while the files were reviewed (S. Demick, personnel communication, May 2005). It was estimated that it would take two weeks onsite to collect the data for around sixty cases. Given budget limitations and the difficulty in gaining cooperation of the state office in Alabama, these cases were not included in the sample.[19] The Mississippi Department of Environmental Quality never responded to the records request. Since there were only seventeen facilities located in Mississippi to be included in the sample, the request was not pursued. In short, the study included facilities from five of the eight states that make up EPA

[17] The term "air office" in this study means a state or local agency designated with the responsibility to issue permits and enforce air pollution regulations.
[18] Per personnel communication, email from Doug Spillman of the Louisville, Kentucky Metro Air Pollution Control Office on April 18, 2005 "The Title V permitting info is on our web site, www.apcd.org under permitting. You can get those issue dates there. We cannot send any violation information you requested as the agency has a big back-log of enforcement actions and all of those for the time period you requested have not been sent out yet."
[19] Personnel communications, email and phone conversations with Scott Demick at the Alabama Department of Environmental Management. May, 2005.

Region 4. This means that roughly 60% of chemical facilities with major air permits in Region 4 were included in the sample.[20] This equates to roughly 15% of chemical facilities with major air permits nationwide.[21]

With the exception of Florida, the offices that responded to the records request were unsure about the accuracy of their enforcement records before fiscal year 2001, which started October 1, 2000; therefore, the dataset only looks at violations between fiscal year 2001 and 2004 (B. Gatano, K Frost, personnel communication, May 2005).[22]

In summary, an original dataset that spans from Oct. 1, 2001 through Oct, 1, 2004 was created from public and private sources, in order to test the proposed hypotheses. The final sample consists of all facilities that had each of the following characteristics:

1) a Title V or synthetic minor air permit from October 2000- October 2004 (EPA fiscal year 2001- 2004). Minor sources were excluded. [23]

2) an SIC code of 28,

3) located in Tennessee, Georgia, North Carolina, South Carolina or Florida.

4) not part of a company that declared bankruptcy.

---

[20] It is difficult to get an exact percentage, given that it is hard to know how many of the facilities in the dataset are miscoded (e.g. wrong SIC code on permit or wrong type of permit listed in the system)

[21] This rough estimate is based on the number of chemical facilities (SIC 28) with a major air permit in EPA's ECHO database website retrieved May, 16, 2008 from http://www.epa-echo.gov/echo/compliance_report_air.html

[22] This is based on discussions with air enforcement staff at state offices in May 2005 including Betty Gatano at the North Carolina Department of Environment and Natural Resources Division of Air Quality, and Keith Frost the Manager of the Air Enforcement section at the Bureau of Air Quality in South Carolina DHEC.

[23] A minor source has the potential to emit emissions below the major source thresholds. Minor source requirements can vary by state. It is difficult to know whether the inclusion of these sources would alter the results. According to a search run on April 16, 2009 in EPA's ECHO database found at http://www.epa-echo.gov/echo/compliance_report_air.html if minor source air permits were included in the sample, it would result in 928 cases in EPA Region 4, and 732 in the five states included in this sample. The estimate may be low. The EPA, ECHO website section "Limitations for 'Minor' Records (Air)" that pops up when *Other Minor* is chosen as a search parameter says, "Clean Air Act (CAA) Minor facilities are not required to be entered into federal data systems. Some states enter data for minors, others enter partial or very limited amounts of data. Comparisons across states relating to these facilities are not valid, and data may be less reliable." Retrieved April 16, 2009 from http://www.epa-echo.gov/echo/compliance_report_air.html

5) publicly or privately owned; not part of a state owned company, non-profit, co-op or joint venture.

The one mega site, Eastman, Tennessee, proved to be an extreme outlier with over twenty inspections. It was taken out of the sample, since as a mega-site (i.e. a complex site with many units), it is subject to other regulatory requirements (EPA, April 2001, p. 9). To help better understand Title V requirements, the next section provides an overview of the regulatory program.

Overview of Clean Air Act Stationary Source Permitting and Enforcement

The 1990 Clean Air Act amendment modified existing programs to help establish the Title V Operating Permit program that is in effect today. [24] The Title V Operating Permit program is applicable to all "major" stationary sources and is designed to ensure that these sources comply with the CAA requirements (Clean Air Act, 1990, 42 U.S.C. §7661). The operating permit a “major” facility receives outlines details on emissions limits, monitoring, maintenance, record keeping, and reporting requirements that it has to follow. For the most part, Title V permitting and enforcement has been delegated to the states, with EPA oversight.

A Title V Operating Permit is required for facilities that have operations that have the “potential to emit” (“PTE”) a given threshold of major air contaminants as defined by the CAA. The threshold that subjects a facility to “major” status varies, depending on the attainment status of the air quality control region in which it is located. Nonattainment for an air quality control

[24] Operating permits grant stationary sources permission to operate. This is different than preconstruction permits. New Source Review and PSD are pre-construction permits that are not addressed here. Additionally, at the time this study was conducted, minor sources subject to New Source Performance Standards (NSPS) and National Emission Standards for Hazardous Air Pollutants (NESHAP) were deferred from the requirement to apply for a Title V Operating Permit.

region means the region did not meet the national ambient air quality standard for criteria pollutants set by the EPA (Clean Air Act, 1990, 42 U.S.C. § 7602). Criteria pollutants are air pollutants that are regulated to protect public welfare. They include carbon monoxide, particulates, sulfur dioxide, nitrogen dioxide, ozone and lead (Protection of Environment, 40 C.F.R. pt. 50). An air quality control region can be in attainment for one criteria pollutant but in nonattainment for another. Some examples of thresholds are as follows:

- In attainment areas: Facilities with the potential to emit (PTE)10 tons per year (tpy) of any hazardous air pollutant (HAP) or 25 tpy of any combination of HAPs or 100 tpy of any regulated pollutant. Hazardous air pollutants (a.k.a. air toxics) are pollutants that cause or may cause cancer or other serious health effects, such as reproductive effects or birth defects, or adverse environmental and ecological effects. The CAA lists over 188 hazardous air pollutants. (Clean Air Act, 1990, 42 U.S.C. §§ 7412-7513).
- In nonattainment areas: If an area is in nonattainment for a pollutant such as particulates, a major source would be defined as one with particulate emissions of 70 tpy or more. In ozone nonattainment areas, this level is reduced to 50 tpy (of ozone precursors VOCs or NOx) in serious areas, 25 tpy in severe areas, and 10 tpy in extreme ozone nonattainment areas. For serious carbon monoxide nonattainment areas, any source with a PTE of 50 tpy or more would need a permit (Clean Air Act, 1990, 42 U.S.C. §§. 7412 a(4), 7511 a(c-e), 7512, a(c)(1), 7513a(b)(3)).

A stationary source that has the PTE at the level of a Title V permit can opt for a synthetic minor permit (a.k.a. conditional major). To be eligible for a synthetic minor permit a facility has to agree to conditions, such as limit production or operation time (e.g. only operate a stipulated amount of hours a day), that would limit their PTE below “major” source thresholds outlined above. Facilities also have to keep records to demonstrate they are meeting these conditions.

Both Title V and synthetic minor permits are federally enforceable permits. The EPA requires that inspection and enforcement data (limited to high priority violations a.k.a. HPV data) be submitted on Title V and synthetic minor permits (Schaeffer, 1998). The EPA also sets out inspection goals for Title V and synthetic minor facilities. The current EPA compliance monitoring strategy states that a full compliance evaluation should be done once every two years for a Title V permitted facility, once every five years for a synthetic minor facility,[25] and once every three years for mega facilities which are large complex facilities (EPA, April 2001, p. 7). Minor sources, which fall below the PTE for major sources, may still require state operating permits.

The U.S. EPA has defined which types of violation should be reported as High Priority Violations by the states (Schaeffer, 1998, Sec. II, pp. 5-7). These designations are not necessarily what states view as high priority; states will often focus on the impact to human health, duration of violations and recalcitrant nature of the violator.[26] HPV designations are made according to the December 22, 1998 memo: Issuance of Policy on Timely and Appropriate Enforcement Response to High Priority Violations (Schaeffer, 1998). The following criteria can trigger HPV status:

- Failure to obtain a Prevention of Significant Deterioration (PSD) permit
- Violation of an air toxics requirement
- Violation by a synthetic minor of an emission limit that affects the source's regulatory status

[25] This applies to synthetic minor sources that emit or have the PTE to emit 80% of the Title V major source threshold.
[26] This is based on conversations with state regulators in Georgia, Tennessee and North Carolina in May 2005.

- Violation of an administrative or judicial order
- Substantial violations of a sources Title V obligations
- Failure to submit a Title V permit application within 60 days of the deadline
- Testing, monitoring, record keeping or reporting violations that substantially interfere with enforcement or determination of a facility's compliance requirements
- Violation of an allowable emission limit detected during a source test
- Chronic or recalcitrant violations, or
- Substantial violations of 112 (r) requirements. 112 (r) refers to emergency management of certain chemicals that are stored at a facility. Not all facilities use chemicals subject to 112 (r) requirements (Schaeffer, 1998, Sec. II, pp 6-7).

Data Collection and Data Quality Challenges

To develop a complete dataset, information from a variety of databases had to be gathered. See Figure 1 for a list of information sources.

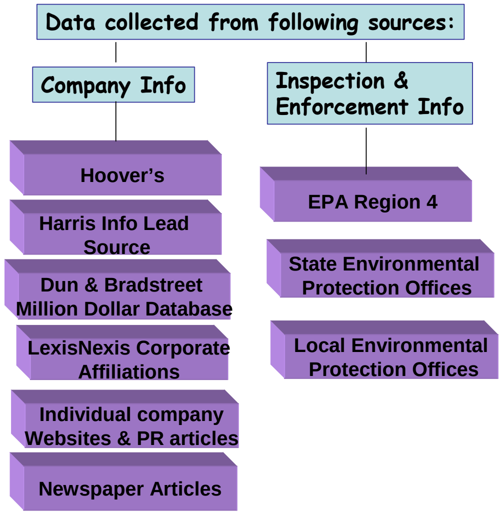

Figure 1: Data Sources

Inspections and Violation Information

EPA data alone were not sufficient for this study because they had several limitations. First of all, EPA only requires that "High Priority Violations" (HPV) be reported, rather than all types of violations (Schaeffer, 1998). Additionally, the EPA AFS database had limited or no information on the type of violation (e.g. reporting, operations, emissions, etc) and appeared to be missing inspection information.[27] Lastly, for some states like South Carolina, it was difficult to discern from EPA data whether a facility was a synthetic minor facility or Title V facility since they were all coded as major facilities. As a result of the limitations with EPA data, state and local program data were considered necessary.[28] The decision to collect local data was made to ensure more accurate results.[29]

The list of chemical facilities with major air permits provided by EPA Region 4 was used to determine which facilities should be included in the sample. Based on this list, a records request was sent to state and local air offices. Inspection, enforcement and permit data were collected from state and local air offices. Data collection mostly involved reviewing and photocopying information from state files.

Some information was available electronically from state offices. Florida Department of Environmental Protection emailed inspection and enforcement information in spreadsheet formats. Permits were downloaded from Florida and Georgia's state websites.[30] When permits were not available from the state websites, local offices in Florida sent missing permits

---

[27] This was confirmed by Errol Reksten at the Chattanooga Air Office, per emails in April 2005. He explained that the EPA Instructions were not always clear as to what needed to be entered in the EPA database.
[28] Other researchers have noted problems with EPA data. Magat and Viscusi (1990) noted that EPA enforcement data was incomplete and therefore they used proxy measures (p. 334).
[29] While most studies in the air pollution compliance area use EPA data, it is data from the 1990's, before the HPV reporting policy went into effect. It was felt that using only EPA data post 2000 may provide misleading results.
[30] Florida DEP maintains a page for Air Permit Document searches at http://www.dep.state.fl.us/air/eproducts/apds/default.asp , and Georgia Department of Natural Resources maintains one at http://airpermit.dnr.state.ga.us/gaairpermits/

electronically, and copies of Georgia permits were obtained from permit files. The South Carolina Air Quality office (via the Freedom of Information Center) and some of the other state and local offices were also able to send some copies of permit data electronically.

This data collection process allowed for “data cleansing” of the original EPA Region 4 list. For example, several facilities that were listed as having major permits were in fact minor facilities; their permit status had not been updated in AFS database. Additionally, several facilities listed as having an SIC code of 28 were not chemical or allied product facilities. For example, a grain facility and gas station were miscoded as SIC 28. It is estimated that roughly 10% percent of facilities were excluded for these reasons. (See Appendix G for a list of state and local air offices that provided data for the study).

Company Information

To obtain firm level data, several databases and public information sources were used. Company data were collected from Hoovers, Lexis Nexus Directory of Corporate Affiliations, Dun and Bradstreet Million Dollar Database, Harris Lead Express, company websites and news articles.[31] Hoovers was used extensively to determine the ultimate parent company information and information on locations and the number of employees worldwide. Furthermore, Hoover’s company history information was used to find out about mergers and acquisitions; additional merger and acquisition information was found in articles or company press releases. To supplement data from Hoovers, Lexis Nexus Directory of Corporate Affiliations and the Dun and Bradstreet Million Dollar Database was utilized to determine U.S. Headquarters and Ultimate

[31] All data was collected in late 2004 through 2005.

Parent information.[32] For small to mid sized company information, data were mainly obtained from the Dun and Bradstreet Million Dollar Database and Harris Lead Express (now called Selectory). In some cases, especially when looking for information on small to mid size companies, company websites needed to be used to help determine or verify ownership.

There were some limitations to using The Dun and Bradstreet Million Dollar Database for collecting firm level data. The database proved to not always be a good source of information on single and multi-facility firms as defined by this study. For the purposes of this study a single facility means a company with one location. The Dun and Bradstreet Million Dollar database used the following definitions:

> Branch - a secondary location of a company that has no locations reporting to it. (D&B, n.d., sec. Single Location Subsidiary)
>
> Single Location - either a subsidiary that has no locations reporting to it, or a company with only one location. (D&B, n.d., sec. Branch)

Additionally, the Dun and Bradstreet Million Dollar Database did not always characterize facility locations as public and private the same way as was done in this study. For purposes of this study, whether a facility was determined to be part of a public or private company was based on whether the ultimate global parent was traded on an exchange. Dun and Bradstreet's Million Dollar Database would sometimes list a facility/location as private, when it had a parent that was on a public exchange. For example, the Bridgestone Firestone North American Tires LLC (not included in the sample, but used for illustrative purposes only) location in Nashville is listed as private, although it is part of Bridgestone Japan which is on a public exchange (D&B, Company

[32] The Directory of Corporate Affiliations "covers major U.S. and international public and private businesses—more than 180,000 companies" (Corporate Affiliations, para. 2, Retrieved May 12, 2009 from http://www.lexisnexis.com/dca/)

Information for Bridgestone Americas Inc).[33] Using Dun & Bradstreet's database, it would be necessary to look up the ultimate global parent of each facility/location to determine public or private status. Subscription limitations to the database prevented access to information on foreign parent companies to determine how they were listed.[34] As a result, company websites needed to be checked to verify public or private ownership status of companies, as defined by this study.

Facility level data, such as number of employees on site, were primarily obtained from the Dun and Bradstreet Million Dollar Database and Harris Lead Express (now called Selectory). For small companies and facilities without a listing in any database, websites and local news articles were used to get data on the facility.

Distance information between the facility and headquarters/parent office were not available in the above mentioned databases, and were obtained using the internet site Mapquest or the internet tool Geobytes City Distance Tool.[35]

Variable Construction

Tables 1through 3 list all the variables that were created from the data collected.

Table 1: List of Control Variables Constructed for Study and Data Sources

Descriptions	Control Variable Name	Type	Source
* Included in final analysis[36]			
Facility Inspections	F_Inspec*	Continuous	State EPA Air offices files
Florida	FL*	0,1 Dummy variable	Facility Address
Georgia	GA*	0,1 Dummy variable 1= facility in GA	Facility Address
North Carolina	NC*	0,1 Dummy variable	Facility Address

[33] Information retrieved March 12, 2007 from Dun & Bradstreet Million Dollar Database via Walker Management Library, Owen School of Management,Vanderbilt University.
[34] Dun & Bradstreet's Million Dollar database was accessed through Vanderbilt University's in 2005-2006.
[35] For Mapquest see www.mapquest.com. For Geobytes City Distance Tool see http://www.geobytes.com/citydistancetool.htm
[36] See variable explanations and correlations section below for an explanation on how variables were chosen for the final analysis.

Descriptions	Control Variable Name	Type	Source
		1= facility in NC	
South Carolina	SC*	0,1 Dummy variable 1= facility in SC	Facility Address
Tennessee	TN*	0,1 Dummy variable 1= facility in TN	Facility Address
Local program office does permitting and enforcement (e.g. not a state air office)	Local	0,1 Dummy variable 1= local program does permitting and enforcement	Facility location
Log of the number of pages of a facilities' air permit	Pages_log	Continuous	Facility Permit
Type of permit, Title V of Synthetic minor (a.k.a. conditional major)	TV_perm*	0,1 Dummy variable 1= a Title V permit, 0 is a synthetic minor (a.k.a. conditional major) permit	Facility Permit
How many quarters a facility had a temporary permit	Temp_Bi	0,1 Dummy variable 1= facility had a temporary permit	State and local EPA files
Did the facility have a fine in the fiscal year prior to the start of the data set	Pre fine*	0,1 Dummy variable 1= facility had a fine in previous year	State and local EPA Air office files[37]
Industrial inorganic chemicals	SIC_281*	0,1 Dummy variable 1= facility permit indicates it is SIC 281 facility	Facility Permit*
Plastic materials and synthetics	SIC_282*	0,1 Dummy variable 1= facility permit indicates it is a SIC 282 facility	Facility Permit
Drugs	SIC_283*	0,1 Dummy variable 1= facility permit indicates it is a SIC 283 facility	Facility Permit
Soaps, cleaners and toilet goods	SIC_284*	0,1 Dummy variable 1= facility permit indicates it is a SIC 284 facility	Facility Permit
Paints and allied products	SIC_285*	0,1 Dummy variable 1= facility permit indicates it is a SIC 285 facility	Facility Permit
Industrial inorganic chemicals	SIC_286*	0,1 Dummy variable 1= facility permit indicates it is a SIC 286 facility	Facility Permit
Agricultural chemicals	SIC_287*	0,1 Dummy variable 1= facility permit indicates it is a SIC 287 facility	Facility Permit
Miscellaneous chemical products	SIC_289*	0,1 Dummy variable 1= facility permit indicates it is a SIC 289 facility	Facility Permit
SIC 28 is secondary	SIC_other	0,1 Dummy variable 1= facility permit indicates that the facility has a secondary SIC of 28	Facility Permit

[37] David Konisky now at the University of Missouri, Department of Political Science provided the 1st quarter, fiscal year 2000 data on fines from data he acquired from EPA.

Table 2: List of Theoretical Variables Constructed for Study and Data Sources

Description	Theoretical Variables	Type	Source
Facility is part of a company with more than 24,839 employees, but less than 315,082 employees. This is the top 3rd of the sample.	BIG*	0,1 Dummy variable 1= facility is part of co. with between 24,839 to 315,082 employees	Hoovers, Dun & Bradstreet, Harris
Facility is part of a company with more than 2,400 employees and less than 24,530 employees. This is the middle third of the sample.	MED*	0,1 Dummy variable 1= facility is part of co. with between 2,400 to 24,530 employees	Hoovers, Dun & Bradstreet, Harris
Facility is part of a company with more than 5 employees, but less than 2,310 employees. This is the bottom third of the sample.	SM*	0,1 Dummy variable 1= facility is part of co. with between 5 to 2,310 employees	Hoovers, Dun & Bradstreet, Harris
Log of the number of employees at the facility	LN_EMP*	Continuous	Hoovers, Dun & Bradstreet, Harris, company websites
Facility is part of a foreign-owned company.	FOREIGN*	0,1 Dummy variable 1= facility is part of a foreign-owned company	Hoovers, Dun & Bradstreet, Harris, Lexis Nexus Directory of Corporate Affiliations, company websites
Facility is not in the same state as company parent/U.S. head office	OUT_ST*	0,1 Dummy variable 1= facility has an out-of-state parent/head office	Hoovers, Dun & Bradstreet, Harris, Lexis Nexus Directory of Corporate Affiliations, company websites, Mapquest and Geobytes city distance tool
Is the facility part of a company on a public exchange (e.g. not part of a privately owned company)	PUBLIC*	0,1 Dummy variable 1= facility is part of a public company	Hoovers, Dun & Bradstreet, Harris, Lexis Nexus Directory of Corporate Affiliations, company websites
Is the facility a single location or part of a company with multiple locations	SINGLE*	0,1 Dummy variable 1= facility is a single location	Articles, company websites, Dun & Bradstreet, Lexis Nexus Directory of Corporate Affiliations,

Table 3: List of Exploratory and Other Variables Constructed for Study and Data Sources

Description	Exploratory Variable Name	Type	Source
Facility ownership changed- the facility was part of a company that went through a merger or acquisition during the time period of the dataset	MERGER	0,1 Dummy variable 1= facility is part of a company that went through a merger or acquisition	Articles, company websites, notes in EPA files, Hoovers
Description	Dependent Variable Name	Type	Source
Total number of violations over 16 calendar quarters	Total_viol*	Continuous	State and local EPA files and AFS Database
Severity of facility violations.	Severity	Ordinal (0 to 3 scale, 3 being worst)	State and local EPA files and AFS Database
Description	Other	Type	Source
Did the facility have a fine during the period of the study	Fine	0,1 Dummy variable 1= the facility had a fine during the period of the study	State and local EPA files and AFS Database
Amount of fine issued to a facility during the period of the study	Fine Amount	Continuous	State and local EPA files and AFS Database
Permit Violation	Viol_1	Continuous	State and local EPA enforcement files
Operations Violation	Viol_2	Continuous	State and local EPA enforcement files
Recordkeeping violation	Viol_3	Continuous	State and local EPA enforcement files
Emissions Violations	Viol_4	Continuous	State and local EPA enforcement files
Other violation	Viol_5	Continuous	State and local EPA enforcement files

Dependent Variables

Several dependent variables were created to test the proposed hypotheses. The dependent variables used in this study were created using both EPA and state violation data. [38]

[38] Phillip Barnett at EPA Region 4 provided electronic files of all state data submitted into the AFS system from all eight states in it's region and data on EPA enforcement actions. As of the third quarter of the fiscal year 1999, EPA only required that High Priority Violations (HPV) be reported. EPA policy, per memo dated 1998, specifically outlines what violations constitute an HPV violation. The policy focuses on Title V permit holders, but also applies to synthetic minor sources that violate a condition of their permit that makes them a synthetic minor source. At the state level enforcement actions are not limited to HPV violations; therefore, enforcement data on all types of violations were collected from state and local air program/enforcement offices.

Total Violations

The variable TOTAL_VIOL represents the total number of different violations discovered at a facility each quarter, added up over 16 quarters. It includes state and EPA reported enforcement actions. [39] To keep the violations as comparable as possible across the states, the violations included as part of this study consist of all recorded violations from Florida's database, and all violations that resulted in some sort of formal Notice of Violation (NOV) from the other four states. The was done because all the states included in the study, except for Florida, sent out some form of formal NOV for both major and minor violations.[40] Florida only issued an NOV if they planned to take formal enforcement action, i.e. legal action resulting in a consent order or penalty against a violator (FL DEP, 2003, Chapter 3: Enforcement Options, p. 16). Although the Florida enforcement officials did not issue NOV's for all types of violations, the state of Florida did keep electronic records of all violations, both minor and major violations.

Using Florida enforcement documents as guidance and input from state enforcement personnel at state air offices, facility violations were broken down into common categories outlined below: permitting, operations, reporting, emissions and other violations (FL DEP, 2003 "Guidelines for Characterizing..," pp. 3-8; J. Aslinger, personnel communication, December 2005). [41] The violation categories include the following:

1. Permit Violations:

- Failure to operate with a permit

[39] In most cases, before an EPA enforcement action, the state would issue an NOV. Most EPA enforcement actions consisted of late Title V yearly certifications.
[40] Georgia sends out a Letter of Non Compliance for less significant violations and an NOV for more significant violations. All other states, with the exception of Florida as noted above, send out NOV's for both minor and major violations.
[41] Based on personnel communication (email correspondence) with Julie Aslinger of the Tennessee Department of Environmental Conservation in December 2005.

- Late permit renewal.
- Construction without a valid permit

2. Operations violations, which are violations of permit requirements to show that equipment is operating properly:

- Failure to properly maintain or operate any equipment or pollution control device
- Failure to timely install or maintain continuous monitoring systems
- Failure to timely conduct required testing and monitoring
- Visual emission violations, a.k.a. opacity violation which measure smoke density and particulate emissions. This was included in this category and not in the "emissions exceedence" category, since evidence of a visual emission violations does not necessarily mean there has been an exceedence, but does mean equipment should be checked for exceedences (J. Aslinger, personnel communication, December 2005).[42]
- Circumvention of all or part of control device
- Excursions are defined as a facility operating equipment outside the required parameters that are set in the permit. This is typically called a surrogate limit violation. Since most facility do not have continuous emissions monitoring systems the best thing states can do to "make sure facilities do not exceed their permit emissions limit is to have them monitor something else, like baghouse pressure drop or scrubber flow rate. When a source has a pressure drop or flow rate that falls out of the specified operating range, it is listed as an excursion. It does not necessarily mean that an emission limit exceedance occurred, although it is possible." Rather, it technically is considered an operations

[42] Based on email correspondence with Julie Aslinger of the Tennessee Department of Environmental Conservation in December 2005.

violation and means that the equipment needs to be checked (J. Aslinger, personnel communication, December 2005).[43].

3. Recordkeeping and reporting requirement violations

- Late reports
- Inadequate reports
- No reports kept or submitted to the air office as required by the facility permit.

4. Emission exceedence

- Malfunction which exceeds allowable time
- Exceedences. An exceedance is defined as a short period above the emissions limitation.
- Release
- Emission exceedences based on stack test results

5. Other

- Any type of cited violation that does not fall within the prior four categories such as going over the allowed hours of operations specified in the permit. [44]

For each calendar quarter, it was recorded whether any one of these five types of violations were discovered. The discovery date was used, since the actual date was often not known. For example, it was often difficult to determine how long a piece of machinery had not been maintained; thus, it was difficult to determine the exact dates of many operations violations. If it

---

[43] Based on email correspondence with Julie Aslinger of the Tennessee Department of Environmental Conservation in December 2005.

[44] Clean Air Act 112(r) requires that facilities that keep on hand certain thresholds of chemicals have a program of preparedness, response and prevention to prevent chemical accidents. These risk management violations were not included, since this type of inspection is not always handled by the air office, so information on these violations were not readily available in facility files. (Based on email correspondence with Christopher Bradley, Enforcement Coordinator for the Southwest District Air Resources Management, Florida Department of Environmental protection, December 21, 2005). Also not all facilities are subject to CAA 112(r) requirements, so it was difficult to control for those who were subject to these requirements, and those facilities who were not.

was not clear when the violation was discovered, it was assumed that it was in the quarter the Notice of Violation or other warning letter was sent.[45]

A facility could receive a maximum score of five violations per quarter, i.e. one for each of the five categories. This meant that if a facility had two of the same kinds of violations, for example two reporting violations in a quarter, they would receive a score of one for reporting violation in that quarter. The maximum number of violations a facility could receive during the entire period of the study was 80 (5 maximum violations per quarter x 16 quarters= 80).

This approach was taken because states were inconsistent in the way they recorded violations. For example, some air offices would specify in their Notice of Violations' which reports were late when a reporting violation occurred. Others would generically state that "reports were late"- making it difficult to discern the number of late reports and when they should have been submitted. By assigning a one for any type of reporting violation in a quarter, regardless of the number of late reports, it allowed for a more consistent measure across states. This was also true with the way operations violations were reported. This approach did not really impact emissions violations. Only two facilities had duplicate emission violations in a quarter.

Severity

The TOTAL_VIOL dependent variable did not account for the severity of a violation, it only accounted for the frequency of violations. For example, some facilities had emissions violations that affected human health, while other facilities had a late report. In an attempt to account for severity, a severity variable was created using information on fines. Fines seem to indicate which violations the states viewed as serious violations (the higher the fine the more serious the violation) and were therefore used as a proxy measure for severity. The types of

[45] Few facilities received more than one NOV in a quarter.

violations (operations, permit, reporting, emissions, other) were ranked, based on average fines for each violation type.

The reason an ordinal variable was created rather than a continuous variable (e.g. total amount of fines received by a facility), was because using total fines did not provide an accurate way to assess the severity of a violation. Many states base fines on the extent of the harm caused by the violation and the recalcitrant nature of the violator. This makes it difficult to assess what portion of the fine was due to the recalcitrant nature of the violator compared to the severity of the violation. Additionally, many fines that are issued by enforcement agencies are "bundled," meaning that one large fine is issued for all violations uncovered by the enforcement agency during the same period of time. This makes it almost impossible to assess what portion of the fine should be attributed to a particular violation. Lastly, it can take time for a fine to be assessed; for many of the violations discovered during the sixteen quarters of the study, fines had not yet been issued.

To use fines as a proxy for seriousness of a particular type of violation, it was necessary to look at facilities that only received a fine for one type of violation and that did not have a fine in the year prior to the study (in order to eliminate the effect of a higher fine being meted out to repeat offenders).[46] In doing this, it was found that the category "other violation" had too few violations to be ranked, so they were not included in the ranking used to create the severity variable.[47] Emissions violations received the highest average fines at $5461. Permit violations received an average fine of $1262, while reporting and operations violations received similar average fines of $125 and $388 respectively. Given the results, emissions violations were ranked 3, permit violations were ranked 2, and reporting and operations violations were ranked 1, since

[46] Previous fines were limited to only one year prior to the study, given that only data back to 2000 was available.
[47] There were six "other" violations and only one violator had no other violation.

they were close in average amount (See Appendix B for a break down of fines by violation type). Based on this ranking system, facilities were given a rank based on the worst type of violation recorded in their record. For example, if a facility received a reporting violation and an emissions violation they would be given a rank of 3.

Control Variables

Several control variables were included in the analysis. Prior theory and studies in the area of environmental enforcement, discussed in Chapter One, suggest that facility inspections and fines may play a role in deterring permit violations; therefore, it is important to account for these influences. State air enforcement officials also pointed out that permit complexity and the type of permit a facility has also affects violation rates of facilities and should be controlled for in the study (J. Aslinger, personnel communication, December 2005).[48] Furthermore, given that states pursue enforcement of permit violations differently, and that different sub-classifications of chemical companies could have different violation rates due to the nature of their processes, it was necessary to account for these differences as well (U.S. GAO, June 2000, p. 2). For these reasons, the control variables discussed below were included in the study. All data for these variables came from state and local air offices and EPA files and databases.

Inspections

The regressor F_INSPEC, represents the total number of field compliance inspections at a facility during fiscal year 2001-2004 (the fiscal year spans from October through September). The F_INSPEC variable did not appear highly skewed; therefore, it was left as a continuous variable.

[48] Discussion with Julie Aslinger at Tennessee Department of Environmental Protection in December 2005.

Field compliance inspections included in the measure consist of all Level 2 and Partial Compliance Evaluations. The EPA and the states used one set of inspection codes prior to October 1, 2001 for "official inspections" under the CAA that consisted of different inspection levels. Level 2 and above were considered adequate for compliance assurance and were given a code so it could be reported in the AFS database. Beginning on October 1, 2001, a policy and subsequent database change created a set of new codes that are used to designate "compliance evaluations" (EPA, April 2001, pp. 4-6). EPA inspections were reclassified as Full Compliance Evaluations (FCE) or Partial Compliance Evaluations (PCE) (EPA, April 2001, pp. 4-6). The new codes were not uniformly adopted right away, with some states still reporting Level 2 and above inspections rather than FCE or PCE inspections. Level 2 inspections are defined as:

> A selective type of inspection in which the control device and process operating conditions are recorded as part of the source evaluation in addition to visible emission observations. In a typical application, the inspector would record such process items as feed rates, temperatures, raw material compositions, process rates, and such control equipment performance parameters as water flow rates, water pressure, static pressure drop, and electrostatic precipitator power levels. The inspector would then use these values to determine any significant change since the last inspection or any process operations outside normal or permitted conditions. Anything less than a Level 2 inspection is largely inadequate to determine compliance with many of the sources that DEP inspects. (McCabe, 1996, Chapter 3, para. 1)

A “Full Compliance Evaluations” (FCE’s) should include:

- Review of all required reports
- An assessment of control device and process operating conditions as appropriate
- A visible emissions observation as needed
- A review of facility records and operating logs
- An assessment of process parameters such as feed rates, raw material compositions and process rates
- An assessment of control equipment performance parameters

- A stack test when there is no other means for determining compliance with emissions limits (EPA, April 2001, pp.4-6).[49]

An FCE can be "accomplished piecemeal through a series of partial compliance evaluations" (p. 6).

Since there was not a consistent approach to reporting inspection data, all Level 2 inspections or partial compliance evaluations (PCE's) were counted. Some inspections can take place over several consecutive days; in that case, most states only reported the last date. Therefore, an inspection over consecutive days was counted as one inspection. By counting all level 2 or greater and/or PCE inspections, it was felt it created a better measure of enforcement pressure, since large facilities may require numerous PCE's over time to get one FCE inspection, due to the size of the facility. If only FCE inspections were counted, it could appear that large facilities are subject to less inspection pressure than small facilities. This may not necessarily be the case, since large facilities could have more PCE's in the same time period as smaller facilities- but overall have less FCE's.

Several types of inspections were not controlled for in the study. Drive-by inspections or investigations of complaint reports and report reviews were excluded, due to being inconsistently reported across states,[50] and not considered by the EPA to be an "acceptable compliance assurance method" (EPA, OIG, September 2000, p. 1). Although stack tests can be required by a regulatory agency whenever they think it is appropriate, stack tests were not included as part of the inspection measure since stack test data was not necessarily accurate (EPA, OIG, Sept 2000, p. i). According to the EPA's Office of Inspector General's report from September 2000, there was confusion over how stack tests should be performed and evaluated, and whether and what

[49] Stack tests can be required by a state/local agency "whenever they deem it appropriate" (EPA, April 2001, p. 5)
[50] Out of these types of inspections reported in the Florida data, none resulted in a violation

stack test information needed to be reported to the EPA (p. i). Risk management inspections required by CAA 112(r) were also not counted, since they were not always done by the local and state air offices as part of an air permit inspection, so it was difficult to assure consistent CAA 112(r) inspection data across states.[51]

Inspections- General Deterrence

Studies of this nature often include a general deterrence measure (Gray & Deily, 1991; Gray & Shadbegian, 2007). General deterrence is indirect deterrence, e.g. deterrence not targeted at a specific site, but deterrence based on how others are publicly monitored. It is thought that a facility will be more deterred from committing a violation if the potential to be inspected is higher in their state compared to facilities in other states. [52] Ultimately, a general deterrence measure was not created because there was not a significant difference in inspection rates across states.

To determine this, the number of facilities in the sample with a Title V permit in a given state, was divided by the number of total inspections (i.e. all inspections over the four year period of the study) of these facilities in the state. This number was then divided by four to get an average inspection rate per year.[53] The same was done for facilities with synthetic minor permits.

[51] As mentioned above, Clean Air Act § 112(r) requires that facilities that keep on hand certain thresholds of chemicals have a program of preparedness, response and prevention to prevent chemical accidents. These risk management violations were not included, since this type of inspection is not always handled by the air office, so information on these violations were not readily available in facility files. (Based on email correspondence with Christopher Bradley, Enforcement Coordinator for the Southwest District Air Resources Management, Florida Department of Environmental protection, December 21, 2005). Also not all facilities are subject to 112r requirements, so it was difficult to control for those who were subject to these requirements, and those facilities who were not.

[52] This is thought to have an effect even if the individual facility is not inspected frequently.

[53] This is a rough estimate, since some companies were inspected more than once in a year.

Each state inspected Title V facilities at a rate of 0.95 inspections per year or above. The change in policy that took effect in October 2001 did not seem to affect the inspection rates (EPA, April 2001). Georgia synthetic minor facilities had the lowest inspection rate of 0.34 per year per facility. The second lowest was North Carolina with synthetic minors having a 0.78 average inspection rate per year. The rest of the states inspected synthetic minor facilities at above 0.85 times per year. Given that only Georgia synthetic minor facilities were inspected at a very dissimilar rate compared to permittees in other states, it did not seem necessary to create a general deterrence measure since state control variables and permit control variables could adequately account for this difference.

Fines Issued in Previous Periods

To account for the potential deterrent effect of fines issued in prior periods, a dummy regressor PRE_FINE was used to identify facilities that had a fine in the year prior to the study, which was fiscal year 2000. Information on fines was obtained from state and local air office files, state enforcement databases, and EPA AFS database for EPA issued fines. Only fines issued one year prior to the study were used given the lack of available and accurate data prior to fiscal year 2000.[54] A value of one was assigned if the facility had a fine in the year prior to the study, and a zero was assigned if a facility did not have a fine a year prior to the study.

[54] David Konisky from University of Missouri, Department of Political Science provided the 1st quarter, fiscal year 2000, data for fines.

Complexity of Permit

PAGES is a proxy measure for the complexity of a facility's permit. Complexity was measured by the length of the permit. It was thought that the longer the permit, the more requirements a facility had to meet; thus, the more opportunities for a facility to have a violation.

Permits were obtained from state and local air offices. The permit that covered the time period of this study was used if available. Otherwise, the most recent permit available was used. Only the Title V or Synthetic Minor permits were measured; construction permits were not included as part of this variable.

How a permit was measured depended on whether an electronic version or hard copy was obtained. In the cases where an electronic copy was available, the permits were converted to 12 point, Times New Roman font before the numbers of pages were counted. When only a hard copy of a permit was available, 0.20 length per page was added (i.e. 20% per page) for 10 point font, and 0.10 length per page was added (10% per page) for 11 point font.

The PAGES variable was highly skewed to the right. To help reduce the skewness, the log of pages was taken and a new variable was created called PAGES_LOG. This helped reduce the skewness significantly. Running a Negative Binomial Regression showed the pseudo r-square went from .0248 to .0326 when using the transformed variable, indicating a slightly better fit.[55]

Type of Permit

All major permits, both Title V permits and synthetic minor permits (a.k.a. conditional major), were included in the study. Because reporting and emission requirements, as well as

[55] A Negative Binomial regression was run of just the variable pages against the dependent variable total_viol to get a pseudo r-square. A negative binomial regression was used because the dependent variable is a much skewed count variable.

inspection requirements, are different for Title V and synthetic minor permit holders a 0,1 dummy variable called TV_PERM was created to account for the differences among permits. A one was assigned to Title V permit holders, while a zero was assigned to synthetic minor permit holders.[56]

Temporary Permit

Many Title V permits took a while to process, and in the interim facilities were operating with a temporary permit that had the same operational requirements, but less reporting requirements than the final Title V permit. To account for potential differences in reporting violations due to facilities having temporary permits, a dummy variable TEMP_BI was created.[57] A one was given to a facility that had a temporary permit in any quarter, and a zero if the facility did not have a temporary permit in any quarter.

Facility SIC Code

Dummy variables were created for the nine, three-digit SIC 28 categories to control for any differences in results due to industry sub-classifications, e.g. differences in production requirements for the different classes of chemical industry goods. This means that a one was assigned to a facility that was assigned the respective SIC classification and a zero if it was not

[56] Pages and permit help control for the attainment status of the area a facility is located in, since the attainment status helps determine the type of permit a facility is required to have and the requirements that are included in the facility permit. For this reason, a separate attainment variable was not included. Nonattainment status affects existing firms by changing the requirements (often making them more stringent) in the permit the next permit cycle.

[57] A continuous variable was created titled TEMP_PERM. This variable is the numbers of quarters a facility had a temporary permit. This variable is highly skewed and transformation did not help to adjust for the skewness. Creating a variable based on the number of quarters a facility had a temporary permit, divided by the number of quarters in the study (sixteen) was also tried, but this variable, TEMP_Q, was also highly skewed and transformation did not help. To help alleviate any results due to skewness of the continuous variables, the dummy variable was created.

assigned the SIC classification.[58] The three-digit SIC code for the facility was determined by what was listed on the facility air permit.[59] The distribution of SIC codes is shown in Table 4.

Table 4: SIC Classifications Included in Study

SIC 28: Chemical and Allied Products

SIC-3 digit	Description	Total facilities
281	Industrial inorganic chemicals	41
282	Plastic materials and synthetics	69
283	Drugs	30
284	Soaps, cleaners and toilet goods	15
285	Paints and allied products	26
286	Industrial organic chemicals	48
287	Agricultural chemicals	25
289	Miscellaneous chemical products	28
****	SIC 28 is secondary	14
	TOTAL	296

Facility Location

Most states have petitioned the EPA to regulate and enforce environmental programs. This means that environmental programs have been delegated to the states or local air offices with oversight by the EPA. The United States Government Accounting Office (GAO) and academics have noted that states vary in stringency of permit requirements and enforcement of those requirements (U.S. GAO, June 2000). The GAO, in a report from 2000, gives a number of reasons for the variations in state enforcement including: budget constraints, differences in state's philosophy (using cooperative approaches vs. enforcement), and variations in state law granting state agency's authority to resolve complaints and recover penalties (p. 8). Given the differences of enforcement strategies across states, many studies that have looked at compliance

[58] SIC 284 and secondary SIC code facilities were put into the default category, since they constituted the smallest number of facilities. The default category becomes part of the constant value.

[59] The SIC code listed on the permit often differed from the SIC code listed in databases like Dun and Bradstreet's Million Dollar Database.

over one or more regions, controlled for regional and state variations (Nadeau, 1997; Gray & Deily 1996; Magat & Viscusi 1990). For these reasons state 0,1 dummy variables were created. A one was assigned if the facility was located in the respective state, and a zero if it was not.[60] Table 5 reports on the number of facilities by state.

Facilities that were part of local county programs were included in the state variable, since the number of facilities in the sample that came under the jurisdiction of these programs was small. Additionally, a local program dummy variable was created to see if a different effect existed for local versus state programs.

Table 5: Number of Facilities Per State[61]

State	Number of facilities
Florida	26
Georgia	70
North Carolina	83
South Carolina	60
Tennessee	57

Local Air Program

A 0,1 dummy variable called LOCAL was created to identify if a local county program office handled permitting and enforcement and was not a branch of a state office. A one was assigned if the facility was in a jurisdiction run by a local office, while a zero was assigned if the facility was in a jurisdiction run by a state office. Tennessee, North Carolina and Florida have local program offices. (See Appendix G for a list of state and local programs.)

[60] Florida was the default when the statistics were run, since Florida had the smallest number of facilities in the sample. The default category becomes part of the constant value.
[61] This represents roughly 60% of chemical facilities with major permit in Region 4, and 10% of all major permitted privately held facilities (e.g. not state or federal facilities) in Region 4. These estimates are based on facilities listed in EPA ECHO database that meet these criteria when retrieved in December 2005 and May 2009 from http://www.epa-echo.gov/echo/compliance_report_air.html.

Theoretical Variables

A number of theoretical variables, outlined below, were created to help test the proposed hypotheses.

Public or Privately Held Company

A 0,1 dummy variable was created called PUBLIC. One indicates that the facility is part of a public company (i.e. listed on a stock exchange), while zero indicates the facility is part of a privately held company. Several facilities changed ownership during the period of the study. In those cases, the facility was designated public if it was part of a public company for the majority of the study, or private if it was part of a private company for the majority of the study. The dataset includes 42 facilities out of 296 that were part of a company that went through a merger or acquisition; out of those only six changed from public to private or vice-versa.

Distance of Facility from Headquarters

A continuous variable was created based on miles to headquarters located in the United States. This could be either the U.S. headquarters, for foreign-owned companies, or the parent head office for U.S. based companies. The variable was highly skewed and transforming the variable did not improve skewness, so a dummy variable was created called OUT_ST. A facility was assigned a one if the headquarters was in a different state than where the facility was located and a zero if the headquarters was in the same state.

Single Facility Operations

A 0,1 dummy variable called SINGLE was created to look at the difference between single facility companies and firms with multi-facility operations. One represents that the facility is a single facility company as opposed to part of a multi-facility firm.

Foreign Parent Company

The regressor FOREIGN is a 0,1 dummy variable. A zero indicates the ultimate parent company that owns the facility is domestically located (in the United States), while a 1 indicates that the ultimate parent is located abroad. This is consistent with the approach taken by King and Shaver (2001); they identified foreign ownership as the “ultimate owner of the facility” (p. 1074). Several facilities changed ownership during the period of the current study. In those cases, the facility was designated “foreign” if it was part of a company with a foreign, ultimate parent for the majority of the study or domestic if it was part of a domestic company for the majority of the study.

Number of Employees at Facility

The regressor LN_EMP is a continuous variable that represents the total number of employees at a facility. Most employee numbers are reported yearly. If a company had 50 employees for two years and 100 for the next two years an average of 75 was taken. Most employee numbers did not fluctuate much over the four-year period; however ten cases did change substantially. This variable had to be transformed into a log variable to help reduce the skewness. Running a negative binomial regression, the r-square went from .0056 to .0072 indicating the transformation helped slightly.[62]

Firm Size

SM, MED, and BIG are dummy regressors. For SM, a one indicates that the facility is part of a company with more than five employees, but less than 2,310 employees (bottom 33.3% of sample in terms of number of employees at the company).[63] For MED, a one indicates that the facility is part of a company with more than 2400 employees and less than 24,530 employees (33.3 to 66.6 % of sample). For BIG, a one indicates that the facility is part of a company with more than 24,839 employees, but less than 315,082 employees (i.e. the top third of cases in the sample in terms of company size).[64]

Exploratory Variable

Merger

This variable is a 0,1 dummy variable where zero indicates that no merger or acquisition of the company that owned the facility occurred during the period of the study, and one represents that the company that owned the facility went through a merger or acquisition during the period of the study. Four facilities that were part of companies who were acquired as a result of a bankruptcy were taken out of the data set. This was done because unlike mergers and

[62] A Negative Binomial regression was run with the untransformed variable and again with the transformed variable to get a pseudo r-square. A negative binomial regression was used because the dependent variable is a very skewed count variable. The dependent variable was total_viol

[63] SM was the default when the statistics were run. The default category becomes part of the constant value.

[64] A similar continuous variable was created. The log form was used to reduce skewness; however, the log form did seem to improve the pseudo r-square. It should be noted, however, that when run in the model, it got the same results as with the dummy variables.

acquisitions, environmental liabilities are not negotiated between parties when there is a bankruptcy (see exploratory analysis section in the last chapter for a discussion about why this was necessary). Forty-two companies (out of 296) in the dataset went through some from of merger or acquisition.

Sample Size

Data was collected on 314 facilities. This dropped to 296 for the final sample. All facilities that went through a bankruptcy, were part of a joint venture, state owned company or nonprofit or coop were dropped from the sample. Also, the one mega facility in the sample was dropped, Eastman in Tennessee, since as a mega facility it was subject to different requirements than Title V permit holders and synthetic minor permit holders (EPA, April 2001, p. 7). Additionally, Eastman proved to be an outlier for many variables such as inspections, number of employees at the facility, and permit length.[65]

Descriptive Statistics

Tables 6 through 9 provide descriptive statistics for all for of the variables used in this study.

[65] The exclusion of this variable did not impact the results discussed in Chapter Three.

Table 6: Control Variables Descriptive Statistics

Variable	Variable Name	Observations	Mean	Standard Deviation	Minimum	Maximum
Facility Inspections	F_inspec	296	3.466	1.824	0	21
Florida	FL	296	0.088	0.284	0	1
Georgia	GA	296	0.236	0.425	0	1
North Carolina	NC	296	0.280	0.449	0	1
South Carolina	SC	296	0.202	0.402	0	1
Tennessee	TN	296	0.192	0.394	0	1
Local program	Local	296	0.166	0.373	0	1
Permit length	Pages	291	30.481	30.382	2	250
Log of permit length	Pages_log	291	3.000	0.936	0.69	5.52
Temp Permit	Tempbi	290	0.258	0.438	0	1
Fine in previous period	Pre_fine	295	0.098	0.298	0	1
Industrial inorganic chemicals	SIC_281	296	0.138	0.346	0	1
Plastic materials and synthetics	SIC_282	296	0.233	0.423	0	1
Drugs	SIC_283	296	0.101	0.302	0	1
Soaps, cleaners and toilet goods	SIC_284	296	0.050	0.219	0	1
Paints and Allied Products	SIC_285	296	0.087	0.283	0	1
Industrial organic chemicals	SIC_286	296	0.162	0.369	0	1
Agricultural chemicals	SIC_287	296	0.084	0.278	0	1
Miscellaneous chemical products	SIC_289	296	0.094	0.293	0	1
SIC 28 is secondary	SIC_Other	296	0.047	0.213	0	1

Table 7: Theoretical Variables Descriptive Statistics

Variable	Variable Name	Observations	Mean	Standard Deviation	Minimum	Maximum
Facility is part of large sized firm	Big	294	0.336	0.473	0	1
Facility is part of medium sized firm	Med	294	0.330	0.471	0	1
Facility is part of small firm	Small	294	0.333	0.472	0	1
Log of the number of employees at the facility	Ln_emp	295	4.844	1.251	1.39	7.83
Foreign ultimate parent company	Foreign	296	0.314	0.465	0	1
Headquarters are in a different state	Out_St	295	0.783	0.413	0	1
Facility is part of publicly traded firm	Public	296	0.652	0.477	0	1
Single plant facility	Single	296	0.074	0.263	0	1

Table 8: Dependent Variables Descriptive Statistics[66]

Variable	Variable Name	Observations	Mean	Standard Deviation	Minimum	Maximum
Total number of violations	Total_viol	296	1.033	1.477	0	9
Severity of violation	Severity	293	0.812	1.058	0	3

Table 9: Exploratory Variable Descriptive Statistics

Variable	Variable Name	Observatio ns	Mean	Standard Deviation	Minimum	Maximum
Facility was part of company that went through a merger or acquisition	Merger	296	0.142	0.350	0	1

[66] See variable construction section for a complete understanding of these dependent variables.

Correlations

If independent variables in a model are highly correlated multicollinearity could occur, resulting in coefficients for the affected variables that might be inaccurate. Since the control variables PAGES_LOG, which measures the length of a facility's permit, and TV_PERM which indicates whether a facility has a Title V permit of synthetic minor permit, were highly correlated (0.74), PAGES_LOG was dropped from the final analysis to avoid multicollinearity. Essentially, both variables appear to measure permit complexity. Having one of these variables in the model was thought to be sufficient to control for this effect.

Correlations between the independent variables in the model and the dependent variable TOTAL_VIOL (total violations over 16 quarters) were also obtained. Based on the bivariate correlations in Table 10, the variables with the highest correlation with the dependent variable were the control variables F_INSPEC (number of facility inspections) at 0.28, and TV_PERM (facility has a Title V permit) at 0.20. The theoretical variables with the highest correlations were LN_EMP (the log of the number of employees at the facility), at 0.17, and SINGLE (single facility firm) at -0.10. The firm-level variables have the lowest correlations, which seems to indicate that firm-level variables may be poor predictors, on their own, of facility violations.

Table 10: Bivariate Correlations

Correlations between the following variables and the dependent variable Total_Viol		
Variable	**Definition**	**Correlation with Total_Viol**
PUBLIC	Facility is part of publicly traded firm	-0.05
PRE_FINE	Fine in previous period	0.10
BIG	Facility is part of large sized firm	-0.03
MED	Facility is part of medium sized firm	0.03
SMALL	Facility is part of small firm	0.01
OUT_ST	Facility is part of a company with headquarters in a different state than where the facility is located	-0.07
SINGLE	Single facility firm	-0.10
F_INSPEC	Facility inspections	0.28
LN_EMP	The log of the number of employees at the facility	0.17
TV_PERM	Title V Permit	0.20
FOREIGN	Foreign parent company	-0.01
MERGER	Facility was part of company that went through a merger or acquisition	-0.05

For a complete table of pairwise correlations see Appendix C.

Model Selection

Several statistical techniques were employed to examine the dependent variable SEVERITY. Both ordinal logistic regression and multinomial regression were attempted. Neither ordinal logistic regression nor multinomial logistic regression resulted in models that fit the data, meaning they could not be used to analyze the dependent variable SEVERITY (see Appendix D & E for the results of the ordinal and multinomial logistic regression analysis for the variable SEVERITY).

Panel data analysis, poisson and negative binomial regression techniques were tried to analyze the amount of violations over 16 quarters.[67] Given the lack of sufficient variation over time in the independent and dependent variables, panel data analysis was excluded as a sufficient model to help examine facility violations (see Appendix F for more detail about panel data analysis). Therefore, it was necessary to conduct a cross-sectional analysis to test the hypotheses, e.g. compare the differences among the facilities in the sample without regard to changes over time. It was found that the best technique for analyzing the cross sectional data was a negative binomial regression model.

Poisson and Negative Binomial Regression Techniques

To analyze the data as a cross section, the data were collapsed into one value for all 16 quarters (see section on variable construction for more detail). For example, all violations over the 16 quarter period were summed to construct the dependent variable TOTAL_VIOL, and all inspections over the 16 quarter period were added up to make the independent variable F_INSPECT (i.e. the total number of facility inspections), etc.

After looking at the distribution of the dependent variable, it was determined that a count model such as Poisson Regression, Negative Binomial Regression (NBRM), Zero Inflated Poisson or Zero Inflated NBRM should be used. These models are used for non-negative distributions that are highly skewed. As can be seen from the Figure 2 and Table 11, the dependent variable is highly skewed to the right. Given the skewed nature of the variable, it made sense to use one of these count models. These models adjust for the skewness of the outcome variable, thus, preventing the model from producing negative predicted values. In

[67] For panel analysis a binary variable was constructed as the dependent variable that assigned a one if the facility had a violation that quarter and a 0 if it did not. For the Poisson and negative binomial regression total_viol was used as the dependent variable.

contrast, if linear regression were used on highly skewed data, it would result in negative predicted values, since linear regression assumes the variance is constant, which means little to no skewness.

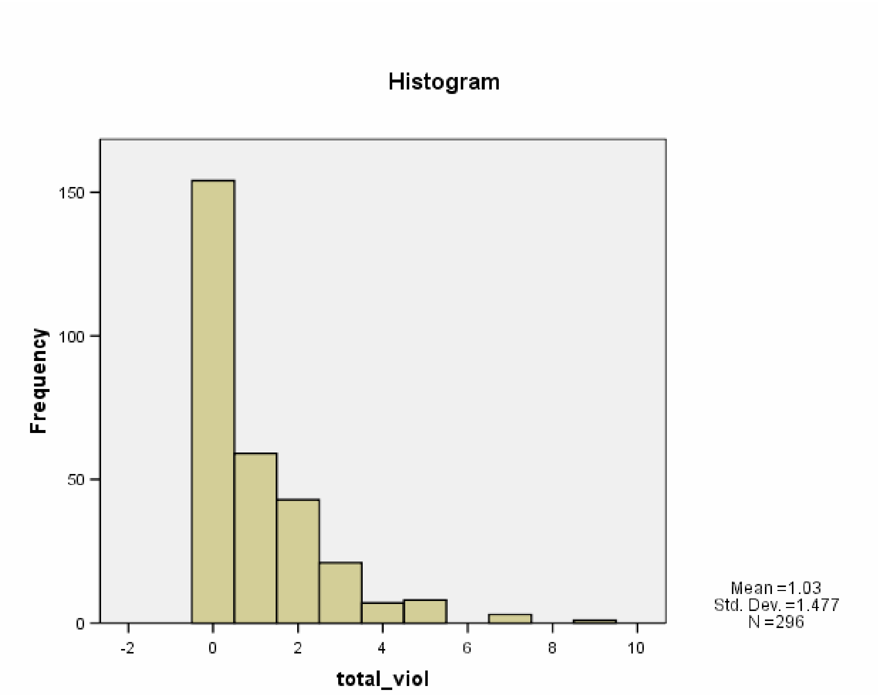

Figure 2: Bar Graph of Total Violations

Table 11: Frequencies of Violations (TOTAL_VIOL)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	154	52.0	52.0	52.0
	1	59	19.9	19.9	72.0
	2	43	14.5	14.5	86.5
	3	21	7.1	7.1	93.6
	4	7	2.4	2.4	95.9
	5	8	2.7	2.7	98.6
	7	3	1.0	1.0	99.7
	9	1	.3	.3	100.0
	Total	296	100.0	100.0	

As suggested by Long (1997), all four count models were evaluated to determine which one was the most appropriate model for the data set (p .218). Different assumptions underlie the four models. The Poisson distribution assumes the mean and variance are equal. In contrast, the NRBM allows for a variance that is larger than the mean. By doing this, the NRBM takes into account unobserved between-subject heterogeneity, e.g. unobserved differences between facilities in the sample that cause violation rates to differ across facilities and resort in overdispersion; however, the same heterogeneity that causes overdispersion can cause zero inflation (Long, 1997 p. 230; Drukker, 2000, sec. Long Answer). Zero-inflation means a high probability of zero counts. If the dependent variable has a lot of zero values, Zero Inflated Poisson (ZIP) and Zero Inflated NRBM's (ZINB) can be employed, since the NBRM and Poisson Regression models will often under-predict the amount of zeros. ZIP models allow the probability of a zero to be larger than what a Poisson or NRBM model would predict (Long, 1997, p. 218).

Where the PRM and NBRM models would assume that each facility has a positive probability of having any given number of violations, the ZIP models allow for the probability to differ by assuming the population consists of two groups: one group that will always be zero; and one group that has the probability of having a positive count (Long, 1997, p. 243). Long gives the example of a study looking at the publication rates of scientists. He points out that if there are two groups, i.e. one for scientists that had jobs where publishing was not possible, and one for scientists that could publish, than zero modified models could help account for the excess zeros caused by the differences between groups (p. 242).

There are several tests that can compare models. PRM and NBRM can be compared by looking at the alpha dispersion parameter. If the *t*-test of the alpha parameter is statistically significant at $p < 0.05$, then it suggests there is significant overdispersion, i.e. the NBRM is suited to the data (Long, 1997, pp. 237, 247). Vuong test statistic compares non-nested models. Specifically, it compares the PRM model to the ZIP model, and the ZINB model to the ZINB model. If the Vuong test statistic is greater than the "critical value of +1.96 than the first model is favored, if the Vuong test statistic is less than -1.96, the second model is favored; otherwise neither model is preferred" (p. 248). There is not a readily available test to compare ZIP to ZINB models. All four statistical techniques and available tests were run to determine which one was the best fit for the data used in this study.

A poisson regression, negative binomial regression, zero inflated poisson and zero negative binomial regressions were run to test the hypotheses that foreign companies have higher costs developing the expertise to comply and that various types of firms have agency costs making it harder to comply with environmental regulations. According to the results of the negative binomial regression, the alpha dispersion parameter was significant at 0.000. The Vuong test statistics for the zero inflated models were as follows: V(ZIP/PRM)= 3.74 and V(NBREG/ZINB)= 1.36. These test statistics suggest that either the ZIP or NRBM models suit the data. Since the test statistics indicate that both models could fit the data, a determination needed to be made about which model made sense substantively to use (Long, 1997, p. 249). The ZIP model did not seem well suited, since ZIP models assume that there are two groups in the sample with different probabilities of having zero counts: one with no opportunity to commit a violation and one that does. Two different groups do not exist within this data set; all subjects had the opportunity to violate the regulations. As a result, it seems more likely that the heterogeneity that caused the overdispersion was also causing the excess zeros. This meant that it made more sense to use the NRBM.[68]

[68] While the NRBM was used, it should be noted that both models produced similar results.

CHAPTER III

RESULTS

This first part of this chapter reports the outcomes of the statistical analysis and examines the results in light of the hypothesis proposed in the first chapter. Included in this section is a review of an exploratory analysis of the effect of mergers on compliance. The last portion of this chapter discusses how the study contributes to theory and policy, outlines the limitations of the analysis and presents ideas for future research.

Statistical Results

A negative binomial regression (NRBM) was run to analyze the effect of firm and facility characteristics discussed in Chapter I, on facility violations rates with air permit requirements.

The following is the model fit information:

Negative binomial regression	Number of obs = 291
	LR chi2(21) = 48.40
	Prob > chi2 = 0.0006
Log likelihood = -386.18071	Pseudo R2 = 0.0590

The chi-square value suggests that the independent variables help predict violations rate outcomes; however, the pseudo r-square value indicates that the overall model was not robust at predicting facility violation rates. A pseudo r-square value closer to one means the model is a better predictor of the outcome compared to a value closer to zero. It should be noted that the pseudo r-square for this model is similar to the pseudo r-square for other recently published work in this area (see Gray & Shadbegian, 2007; Gray & Shadbegian, 2005). While the pseudo r-square value for the overall model fit was low, some individual coefficients were significant.

Tables 12 reports the coefficients for the independent variables. Values highlighted in red indicate variables that are significant at p≤0.05

Table 12: Negative Binomial Regression Coefficients

Dependent Variable Total Violations[69]						
Variable	**Coefficient**	**Std. Error**	**Z**	**P>|z|**	**95% confidence interval**	
PUBLIC	-0.01	0.24	-0.06	0.96	-0.48	0.46
PRE_FINE	0.26	0.27	0.96	0.34	-0.27	0.78
BIG	-0.15	0.29	-0.54	0.59	-0.72	0.41
MED	-0.08	0.28	-0.30	0.77	-0.63	0.46
OUT_ST	-0.55	0.25	-2.20	0.03	-1.04	-0.06
SINGLE	-0.99	0.42	-2.38	0.02	-1.81	-0.18
F_INSPEC	0.14	0.06	2.47	0.01	0.03	0.24
LOG_EMP	0.16	0.08	2.03	0.04	0.01	0.32
TV_PERM	0.37	0.19	1.97	0.05	-0.00	0.74
FOREIGN	-0.08	0.19	-0.41	0.68	-0.45	0.30
SC	0.26	0.35	0.77	0.44	-0.41	0.94
NC	0.44	0.33	1.32	0.19	-0.21	1.10
TN	-0.06	0.37	-0.17	0.87	-0.78	0.66
GA	0.12	0.37	0.33	0.74	-0.60	0.84
SIC_281	-0.07	0.36	-0.19	0.85	-0.78	0.64
SIC_282	-0.15	0.32	-0.47	0.64	-0.79	0.48
SIC-283	-0.35	0.39	-0.90	0.37	-1.2	0.41
SIC-285	0.17	0.39	0.43	0.66	-0.60	0.93
SIC_286	0.20	0.33	0.61	0.54	-0.44	0.84
SIC_287	0.32	0.39	0.82	0.41	-0.45	1.10
SIC_289	0.23	0.37	0.62	0.53	-0.49	0.95
Constant	-1.17	0.64	-1.84	0.07	-2.42	0.08
Lnalpha | -.2879359 .236685 -.75183 .1759581						
Alpha | .7498096 .1774687 .4715029 1.192388						
Likelihood-ratio test of alpha=0: chibar2(01) = 48.28 Prob>=chibar2 = 0.000						

Number of observations = 291

[69] The control variables TEMP_PERM, that looked at whether a facility had a temporary permit, and LOCAL which looked at whether a facility was in the jurisdiction of a local county program, were not significant when added, so they were excluded from the final model to allow for a more parsimonious fit.
A sensitivity analysis was also run to test the effect of the cases having gone through a merger or acquisition. The model was run excluding all cases that had gone through a merger or acquisition, which brought the sample down to 250 cases. The same variables were significant at p<.05, except for TV_PERM which was significant at p=0.06

Table 13 reports the percentage change in expected count of violations, for each of the independent variables. These estimates could be interpreted, for example, as follows: being a facility owned by a publicly held firm (PUBLIC), holding all other variables constant, decreases the expected number of violations by 1.30%, compared to facilities owned by privately held firms.

Table 13: List of Percentage Change In Expected Count

Dependent Variable Total Violations						
Total violations	B	Z	P>|z|	%	%StdX	SDofX
PUBLIC	-0.01	-0.05	0.96	-1.30	-0.60	0.48
PRE_FINE	0.26	0.96	0.34	29.20	8.00	0.30
BIG	-0.15	-0.57	0.60	-14.30	-7.10	0.47
MED	-0.08	-0.30	0.77	-8.0	-3.90	0.47
OUT_ST	-0.55	-2.20	0.03	-42.30	-20.20	0.41
SINGLE	-0.99	-2.38	0.02	-63.00	-23.10	0.26
F_INSPEC	0.14	2.47	0.01	14.60	28.20	1.82
LOG_EMP	0.16	2.03	0.04	17.40	22.20	1.25
TV_PERM	0.37	1.98	0.05	45.00	20.50	0.50
FOREIGN	-0.08	-0.41	0.68	-7.40	-3.50	0.46
SC	0.26	0.77	0.44	30.30	11.20	0.40
NC	0.44	1.33	0.19	55.70	22.00	0.45
TN	-0.06	-0.17	0.87	-5.90	-2.40	0.40
GA	0.12	0.33	0.74	12.70	5.20	0.43
SIC_281	-0.07	-0.19	0.85	-6.70	-2.30	0.34
SIC_282	-0.15	-0.47	0.64	-14.20	-6.20	0.42
SIC_283	-0.35	-0.91	0.37	-29.70	-10.20	0.31
SIC_285	0.17	0.43	0.66	18.40	4.90	0.28
SIC_286	0.20	0.61	0.54	22.10	7.70	0.37
SIC_287	0.32	0.82	0.41	38.30	9.50	0.28
SIC_289	0.23	0.62	0.53	25.80	7.00	0.30
Number of observations = 291			Dichotomous 0-1 expected change			
Observed SD: 1.4862406			SDofX = standard deviation of X			
Ln alpha | -0.28794 -1.217						
B = raw coefficient						
Z = z-score for test of b=0						
P>|z| = p-value for z-test						
% = percent change in expected count for unit increase in X						
%StdX = percent change in expected count for SD increase in X						

Table 14 reports the marginal effects (note, that for dichotomous variables the discrete change is reported). These estimates, can be interpreted, for example, as follows: being a facility owned by a publicly held firm (PUBLIC), compared to a privately held firm, decreases the expected count of violations by 0.01 violations, holding all other variables at their mean.

Table 14: Marginal effects

Dependent Variable Total Violations							
Total violations	Dy/dx	Std. Err.	Z	P>|z|	95% confidence interval		X
PUBLIC*	-0.01	0.21	-0.05	0.96	-0.43	0.41	0.66
PRE_FINE*	0.25	0.29	0.87	0.36	-0.32	0.82	0.10
BIG*	-0.13	0.24	-0.55	0.58	-0.61	0.34	0.34
MED*	-0.07	0.24	-0.30	0.76	-0.54	0.40	0.33
OUT_ST*	-0.58	0.31	-1.87	0.06	-1.18	0.03	0.79
SINGLE*	-0.60	0.17	-3.58	0.00	-0.93	-0.27	0.08
F_INSPEC	0.12	0.05	2.46	0.01	0.02	0.22	3.45
LOG_EMP	0.14	0.07	2.03	0.04	0.00	0.28	4.85
TV_PERM*	0.33	0.17	1.95	0.05	-0.00	0.66	0.49
FOREIGN*	-0.07	0.16	-0.41	0.68	-0.38	0.25	0.31
SC*	0.25	0.36	0.71	0.48	-0.45	0.95	0.20
NC*	0.43	0.37	1.19	0.23	-0.28	1.15	0.28
TN*	-0.05	0.31	-0.17	0.87	-0.66	0.56	0.19
GA*	0.11	0.34	0.32	0.75	-0.57	0.78	0.24
SIC_281*	-0.06	0.31	-0.19	0.85	-0.66	0.54	0.13
SIC_282*	-0.13	0.26	-0.49	0.62	-0.65	0.39	0.23
SIC_283*	-0.27	0.26	-1.04	0.30	-0.78	0.24	0.10
SIC_285*	0.16	0.40	0.40	0.69	-0.62	0.94	0.09
SIC_286*	0.19	0.33	0.57	0.57	-0.46	0.84	0.16
SIC_287*	0.33	0.46	0.72	0.47	-0.57	1.23	0.09
SIC_289*	0.22	0.30	0.57	0.57	-0.55	0.99	0.10
Y= predicted number of events (predict)							
= 0.88329757							
* dy/dx is for discrete change of dummy from 0 to 1							

Discussion

The results did not provide any evidence that foreign companies had higher violation rates, but did provide some mixed evidence that principal-agent problems affect facility environmental compliance.

Agency Costs and Firm/Facility Size

As discussed earlier, agency theory suggests that facilities with more employees, facilities that are part of multi-facility firms and facilities that are part of large firms will incur higher agency costs to monitor employees, resulting in higher compliance costs compared to smaller firms/facilities and single facility firms. Becker's optimal penalty model indicates that the higher a firm's compliance costs and the lower the expected penalty, the less incentive firms have to comply (Becker, 1968). This would suggest, controlling for government monitoring and enforcement, that large facilities, facilities part of large firms, and multi-facility firms would have lower compliance rates.

Both SINGLE (single location firm) and LN_EMP (the log of the number of employees at the facility) results are statistically significant and the coefficients are in the direction expected. The coefficients for these variables suggest that both multi-facility firms and larger plants are more likely to be non-compliant. More specifically, Table 13 shows that being a single facility, holding all other variables constant, decreases the expected number of violations by 63%. According to Table 14, this would mean that being a single facility decreases the expected count of violations by 0.60 violations, holding all other variables at their mean. Table 12 shows that LN_EMP had a coefficient of B= 0.16. This suggests that a 1% increase in employees is associated with a roughly 0.16% increase in expected count of violations holding all else in the model fixed. This means that the more employees at the facility, the more likely the facility will have a violation.[70]

These results are consistent with the proposed agency theory hypotheses; both multi-facility firms and larger facilities are more likely to have higher violation rates. Some may, however, call into question whether the significance of LN_EMP is a result of agency costs. It could be the case that larger facilities have the opportunity to have more violations, since they are often subject to more permit requirements. However, the TV_PERM variable was included in the model to help control for permit complexity and was significant (see discussion of permit complexity in Chapter Two).

Contrary to what agency theory would indicate, firm size (BIG) did not prove to be significant. Even though this variable was not significant, there wasn't any clear evidence to support the alternative theory of economies of scale, since larger firms did not perform better than smaller ones. This does not mean that economies of scale or agency costs do not exist with regards to firm size and compliance. Both effects may exist and it could be the case that one effect may offset the other.

Table 15 compares the results of this study to prior published research in the area. As shown in the table, the result on facility size was consistent with prior research on compliance, but not all research looking at TRI emissions.[71] On the other hand, the lack of significance of firm size is also consistent with prior research on both compliance and TRI emissions. The evidence for SINGLE (single facility firm) is mixed. The finding that multi-facility firms had lower compliance rates than single facility firms is consistent with earlier work on air pollution

[70] This interpretation is based on the interpretation found in Simonoff (2003, p. 136)
[71] While the results are consistent with studies on compliance, the results of facility size and it's effect on TRI emissions was mixed.

by Gray and Deily (1996) and studies on TRI emissions, but not with more current research on air pollution by Gray and Shadbegian (2005, 2007).

Agency Theory and Distant Headquarters

Similar to firm and facility size, it was thought that firms with facilities that are in a different state than the headquarters would incur more agency costs monitoring employees. This added compliance cost could mean that facilities not located in the same state as the head office would have higher violations rates compared to those whose headquarters are situated in the same state. The variable OUT_ST (facility is part of a company with headquarters in a different state than where the facility is located) was statistically significant at $p < .05$, but produced an opposite result from what was expected. The coefficient for OUT_ST indicated that facilities that are distant from the firm headquarters are more likely to be in compliance. In fact, Table 13 indicates that if the facility is part of a firm with headquarters in a different state, the facility would have 42.3% less expected violations than facilities that are part of firms with headquarters in the same state, holding all other variables constant. Table 14 suggests that facilities that are not in the same state as their company headquarters, have a lower expected count of violations by 0.58 violations, holding all other variables at their mean. It should be noted that not only are these results not as hypothesized, they are also not in line with prior research using TRI data. Grant, Jones and Trautner (2004) found when looking at TRI emissions that out-of- state headquarters did not have much of an effect on TRI emissions.

It is difficult to explain the outcome for the variable OUT_ST. One explanation put forth is that the out-of-state headquarters may be in states with strict environmental enforcement programs, thereby influencing company wide compliance programs. To test for this potential effect, a proxy measure of strictness was created. It was assumed that states with maximum contaminant levels for benzene that are lower than what federal law requires for drinking water would have ‘stricter” enforcement programs. The states with stricter requirements are New Jersey, California and Florida. Facilities with out-of-state headquarters located in these states were coded a one; a zero was assigned to the other cases. The variable “strict” was then added to the model. The variable did not prove to be significant.[72] A potential alternative explanation is that facilities, whose company’s headquarters are out-of-state, have managers that take more ownership in their facilities. This could mean these facilities have characteristics that are more like single facility firms.

Agency Theory and Corporate Ownership

Analogous to the previous discussions on agency costs, agency theory suggests that the more owners/managers that a company has, the more agency costs the company would incur monitoring employees. As a result, it was thought that facilities that are part of publicly traded firms with many shareholder/owners would have lower compliance rates than facilities owned by privately held companies which tend to have fewer owners. However, it was found that PUBLIC was not a statistically significant predictor of expected violation rates. This implies that this firm level variable is not a helpful indicator of facility violation rates.

Cost to Develop Compliance Expertise

In addition to agency theory, another idea discussed in the literature that could affect an individual facility’s ability to comply is whether the facility incurs more costs than others

[72] Strict included 50 cases with U.S. headquarters in NJ, CA and FL. Pairwise correlations were run with the other independent variables in the model; strict was not highly correlated with the other independent variables. The p-value of strict when included in the model was 0.965.

developing the expertise to comply with local regulations. It was thought that foreign-owned facilities may face higher costs developing this expertise, since they may have less experience understanding and complying with domestic U.S. regulations and the domestic regulatory structure. Assuming firms are profit maximizers, Becker's optimal penalty model suggests that if foreign-owned facilities have higher compliance costs than domestically owned facilities, they would have less incentive to comply with regulations, given the same level of government monitoring and enforcement. This would imply that, controlling for government monitoring and enforcement, foreign-owned facilities would have lower compliance rates.

The variable FOREIGN (foreign-owned facilities) was not statistically significant. This implies that foreign ownership of facilities is not a good predictor of expected violations. This result is not consistent with the results of King and Shaver (2001) that found, using TRI data, that foreign-owned facilities generate and manage more waste than their domestically owned counterparts (p. 1069).

It should be noted that this analysis may have limitations. It was suggested that if a company has operated for a long time in the United States they could have already developed the expertise to comply with local environmental regulations. It would be useful in future research to control for the date in which the foreign owner entered the U.S. market.

Current Study Comparisons with TRI Studies

It should be mentioned that the reason some of the current study findings including FOREIGN (facility owned by foreign company), OUT_ST (facility is part of a company with headquarters in a different state than where the facility is located) and LN_EMP (the log of the number of employees at the facility) are inconsistent with a number of the prior published studies looking at TRI emissions may be due to the nature of TRI releases compared to permit compliance requirements. Many of the chemicals required to be reported under the TRI program are not necessarily permitted emissions, meaning facilities are not required to, for example, monitor or meet emissions limits for these chemicals. In contrast, failing to meet the requirements outlined in a Clean Air Act permit can result in enforcement by government agencies and fines.

Control Variables

Turning to control variables, the results show that most of the control variables do not have a significant effect on the outcome, with p-values above 0.10, except for TV_PERM (facilities with a Title V permit), which is significant at $p < 0.10$, and F_INSPEC (number of facility inspections) which is significant at $p < 0.05$. According to Table 13, facilities with Title V permit, holding all other variables constant, have a 45% expected increase in total violations compared to facilities with synthetic minor permits. This would mean, according to Table 14, that a facility with a Title V permit would have 0.33 more expected violations than a facility with a synthetic minor permit, holding all other variables at their means. TV_PERM helps account for permit complexity, so this result is likely due to the fact that facilities with more complex permits have more opportunities to have violations. For each inspection, Table 13 indicates that there is an expected increase in violations of 14.6%, holding all other variables constant. This means the more inspections a facility is subject to, the more violations that could be discovered by regulatory agencies; hence, the importance of controlling for this effect. The lack of significance of the other state and SIC control variables suggest that the SIC sub-classification and state in which a facility is located in did not have an effect on predicting the observed outcomes. The control variable PRE_FINE (fine issued in fiscal year prior to the study) also did not appear to have a significant effect on violation rates.

Table 15: Comparison of Results with Published Research

Study	Sample	Facility Size	Firm Size	Multi-facility location	Distant Headquarters	Foreign-owned facility	Public vs. Private Ownership
Current Study Results	Compliance with air pollution requirements (chemical industry)	Large facilities are less likely to be in compliance	No effect found on compliance rates	Facilities part of multi-facility firms are less likely to be in compliance	Facilities that are distant from headquarters are less likely to be in compliance	No effect found on compliance rates	No effect found on compliance rates
Gray and Deily (1996)	Compliance with air pollution requirements (steel plants)	Large facilities are less likely to be in compliance	___	Firms with multiple plants are less likely to be in compliance	___	___	___
Gray and Shadbegian (2005)	Compliance with air pollution requirements (pulp and paper)	Large facilities are less likely to be in compliance	No effect found on compliance rates	No effect found on compliance rates	___	___	___
Gray and Shadbegian (2007)	Compliance with air pollution requirements and emissions (TRI emissions and conventional air pollutants $PM_{2.5}$, SO_2. (manufacturing)	1)Large facilities are less likely to be in compliance 2) Larger facilities had lower emissions.[73]	___	Facilities part of multi-facility firms are more likely to be in compliance	___	___	___
Grant, Bergesen and Jones (2002)	Toxic Release Inventory (TRI) Emissions (chemical industry)	Large plants emit TRI toxins at a higher rate than smaller plants	Firm size did not have a direct, significant impact on TRI emission rates.	Large plants emit more toxins… especially if "embedded in a wider corporate structure"(pp. 389, 402)	___	___	___
King and Shaver (2001)	Toxic Release Inventory Emissions (chemical and petroleum industries)	___	___	Firms generate more waste if 'they operate multiple facilities across multiple jurisdictions in the United States' (p1079)	___	Foreign-owned firms tend to 'generate and manage more hazardous waste' than domestic firms (p. 1069)	___
Grant, Jones and Trautner (2004)	Toxic Release Inventory Emissions (chemical industry)	___	___	___	No effect found on TRI emission rates.	___	___

[73] Gray and Shadbegian (2007) note that "since emissions are calculated relative to plant size, and only plants with relatively large emissions are included in the EPA database" the results for plant size "can hardly be treated as evidence of economies of scale in controlling emissions" (pp. 75-76).

Exploratory Analysis of Merger

The main scope of this study was to look at the effects of observable firm and/or facility characteristics on compliance rates. A related issue suggested in the legal literature is that mergers and acquisition may impact compliance rates. Since the effect of mergers and acquisitions on facility compliance rates could be tested by adding a variable to the model discussed in the previous section, an exploratory analysis was conducted to investigate whether mergers and acquisitions had an impact on violation rates.

It is suggested that companies that undergo a merger or acquisition may have an incentive to comply or over comply with regulations (Vandenbergh, 2003). Vandenbergh suggests that sellers, before announcing a sale, may want to do a housecleaning to reduce environmental liabilities (pp.2047-2048). Additionally, he notes that a buyer, as part of assessing liabilities of an acquisition, may discover environmental non-compliance and pressure the seller to correct any non-compliance (pp. 2049-2051). Also, he mentions that if the seller indemnifies the buyer of environmental liabilities as part of their negotiations, after closing, the seller may engage in monitoring of the buyer to ensure the "buyer does not act in a way that will trigger indemnified liability" (p. 2050-2051). This suggests that there may be an incentive to increase compliance at facilities that are part of a company going through a merger or acquisition, both before and after the merger or acquisition commences.[74] To test the effect of mergers and acquisitions, a merger variable was created. The merger variable was coded one if a facility was part of a company that went through a merger or acquisition, and zero otherwise. This variable was added to the NRBM model.

[74] To test the effect of mergers, all of the facilities that were part of companies that went through a bankruptcy were taken out of the sample, since companies going through bankruptcies may have environmental claims get wiped out; this means that a buyer and seller would not negotiate liabilities.

The results of including the merger and acquisition variable called MERGER into the model is shown in Table 16. The p-value for the merger variable was not significant, suggesting the variable is not a good predictor of facility violations rates. It had a p-value of > 0.10. Additionally, the pseudo r-square for the model did not improve much by adding MERGER. The pseudo r-square went from 0.059 to 0.0596, indicating that adding the merger variable did not really make the model more robust at predicting facility violation rates. It should be noted that this analysis may not have been sensitive enough to identify the effect of mergers and acquisitions as discussed in Vandenbergh (2003). The variable did not account for how liabilities between parties were allocated. It was not possible to account for this in this study, because only publicly traded company agreements are available and this study includes both public and privately held firms.[75]

[75] Publicly traded companies need to file material acquisition agreements with the Securities and Exchange Commission (See discussion in Vandenbergh, 2003, pp. 2045-2046).

Table 16: Results Including Merger

Dependent Variable Total Violations				
	Without merger		With Merger	
Variable	**Coefficient**	**P>|z|**	**Coefficient**	**P>|z|**
MERGER	----	---	-0.18	0.47
PUBLIC	-0.01	0.96	-0.02	0.94
PRE_FINE	0.26	0.34	0.26	0.34
BIG	-0.15	0.59	-0.15	0.61
MED	-0.08	0.77	-0.07	0.80
OUT_ST	-0.55	0.03	-0.54	0.03
SINGLE	-0.99	0.02	-0.99	0.02
F_INSPEC	0.14	0.01	0.13	0.02
LOG_EMP	0.16	0.04	0.17	0.04
TV_PERM	0.37	0.05	0.38	0.05
FOREIGN	-0.08	0.68	-0.07	0.70
SC	0.26	0.44	0.28	0.42
NC	0.44	0.19	0.45	0.17
TN	-0.06	0.87	-0.05	0.88
GA	0.12	0.74	0.11	0.76
SIC_281	-0.07	0.85	-0.05	0.89
SIC_282	-0.15	0.64	-0.15	0.65
SIC-283	-0.35	0.37	-0.37	0.35
SIC-285	0.17	0.66	0.19	0.63
SIC_286	0.20	0.54	0.19	0.56
SIC_287	0.32	0.41	0.30	0.44
SIC_289	0.23	0.53	0.21	0.56
Constant	-1.23	0.07	-1.19	0.06

Number of observations = 291

Conclusion

The effect of several firm and facility-level characteristics on violation rates were tested, as measured by compliance with facility air permit requirements. A large amount of the explanatory power of the model came from several theoretical variables including SINGLE (single facility firm), LN_EMP (the log of the number of employees at the facility), and OUT_ST (facility is part of a company with headquarters in a different state than where the facility is located). SINGLE and LN_EMP are consistent with the agency theory hypothesis put forward, while OUT_ST was opposite from what was expected. Firm level characteristics, such as BIG (facility is part of a large firm), PUBLIC (facility is part of a publicly held company), and FOREIGN (facility owned by a foreign parent) were not significant. The results suggest that these firm level variables are not good predictors of facility violation rates. In summary, some evidence is found that multi-facility firms and larger facilities have higher agency costs that could add to their compliance costs, which results in these facilities having lower compliance rates. No evidence is found that large firms, facilities that are in a different state than their company headquarters, or facilities that are part of publicly owned companies incur agency costs that impact their compliance rates. Additionally, no evidence is found that foreign-owned facilities face higher costs developing the expertise to comply with local environmental regulations, which would lead to lower compliance rates than domestically owned firms. Lastly, an exploratory analysis of the effect of merger and acquisitions on compliance did not yield any significant results.

Contribution to Theory and Policy

Despite the mixed results, the evidence does help improve upon theory by adding to the literature's understanding of agency costs and firm compliance rates, as well as whether foreign-owned facilities have a higher cost developing the expertise to comply with local requirements. Further empirical research is needed to better comprehend the existence (or non-existence) of principal-agency problems in the context of facility environmental compliance rates. Additionally, further work could be done to explore which types of firms may have higher costs developing the expertise to comply with regulations.

While the results are not definitive regarding the effects of firm and facility characteristics on environmental compliance, there are some noteworthy results that could aid policymakers. Consistent with prior research, there appears to be observable characteristics that are associated with compliance, such as facility size. On the other hand, there appears to be other characteristics that are not likely to be indicative of compliance, such as firm size. Other characteristics looked at in this study are less conclusive. Given these results, it could be beneficial to further explore firm and facility characteristics to determine whether these results are similar when looking at other media or industries. This information could prove useful to enforcement officials who want to create more targeted environmental enforcement strategies.

Study Limitations

It should be noted that this study has several limitations. In addition to data quality problems, the sample was limited to one type of media, a limited number of states, and one industry. Given these limitations, this study could not be considered the final word on agency theory and facility compliance rates and whether foreign-owned facilities face higher costs developing the expertise to comply. This means, for example, that the results should not be generalized to other media, such as Clean Water Act and hazardous waste management compliance, since those programs have a unique structure and set of rules. Empirical testing would be required to see if the same results would hold, if applied to those media. Additionally, a larger sample that included more regions may also provide more insight into whether firm and facilities characteristics affect facility compliance rates in other geographical locations.

Moreover, the results should not be generalized to other industries, since different industries may face different compliance challenges. For example, different industries may manage more hazardous waste then others, or have more (or less) complex operations. Different industries may also have different external pressures that affect their compliance rates. In a survey of industrial managers in the chemical and electroplating industries done by Gunningham, Thorton, and Kagen (2005), some respondents from the chemical sector were concerned about maintaining their “social license to operate” in their respective communities after major accidents like Bhophal (pp.310-313). Their paper notes that these types of concerns are why the chemical industry started programs like Responsible Care (pp. 310-313). As a result, it could be the case that the chemical industry is more sophisticated than other industries, meaning it could have more compliers and less violators than other industries. Empirical testing would be required to see if the same results would hold, if applied to different industries.

Lastly, it should also be noted that the information used for this study had data quality issues that could have affected the outcome. To conduct this study it was decided to collect state level data, so that high and low priority violations could be included and more accurate inspection data could be obtained for control purposes. Despite this effort, data from several states in EPA Region 4 could not be obtained. Additionally, while state level data were collected, some control variables could not be included such as stack tests, since they were not consistently reported across states (see Chapter 2 for a more in depth discussion on data quality issues). In any event, it should be kept in mind that data quality issues may have hampered the ability to adequately test the hypotheses.

Future Research

Expansion of Current Work

Future empirical research, which looks at the effect of firm and facility characteristics on compliance rates, could be expanded in a number of different directions. One direction would be to better refine some of the theoretical variables such as FOREIGN, e.g. control for how long a company has operated in the local environment, and MERGER, e.g. look at how specific liabilities were negotiated during the merger and acquisitions process. Another direction would be to broaden the scope of the sample to include plants from more regions and plants in other industries to see if the results are similar. Most studies in this area have focused on the chemical, petroleum and paper industries. Additionally, the sample could be expanded to obtain enough cases so that firms with multiple permits, such as Clean Water Act permits and RCRA (hazardous waste management) permits are included. This would provide an opportunity to examine the effects of firm and facility characteristic on multi-media compliance. It would be interesting to see if there were similar outcomes when looking at these media. Additionally, it would be beneficial to determine whether non-compliance in one area means that a facility is likely out of compliance in another, since there is very little research in this area (see Gray & Shadbegian, 2005). Understanding this relationship could help enforcement officials develop multi-media enforcement strategies. Lastly, it would be potentially beneficial to determine a better way to collect and analyze the effect of firm/facility characteristics on the severity of violations. Understanding if there is a relationship between firm/facility characteristics and the severity of violations could help create enforcements strategies that target the most serious violations.

All of these proposed expansions on the current research may be hindered by data collection and quality issues. State data is more difficult to collect than U.S. EPA data, but may provide more accurate results than U.S. EPA data (see discussion in Chapter 2 on data quality issues). While using state level data for this study allowed for some data cleansing, it is difficult to know for sure whether using these data compared to U.S. EPA data made a significant difference in testing the proposed hypotheses. Given the fact that the results of this study are similar to prior research; it may suggest there is no significant difference between state and federal data. However, most prior research was conducted on data from the 1990's, before the EPA implemented it's current High Priority Violation reporting policy (Schaeffer, 1998, p. 6). This policy may have changed the way states reported enforcement date into EPA's database. An interesting study for the future would be to run the same analysis using current state level facility data and U.S. EPA data and compare outcomes. This could help determine whether state level data is better suited for empirical analysis.

Additional Variables to be Included in Future Work

Many variables suggested in the literature that affect firm compliance could not be included in this study due to the limitation of the sample, but would be interesting to explore in future research projects. For example, Konar and Cohen (working paper, March 2000) suggested that those firms closest to the consumer (measured by level of ad expenditures and forms of consumer marketing) vs. products sold as intermediate goods would be more likely to comply, because they may be more apt to be rewarded in the marketplace for good behavior. This also implies that those firms with a strong brand image have more to lose and will be more likely to comply; however, previous studies have not found there to be an effect. Since many facilities in this study produced intermediate goods, these characteristics could not be tested. A sample with a larger range of companies would be needed to further test these theories.

Other variables that were not included in this study due to data availability include the age of capital stock and the financial health of the firm. It is thought that the older the capital stock, the harder it is to comply with regulations because the equipment would be older and less efficient, resulting in more violations (See Konar & Cohen working paper,[76] March 2000). These data were compiled at the company level for publicly traded firms, and are not available at the facility level. It is also thought that poor financial health could result in decisions to not invest in pollution prevention, which tend to be long term investments. Less investment in pollution prevention could result in more violations. This is called the "failing firm hypothesis." If a firm has a lot of debt, and less available capital, a firm may focus on short term results. Konar and Cohen (March 2000 working paper) found those "firms with constrained cash flows were least likely to reduce emissions" (p. 8). Another method of operationalizing the 'failing firm hypothesis' is to look at the rate of sales growth. Cohen and Alexander (1996) found that a low rate of sales and employment growth by a firm, over time, tends to be a good predictor of environmental crime (pp. 421, 433). This theory stems from the principal agent context. The more the agent feels his job is secure, the less likely he will risk 'cheating' (Cohen & Alexander, 1996). Data on firm financial health to test these theories is only available for publicly held companies. Since this study focused on facility level compliance and included facilities owned by both publicly traded and privately owned firms, these variables could not be incorporated into this study. These variables, however, could potentially be included in future studies using a subset of the data.

[76] They talk about this in relationship to environmental performance.

Some other factors to consider for future research projects, which were not incorporated into this study, include accounting for the attitude and expertise of individuals within the firm. Poor performance could be related to a change in Environmental Health and Safety (EH&S) managers. When an experienced EH&S manager leaves, it can take time to find an experienced individual to take over and learn the permit requirements.[77] Good performance could be a result of a belief that penalties are much greater than actual amounts, risk aversion, or that firms feel that compliance is just the right thing to do (Cohen, 1999).[78] Studies that include data on individuals within the organization require a different data collection methodology which was beyond the scope of this current study.

[77] It was mentioned by some of the regulators at TDEC that when a new EH&S manager comes to the job, there can be a bit of a learning curve. During this time mistakes could be made resulting in violations. If this is true, then having training programs for new hires on requirements may be appropriate to help reduce violation rates.
[78] See Cohen (1999) for a complete overview of theory and empirical work in the area of factors that effect corporate and facility level environmental compliance.

A. FREQUENCIES OF VIOLATIONS BY TYPE

Table 17: Frequencies of Permit Violations

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	259	87.5	87.8	87.8
	1	34	11.5	11.5	99.3
	2	2	.7	.7	100.0
	Total	295	99.7	100.0	
Missing	System	1	.3		
Total		296	100.0		

Table 18: Frequencies of Operations Violation

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	221	74.7	74.9	74.9
	1	51	17.2	17.3	92.2
	2	20	6.8	6.8	99.0
	3	2	.7	.7	99.7
	4	1	.3	.3	100.0
	Total	295	99.7	100.0	
Missing	System	1	.3		
Total		296	100.0		

Table 19: Frequencies of Recordkeeping Violations

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	217	73.3	73.6	73.6
	1	59	19.9	20.0	93.6
	2	14	4.7	4.7	98.3
	3	5	1.7	1.7	100.0
	Total	295	99.7	100.0	
Missing	System	1	.3		
Total		296	100.0		

Table 20: Frequencies of Emissions Exceedences

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	255	86.1	86.4	86.4
	1	35	11.8	11.9	98.3
	2	3	1.0	1.0	99.3
	3	1	.3	.3	99.7
	4	1	.3	.3	100.0
	Total	295	99.7	100.0	
Missing	System	1	.3		
Total		296	100.0		

Table 21: Frequencies of Other Violations

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	283	95.6	95.9	95.9
	1	8	2.7	2.7	98.6
	2	4	1.4	1.4	100.0
	Total	295	99.7	100.0	
Missing	System	1	.3		
Total		296	100.0		

B. FINE BY VIOLATION TYPE

The following tables represent the average fine per violation type for facilities with only one violation and no fine in the year previous to the study. These tables were used to determine the ranking for the dependent variable SEVERITY.

Table 22: Permit Violation Fines

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00	6	60.0	60.0	60.0
	1225.00	1	10.0	10.0	70.0
	2000.00	1	10.0	10.0	80.0
	3000.00	1	10.0	10.0	90.0
	6400.00	1	10.0	10.0	100.0
	Total	10	100.0	100.0	

Table 23: Operations Violation Fines

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00	11	91.7	91.7	91.7
	1500.00	1	8.3	8.3	100.0
	Total	12	100.0	100.0	

Table 24: Reporting Violation Fines

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00	15	83.3	83.3	83.3
	1500.00	1	5.6	5.6	88.9
	2000.00	1	5.6	5.6	94.4
	3500.00	1	5.6	5.6	100.0
	Total	18	100.0	100.0	

Table 25: Emissions Violation Fines

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00	7	50.0	50.0	50.0
	1000.00	2	14.3	14.3	64.3
	2560.00	1	7.1	7.1	71.4
	6000.00	1	7.1	7.1	78.6
	15000.00	1	7.1	7.1	85.7
	17000.00	1	7.1	7.1	92.9
	33900.00	1	7.1	7.1	100.0
	Total	14	100.0	100.0	

C. PAIRWISE CORRELATIONS

Table 26 Dependent Variable Pairwise Correlations

	TOTAL_VIOL	SEVERITY
TOTAL_VIOL	1	.651
SEVERITY	.651	1
	293	293

Table 27: Theoretical Variable Pairwise Correlations

	PUBLIC	FOREIGN	SINGLE	MED	SMALL	BIG	OUT_ST	MED_PUB	BIG_PUB	SM_PUB	MERGER	LN_EMP
PUBLIC	1.00											
FOREIGN	0.30	1.00										
SINGLE	-0.39	-0.19	1.00									
MED	0.22	-0.02	-0.20	1.00								
SMALL	-0.63	-0.23	0.40	-0.50	1.00							
BIG	0.41	0.25	-0.20	-0.50	-0.50	1.00						
OUT_ST	0.43	0.11	-0.48	0.24	-0.58	0.33	1.00					
MED_PUB	0.44	0.01	-0.17	**0.86**	-0.42	-0.43	0.24	1.00				
BIG_PUB	0.49	0.25	-0.19	-0.47	-0.48	**0.95**	0.32	-0.40	1.00			
SM_PUB	0.21	0.08	-0.08	-0.20	0.41	-0.21	-0.19	-0.17	-0.19	1.00		
MERGER	0.05	0.08	-0.08	0.11	-0.10	0.00	0.09	0.04	-0.02	0.06	1.00	
LN_EMP	0.30	0.06	-0.25	0.03	-0.27	0.24	0.19	0.07	0.21	0.05	0.16	1.00

Table 28: Control Variable Pairwise Correlations

	TEMP_PERM	TV_PERM	F_INSPEC	TN	GA	NC	SC	SIC_281	SIC_282	SIC_283	SIC_284
TEMP_PERM	1										
TV_PERM	0.46	1.00									
F_INSPEC	0.07	0.34	1.00								
TN	-0.03	-0.02	0.01	1.00							
GA	-0.03	-0.10	-0.34	-0.27	1.00						
NC	0.11	-0.10	0.02	-0.30	-0.35						
SC	0.05	0.16	0.12	-0.25	-0.28	-0.31	1.00				
SIC_281	0.05	0.14	-0.01	0.13	0.08	-0.14	-0.01	1.00			
SIC_282	0.09	0.08	-0.07	-0.11	0.01	0.01	0.08	-0.22	1.00		
SIC_283	0.07	-0.08	-0.05	-0.11	-0.06	0.21	-0.09	-0.13	-0.19	1.00	
SIC_284	-0.10	-0.10	-0.02	-0.03	-0.02	0.03	0.08	-0.09	-0.13	-0.08	1.00
SIC_285	-0.04	-0.02	-0.04	0.09	-0.03	0.07	-0.13	-0.12	-0.17	-0.10	-0.07
SIC_286	0.04	0.05	0.11	-0.03	-0.03	-0.05	0.10	-0.18	-0.24	-0.15	-0.10
SIC_287	-0.14	-0.01	0.14	-0.09	0.06	0.00	-0.12	-0.12	-0.17	-0.10	-0.07
SIC_289	-0.05	-0.13	-0.04	0.08	-0.02	-0.07	0.07	-0.13	-0.18	-0.11	-0.07
PAGES_LOG	0.48	**0.75**	0.42	-0.02	-0.21	-0.06	0.22	0.07	0.11	-0.02	-0.05
PRE_FINE	0.04	0.20	0.14	0.16	0.00	-0.13	-0.03	0.10	-0.02	-0.07	-0.02

	SIC_285	SIC_286	SIC_287	SIC_289	PAGES_LOG	PRE_FINE
SIC_285	1.00					
SIC_286	-0.14	1.00				
SIC_287	-0.09	-0.13	1.00			
SIC_289	-0.10	-0.14	-0.10	1.00		
PAGES_LOG	-0.19	0.13	0.05	-0.21	1.00	
PRE_FINE	-0.06	-0.02	0.10	-0.03	0.19	1.00

Table 29: Pairwise Correlations of Variables in Model

	TOTAL_VIOL	PUBLIC	PRE_FINE	BIG	MED	CBIG_PUB	CMED_PUB	OUT_ST	SINGLE	F_INSPEC	LN_EMP
TOTAL_VIOL	1										
PUBLIC	-0.05	1.00									
PRE_FINE	0.10	0.02	1.00								
BIG	-0.03	0.41	-0.07	1.00							
MED	0.02	0.22	0.08	-0.50	1.00						
CBIG_PUB	-0.08	-0.31	-0.04	0.38	-0.56	1.00					
CMED_PUB	0.09	-0.15	0.04	-0.49	0.19	-0.42	1.00				
OUT_ST	-0.06	0.43	0.06	0.33	0.24	-0.12	-0.17	1.00			
SINGLE	-0.10	-0.39	-0.01	-0.20	-0.20	0.19	0.23	-0.48	1.00		
F_INSPEC	0.28	0.01	0.14	-0.01	0.07	-0.11	0.03	0.05	-0.05	1.00	
LN_EMP	0.17	0.30	0.13	0.24	0.03	-0.14	-0.12	0.19	-0.25	0.20	1.00
TV_PERM	0.20	0.16	0.20	0.13	0.09	-0.11	0.00	0.17	-0.15	0.34	0.33
FOREIGN	-0.01	0.30	-0.03	0.25	-0.02	-0.03	-0.18	0.11	-0.19	0.03	0.06
SC	0.06	0.02	-0.03	0.02	0.03	0.02	-0.04	-0.02	-0.01	0.12	0.04
NC	0.05	-0.02	-0.13	0.07	0.02	0.03	-0.15	0.01	0.00	0.02	0.11
TN	-0.08	-0.09	0.16	-0.08	-0.09	0.08	0.07	-0.03	0.02	0.01	-0.05
GA	-0.10	0.07	0.00	-0.01	0.02	-0.04	0.06	-0.04	0.02	-0.34	-0.17
SIC_281	-0.04	0.01	0.10	-0.11	0.14	-0.07	0.08	0.05	-0.04	-0.01	-0.18
SIC_282	-0.04	0.10	-0.02	0.07	0.04	-0.05	-0.03	0.09	-0.13	-0.07	0.16
SIC_283	-0.06	0.17	-0.07	0.19	-0.09	0.07	-0.15	0.07	-0.10	-0.05	0.28
SIC_285	-0.01	-0.02	-0.06	0.01	-0.09	0.07	0.05	-0.04	0.09	-0.04	-0.12
SIC_286	0.07	-0.08	-0.02	-0.04	0.02	-0.01	0.04	-0.06	0.08	0.11	-0.03
SIC_287	0.12	-0.14	0.10	-0.06	0.07	-0.03	0.02	0.01	0.10	0.14	-0.03
SIC_289	0.00	-0.08	-0.03	-0.04	-0.06	0.08	0.01	-0.03	0.00	-0.04	-0.13

	TV_PERM	FOREIGN	SC	NC	TN	GA	SIC_281	SIC_282	SIC_283	SIC_285	SIC_286	SIC_287	SIC_289
TV_PERM	1.00												
FOREIGN	0.04	1.00											
SC	0.16	0.02	1.00										
NC	-0.10	0.05	-0.31	1.00									
TN	-0.02	-0.04	-0.25	-0.30	1.00								

	TV_PERM	FOREIGN	SC	NC	TN	GA	SIC_281	SIC_282	SIC_283	SIC_285	SIC_286	SIC_287	SIC_289
GA	-0.10	0.02	-0.28	-0.35	-0.27	1.00							
SIC_281	0.14	0.00	-0.01	-0.14	0.13	0.08	1.00						
SIC_282	0.08	-0.05	0.08	0.01	-0.11	0.01	-0.22	1.00					
SIC_283	-0.08	0.04	-0.09	0.21	-0.11	-0.06	-0.13	-0.19	1.00				
SIC_285	-0.02	-0.11	-0.13	0.07	0.09	-0.03	-0.12	-0.17	-0.10	1.00			
SIC_286	0.05	0.02	0.10	-0.05	-0.03	-0.03	-0.18	-0.24	-0.15	-0.14	1.00		
SIC_287	-0.01	-0.02	-0.12	0.00	-0.09	0.06	-0.12	-0.17	-0.10	-0.09	-0.13	1.00	
SIC_289	-0.13	0.08	0.07	-0.07	0.08	-0.02	-0.13	-0.18	-0.11	-0.10	-0.14	-0.10	1

D. ANALYSIS OF SEVERITY USING ORDINAL LOGISTIC REGRESSION

An ordinal logistic regression was used to try and analyze the severity variable. Ordered logistic regression is used when the dependent variable consists of more than two categories that can be ranked in a meaningful way, but the distance between categories is unknown. The severity dependent variable, as discussed in previous sections, ranks violation types committed by facility by seriousness, meaning a facility that received an emissions violation got the highest ranking, and a facility with no violation during the period of the study received 0.

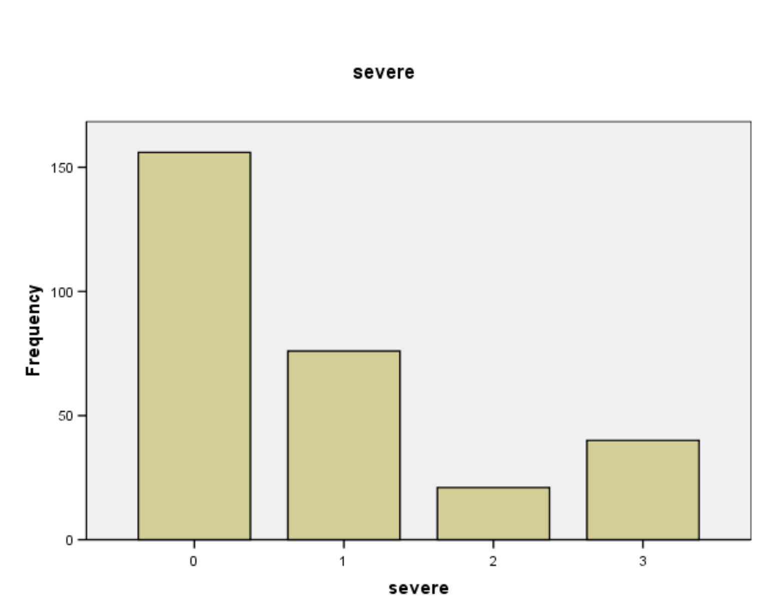

Figure 3: Bar Graph of Severity Dependent Variable

An assumption of ordinal logistic regression is that “the relationship between each pair of outcome groups is the same” also known as the “proportional odds assumption or parallel regression assumption” (UCLA: Academic Technology Services, n.d, Sec. Using the Ordinal Logistic Model). This means that the coefficients that describe the relationship between the different ordered categories of the dependent variable are the same for the lowest to the highest category, so only one model is needed vs. different models for each pair of categories. If this assumption is not met, a different model would need to be used to describe the relationships between the categories. The proportional odds assumption can be tested using a likelihood ratio test. The null hypothesis indicates there is no difference in the coefficients between models; this means if the model is a good fit the chi-square should not be significant (UCLA: Academic Technology Services, n.d., Sec. Using the Ordinal Logistic Model).

When an ordered logistic regression was run, the chi-square indicated it violated the proportionality odds assumption. The chi-square was 110.12 with a probability of 0.000. Since the likelihood ratio test is significant at 0.000, the model is a poor fit to the data and could not be used to reliably analyze the dependent variable.

E. ANALYSIS OF SEVERITY USING MULTINOMINAL LOGISTIC REGRESSION

Multinomial logistic regression was also tried to analyze the severity variable. This technique can be used when the parallel regression assumption of ordered logistic regression is violated. Multinomial regression does not assume that the categories can be ordered in a meaningful way, rather it assumes the categories of the dependent variable are nominal (a set of more than two categories that have no particular order). Multinomial logistic regression results evaluate one group against a comparison group; in this case the comparison group was those facilities that received no violation during the period of the study. When running the multinomial logistic regression using SPSS statistical software, I received the following warning:

> Unexpected singularities in the Hessian matrix are encountered. This indicates that either some predictor variables should be excluded or some categories should be merged. The NOMREG procedure continues despite the above warning(s). Subsequent results shown are based on the last iteration. *Validity of the model fit is uncertain.*

The results of the regression are shown in the following tables:

Table 30: Case Processing Summary

		N	Marginal Percentage
Severe	0	154	53.3%
	1	74	25.6%
	2	21	7.3%
	3	40	13.8%
Valid		289	100.0%
Missing		7	
Total		296	
Subpopulation		288(a)	

a The dependent variable has only one value observed in 288 (100.0%) subpopulations.

Table 31: Model Fitting Information

Model	Model Fitting Criteria	Likelihood Ratio Tests		
	-2 Log Likelihood	Chi-Square	Df	Sig.
Intercept Only	663.831			
Final	521.398	142.433	69	.000

Table 32: Pseudo R-Square

Cox and Snell	.389
Nagelkerke	.433
McFadden	.215

Table 33: Likelihood Ratio Tests for Independent Variables included in the Multinomial Logistic Regression Model of Severity[79]

Effect	Model Fitting Criteria	Likelihood Ratio Tests		
	-2 Log Likelihood of Reduced Model	Chi-Square	Df	Sig.
Intercept	536.000	14.602	3	.002
PUBLIC	524.470	3.072	3	.381
FOREIGN	525.321	3.924	3	.270
TV_PERM	527.491	6.093	3	.107
F_INSPEC	532.914	11.516	3	.009
SINGLE	530.471	9.073	3	.028
PRE_FINE	522.535	1.137	3	.768
MED	521.705	.307	3	.959
BIG	524.933	3.535	3	.316
OUT_ST	529.680	8.282	3	.041
C_BIG_PUB[80]	522.639	1.241	3	.743
C_MED_PUB	528.060	6.662	3	.083
LN_EMP	526.828	5.430	3	.143
TN	526.665	5.267	3	.153
GA	524.187	2.789	3	.425
NC	534.482	13.084	3	.004
SC	534.186	12.788	3	.005
SIC_281	523.728	2.330	3	.507
SIC_282	529.086	7.688	3	.053
SIC_283	533.443	12.045	3	.007
SIC_285	523.839	2.441	3	.486
SIC_286	526.633	5.235	3	.155
SIC_287	525.778	4.380	3	.223
SIC_289	524.712	3.314	3	.346

[79] The chi-square statistic is the difference in -2 log-likelihoods between the final model and a reduced model. The reduced model is formed by omitting an effect from the final model. The null hypothesis is that all parameters of that effect are 0.

[80] It was proposed that facilities owned by publicly traded companies may have fewer violations due to market pressure. An interaction variable was also added to the model to test whether facilities part of large, public firms will have fewer violations than those who are part of large, private firms, because market pressure will have a moderating effect. No effect was found. The result did not change if these interaction variables were taken out.

The likelihood ratio test results in Table 31 and the Pseudo R-Square results in Table 32, suggest that using predictors to determine severity of violations, is better than not including these variables; however, the warning and the likelihood ratio tests results shown in Table 33 for the independent variables in the model indicate many of the predictors used do not help explain the outcomes for severity. The likelihood ratio tests values, according to Pampel (2000) “reflects the likelihood that the data would be observed given the parameter estimates” (p. 45). The test compares the baseline model (with only the constant value, and all other variables set to 0), and a model with the independent variables to help see if adding in the independent variables helps to produce the observed outcomes values (Pampel, 2000, pp.45- 46). Most of the likelihood ratio test values for the independent variables used in the model are not significant (e.g. p-value $<.05$), meaning they do not appear to add value to the model in terms of predicting outcomes of severity; only a few variables show a significance level of $p <.05$. This suggests these predictors may not be good predictors of severity of violation types, when comparing severity as separate groups.

The coefficients produced by the model as shown in Table 34, do not allow for meaningful analysis and most are not significant. Overall, the results of the multinomial regression show that the model poorly fits the data and cannot be used to analyze the severity of violations. It could be the case that the predictors used in the model are poor predictors of severity, or that the severity dependent variable categories are not easily comparable, or the variable severity is not a good measure of severity. The model could also be a poor fit due to the fact there are too few observations to run the multinomial model.

Table 34: Results of Multinomial Regression

Parameter Estimates									
Severe(a)		B	Std. Error	Wald	Df	Sig.	Exp(B)	95% Confidence Interval for Exp(B)	
								Lower Bound	Upper Bound
1	Intercept	-2.45	1.38	3.16	1.00	0.08			
	PUBLIC	-0.09	0.52	0.03	1.00	0.87	0.92	0.33	2.53
	FOREIGN	0.51	0.37	1.90	1.00	0.17	1.66	0.81	3.42
	TV_PERM	0.78	0.38	4.26	1.00	0.04	2.18	1.04	4.59
	F_INSPEC	0.30	0.14	4.81	1.00	0.03	1.35	1.03	1.76
	SINGLE	-0.84	0.69	1.49	1.00	0.22	0.43	0.11	1.66
	PRE_FINE	-0.36	0.59	0.38	1.00	0.54	0.70	0.22	2.19
	MED	-0.01	0.53	0.00	1.00	0.99	0.99	0.35	2.78
	BIG	-0.61	0.66	0.86	1.00	0.35	0.54	0.15	1.97
	OUT_ST	-0.42	0.50	0.70	1.00	0.40	0.66	0.25	1.75
	C_BIG_PUB[81]	0.61	1.39	0.19	1.00	0.66	1.84	0.12	27.87
	C_MED_PUB[82]	1.11	0.94	1.40	1.00	0.24	3.03	0.48	19.05
	LN_EMP	0.05	0.16	0.10	1.00	0.75	1.05	0.77	1.43
	TN	0.56	0.92	0.38	1.00	0.54	1.76	0.29	10.65
	GA	1.31	0.91	2.09	1.00	0.15	3.72	0.63	22.10
	NC	2.31	0.87	7.04	1.00	0.01	10.08	1.83	55.59
	SC	0.96	0.89	1.15	1.00	0.28	2.61	0.45	15.04
	SIC_281	-0.86	0.62	1.91	1.00	0.17	0.42	0.12	1.43
	SIC_282	-1.45	0.59	6.01	1.00	0.01	0.23	0.07	0.75
	SIC_283	-2.42	0.84	8.37	1.00	0.00	0.09	0.02	0.46
	SIC_285	-0.38	0.68	0.30	1.00	0.58	0.69	0.18	2.62
	SIC_286	-0.72	0.60	1.44	1.00	0.23	0.49	0.15	1.58
	SIC_287	-0.77	0.78	0.97	1.00	0.32	0.46	0.10	2.14
	SIC_289	-0.47	0.67	0.50	1.00	0.48	0.63	0.17	2.31
2	Intercept	-23.23	1328.78	0.00	1.00	0.99			
	PUBLIC	5.16	2044.27	0.00	1.00	1.00	173.71	0.00	.(b)
	FOREIGN	0.05	0.66	0.01	1.00	0.94	1.05	0.29	3.80
	TV_PERM	-0.48	0.62	0.59	1.00	0.44	0.62	0.18	2.11
	F_INSPEC	-0.28	0.29	0.98	1.00	0.32	0.75	0.43	1.32
	SINGLE	-0.15	1.14	0.02	1.00	0.89	0.86	0.09	8.03
	PRE_FINE	0.58	0.81	0.52	1.00	0.47	1.79	0.37	8.74
	MED	0.02	0.77	0.00	1.00	0.98	1.02	0.23	4.61
	BIG	-6.54	2168.17	0.00	1.00	1.00	0.00	0.00	.(b)
	OUT_ST	0.12	0.82	0.02	1.00	0.88	1.13	0.22	5.69
	C_BIG_PUB	15.58	6194.77	0.00	1.00	1.00	5843071.11	0.00	.(b)
	C_MED_PUB	0.21	1.40	0.02	1.00	0.88	1.23	0.08	19.18
	LN_EMP	0.14	0.26	0.27	1.00	0.60	1.14	0.69	1.91
	TN	17.90	0.68	691.88	1.00	0.00	59300554.48	15626610.85	225036368.70

[81] It was proposed that facilities owned by publicly traded companies may have fewer violations due to market pressure. An interaction variable was also added to the model to test whether facilities part of large, public firms will have fewer violations than those who are part of large, private firms, because market pressure will have a moderating effect. No effect was found. The result did not change if these interaction variables were taken out.
[82] Ibid.

Parameter Estimates									
Severe(a)		B	Std. Error	Wald	Df	Sig.	Exp(B)	95% Confidence Interval for Exp(B)	
	GA	15.09	1.35	125.02	1.00	0.00	3559207.06	252911.09	50088570.22
	NC	17.20	0.76	519.01	1.00	0.00	29547925.14	6727052.56	129786391.91
	SC	18.38	0.00	.	1.00	.	95608768.47	95608768.47	95608768.47
	SIC_281	-0.23	1.55	0.02	1.00	0.88	0.79	0.04	16.63
	SIC_282	0.95	1.24	0.59	1.00	0.44	2.59	0.23	29.20
	SIC_283	1.29	1.37	0.89	1.00	0.34	3.63	0.25	52.87
	SIC_285	1.66	1.41	1.39	1.00	0.24	5.29	0.33	83.87
	SIC_286	0.98	1.27	0.60	1.00	0.44	2.67	0.22	32.25
	SIC_287	1.73	1.43	1.47	1.00	0.23	5.63	0.34	92.36
	SIC_289	1.70	1.26	1.83	1.00	0.18	5.48	0.47	64.39
3	Intercept	-8.09	600.00	0.00	1.00	0.99			
	PUBLIC	5.79	923.07	0.00	1.00	0.99	325.42	0.00	.(b)
	FOREIGN	-0.47	0.49	0.92	1.00	0.34	0.62	0.24	1.64
	TV_PERM	0.51	0.51	0.99	1.00	0.32	1.66	0.61	4.52
	F_INSPEC	0.38	0.15	6.44	1.00	0.01	1.47	1.09	1.98
	SINGLE	-18.07	2685.80	0.00	1.00	0.99	0.00	0.00	.(b)
	PRE_FINE	-0.23	0.72	0.10	1.00	0.75	0.79	0.19	3.25
	MED	-5.00	979.01	0.00	1.00	1.00	0.01	0.00	.(b)
	BIG	0.77	0.93	0.69	1.00	0.41	2.15	0.35	13.20
	OUT_ST	-1.97	0.72	7.38	1.00	0.01	0.14	0.03	0.58
	C_BIG_PUB	1.60	1.61	0.98	1.00	0.32	4.93	0.21	116.42
	C_MED_PUB	18.41	2797.17	0.00	1.00	0.99	99052175.66	0.00	.(b)
	LN_EMP	0.51	0.23	4.96	1.00	0.03	1.67	1.06	2.62
	TN	-0.40	0.81	0.24	1.00	0.63	0.67	0.14	3.30
	GA	0.50	0.73	0.47	1.00	0.49	1.66	0.39	6.99
	NC	-0.22	0.74	0.09	1.00	0.77	0.80	0.19	3.44
	SC	-1.50	0.82	3.33	1.00	0.07	0.22	0.04	1.12
	SIC_281	-1.02	1.10	0.87	1.00	0.35	0.36	0.04	3.09
	SIC_282	-0.34	0.85	0.16	1.00	0.69	0.71	0.13	3.77
	SIC_283	-0.42	0.93	0.20	1.00	0.65	0.66	0.11	4.09
	SIC_285	-0.79	1.32	0.36	1.00	0.55	0.45	0.03	6.01
	SIC_286	1.03	0.89	1.35	1.00	0.25	2.80	0.49	15.96
	SIC_287	0.89	1.06	0.71	1.00	0.40	2.43	0.31	19.23
	SIC_289	-0.06	1.12	0.00	1.00	0.96	0.94	0.11	8.39
A	The reference category is: 0.								
B	Floating point overflow occurred while computing this statistic. Its value is therefore set to system missing.								

F. PANEL DATA ANALYSIS

Since my dataset covers four years worth of data, panel data analysis was considered, but found inadequate to model the data. Panel data analysis is a form of longitudinal analysis that allows one to analyze variations over time; for example, the effect of the change of ownership and size over time on total violation rates and number of quarters in violation. Panel data analysis is a variant of time series analysis, that has both a cross sectional component and a time series component (e.g. observes a groups of cases over multiple time periods); therefore it is often called cross sectional time-series analysis (Princeton University, Data and Statistical Services, n.d, Sec. Introduction). Panel data analysis, unlike time series analysis (which follows a single unit of observation over time) is used for datasets that have a larger number of cases (in this study facilities are the cases) than time periods being observed (this study is composed of 16 quarters). Since this study is composed of 296 cases that are followed over 16 quarters, panel data analysis was tried; both a fixed effects and random effects panel data model was run.

Fixed effects models are commonly used with panel data. Fixed effects models control for omitted variables that differ between subjects (e.g. facilities are the subject in this study) but are constant over time (Princeton University, Data and Statistical Services, n.d, Sec. Fixed Effects Regression). For example it may be the case that despite controlling for SIC category that a specific facility has unique issues due to the product that they are manufacturing. If this is the case, this effect is not controlled for by the current independent variables.

A limitation of fixed effects analysis is that it can not assess the effect of variables with little “within group” variation (each facility is considered it own group). This is because fixed effects models only look at "within group" variation and ignore "between group" variances (e.g.

differences between groups or facilities in this case). Fixed effects models consider each group's (e.g. facility) variation separately and takes the average variation of each group (facility) over time to get a coefficient. Since the estimates are based on "within group" variations you need a reasonable amount of variation of X variables over time for each case to get an efficient model (Yaffee, Fall 2003, sec. Fixed Effects Pros and Cons).

The fixed effects logistic regression panel data analysis that was run had a dependent variable that consisted of whether a facility has a violation in the quarter (1=violation, 0= no violation in a quarter).[83] All the SIC dummies, as well as the state dummies, were rejected from the model; this is because the state and SIC code for the facilities in the sample stayed the same for the entire time period of the study. When there was no variation within a group, the model kicked out the variable for lack of variation. In order to try and obtain coefficients for the SIC and states variables, it was recommended by Professor Anderson at Vanderbilt University to try a two stage estimation. Some studies estimated a linear probability model (OLS regression) with robust standard errors and then back out the fixed effects (Anderson & Siegfried, 1997). However, this technique does not work for a binary dependent variable, so it could not be used to analyze the dependent variable used in this study.

The fixed effects model, which included the rest of the variables, could not produce a Homer Lemeshow goodness of fit statistic, used to determine how the model fits the data. This is probably due to the fact there was very little variation in the dependent and independent variables that remained in the model. Any variation being picked up in the model on independent variables was primarily due to a facility being a part of a company that went through a merger or acquisition (e.g. companies that went through a merger or acquisition were primarily the ones that changed their status from public to private, foreign to domestic, or vice versa). Additionally,

[83] Logistic regression is used when you have a binary dependent variable.

the dependent variable had little variation: less than half the cases had no violation in any quarter -only 28% of the sample had more than one quarter in violation. Since the change was not occurring in a significant number of firms (roughly 14% went through a merger or acquisition) the effect of mergers and acquisitions could not be sufficiently measured. It was also attempted to use a continuous dependent variable total_viol for each quarter; however, there was again too little variation to produce an adequate model.

A random effects model was also run. Unlike fixed effects models, random effects models model both “between group” and “within in group” variation. Random effects models take into account that “some omitted variables may be constant over time but vary between cases, and others may be fixed between cases but vary over time” (Princeton University, Data and Statistical Services, n.d, Sec. Random Effects). This is done by getting a random-effects estimator based on a weighted average of fixed and between effects. A limitation of the random effects model, like the fixed effects model, is that there needs to be sufficient variation over time of the independent and dependent variable. The random effects model did not exclude any variables but provided poor predictability; the p-values of all of the independent variables were around 0.9 and no Homer Lemeshow goodness of fit statistic was provided. This is likely due to the lack of variation over time with the independent and dependent variables; in other words, as with the fixed effects model, there wasn’t enough variation over time to model.

Given the lack of sufficient variation over time in the independent and dependent variables, panel data analysis was excluded as a sufficient model to help analyze the results of the data set. As a result, it was necessary to do a cross-sectional analysis to test the hypotheses, i.e. compare the differences among subject (facility level data) without regard to changes over time.

G. STATE AND LOCAL AIR OFFICES

Tennessee

- Tennessee Department of Environmental Conservation Division of Air Pollution Control
- Memphis-Shelby County Health Department Air Pollution Control Section
- Hamilton County Air Pollution Control Bureau- (Chattanooga Area)
- Knox County Air Quality Management Division (Knoxville Area)
- Metropolitan Public Health Department of Nashville/Davidson County, Pollution Control Division

Georgia

- Georgia Department of Natural Resources Environmental Protection Division- Air Protection Branch

North Carolina

- Department of Environment and Natural Resources, Division of Air Quality
- Western North Carolina Regional Air Quality Agency, Asheville North Carolina
- Forsyth County Environmental Affairs Department, Winston Salem, North Carolina
- Mecklenburg County Air Quality (Charlotte Area)

South Carolina

- South Carolina Department of Health and Environmental Control, Air Quality

Florida

- Florida Department of Environmental Protection, Air Compliance Assurance Section. The Department "coordinates the statewide air compliance and enforcement activities.

The Air Compliance Assurance (ACA) Section promotes air compliance through the department's district offices and the approved local program offices.: DEP website

There are also eight local county program offices:

- Orange County Environmental Protection Division
- Duval County Environmental Resources Management Department
- Broward County Department of Planning & Environmental Protection
- Miami-Dade County Department of Environmental Resources Management
- Palm Beach County Health Department
- Hillsborough County Environmental Protection Commission
- Pinellas County Department of Environmental Management
- Sarasota County Natural Resources Department

REFERENCES

Alexander, C. R. & Cohen, M. A. (1996). New evidence on the origins of corporate crime. *Managerial and Decision Economics*, 17(4), 421-435.

Alexander, C.R. & Cohen, M.A. (1999). Why do corporations become criminals? Ownership, hidden actions, and crime as an agency cost. *Journal of Corporate Finance*, 5(1), 1-34.

Anderson, K. H. & Siegfried, J. J. (1997). Gender differences in rating the teaching of economics. *Eastern Economic Journal,* 23(3), 347-357.

Becker, G. S. (1968). Crime and punishment: an economic approach. *Journal of Political Economy*, 76(2), 169-217.

Burby, R. J. & Paterson, R. G. (1993). Improving compliance with state environmental regulations. *Journal of Policy Analysis and Management*, 12(4), 753-772.

Clark, M. (2004, November). Future directions for environmental economics research support, *Association of Environmental and Resource Economics Newsletter*, 24(2), 20-24.

Clean Air Act, Air quality and emissions limitations § 101-131, 160-193, 42 U.S.Code § 7401-7515 (2006 ed.)

Clean Air Act, General § 301-328, 42 U.S.Code § 7601-7627 (2006 ed.)

Clean Air Act, Permits § 501-507, 42 U.S.Code § 7661-7661f (2006 ed.)

Cohen, M. A. (1986). The costs and benefits of oil spill prevention and enforcement. *Journal of Environmental Economics and Management*, 13(2), 167-88.

Cohen, M. A. (1997). Firm response to environmental regulation and environmental pressures. *Managerial and Decision Economics*, 18(6), 417-420.

Cohen, M. A. (1999). Monitoring and enforcement policy. *International Yearbook of Environmental Resource Economics*. Ed. Tietenberg, T & Folmer, H. (vol. 3, pp.44-106). UK: Edward Elgar publishers.

Directory of Corporate Affiliations: Who Owns Whom (2003). Skokie, Illinois: National Register Publishing Company.

Drukker, D. (February 2000, revised July 2007). Stata Data Analysis and Statistical Software, FAQ's. My raw count data contains evidence of both overdispersion and "excess zeros." Retrieved March 25, 2008 from http://www.stata.com/support/faqs/stat/nbreg.html

Dun & Bradstreet (n.d) D&B's million dollar database. Retrieved May 29, 2009 from http://www.dnbmdd.com/mddi/family_tree.aspx

Dun & Bradstreet (2007) Company Information for Bridgestone America Inc. Retrieved March 12, 2007 from Dun & Bradstreet Million Dollar Database via Walker Management Library, Owen School of Management, Vanderbilt University.

Environmental Protection Agency, Office of Inspector General (September, 2000). *EPA's oversight of state stack testing programs.* (Report Number 2000-P-00019). Philadelphia, PA. Retrieved from http://epa.gov/oig/reports/2000/stack.pdf

Environmental Protection Agency (April, 2001). *Clean Air Act stationary source compliance monitoring strategy*. Washington, D.C. Retrieved from www.epa.gov/compliance/resources/policies/monitoring/cmspolicy.pdf

Environmental Protection Agency (2003). Enforcement & Compliance History Online (ECHO). Search help. Facility characteristics. Retrieved October 8, 2003 from http://www.epa-echo.gov/echo/help_all_programs.html#nonattinment

Environmental Protection Agency (2005a). Enforcement & Compliance History Online (ECHO). About the data. Retrieved January 20, 2005 from http://www.epa-echo.gov/echo/about_data.html#afs

Environmental Protection Agency (2005b). Enforcement & Compliance History Online (ECHO). Detailed facility report: data dictionary. Retrieved January 20, 2005 from http://www.epa-echo.gov/echo/dfr_data_dictionary.html#fc

Environmental Protection Agency (2005c). Enforcement & Compliance History Online (ECHO). Known data problems. Air. Retrieved January 20, 2005, from http://www.epa-echo.gov/echo/known_data_problems.html#caa

Environmental Protection Agency (2008). Enforcement & Compliance History Online (ECHO). Known data problems. Hazardous Waste. Retrieved October 15, 2008, from http://www.epa-echo.gov/echo/known_data_problems.html#rcra

Environmental Protection Agency. (September, 2008). 2001 National Hazardous Waste Biennial Report. The data presented in the 2001 National Biennial Report**.** Retrieved April 15, 2009 from http://www.epa.gov/epawaste/inforesources/data/brs01/index.htm

Epple, D. & Vischer M. (1984). Environmental pollution modeling occurrence, detection, and deterrence. *Journal of Law and Economics*, 27(1), 29-60.

Florida Department of Environmental Protection (2003). Chapter 3: Enforcement options. In *Enforcement Manual*. Retrieved December 15, 2005 from http://www.dep.state.fl.us/legal/Enforcement/enforcement.htm

Florida Department of Environmental Protection (2003). Guidelines for characterizing air violations for stationary point sources of air pollution. In *Enforcement Manual* (Appendix Civil Policy Memo). Retrieved December 15, 2005, from http://www.dep.state.fl.us/legal/Enforcement/appendix/guidelines/air%20-%20stationary.pdf

Florida Department of Environmental Protection (2004). General air penalty guidelines. In *Enforcement Manual (11/26/07 document removed from manual).* Retrieved December 15, 2005 from http://www.dep.state.fl.us/legal/Enforcement/appendix/guidelines/air-general.pdf

Grant, D., Bergesen, A. J., & Jones, A.W (2002). Organizational size and pollution: The case of the U.S. chemical industry. *American Sociological Review*, 67(3), 389-407.

Grant, D., Jones, A. W. & Trautner, M. N. (2004). Do facilities with distant headquarters pollute more? How civic engagement conditions the environmental performance of absentee managed plants. *Social Forces,* 83(1), 189-214.

Grau, M. V., & Groves, T. (1997). The oil spill process: The effect of Coast Guard monitoring on oil spill. *Environmental and Resource Economics*, 10(4), 315-339.

Gray, W. B., & Deily, M.E. (1991). Enforcement of pollution regulations in a declining industry. *Journal of Environmental Economics and Management*, 21(3), 260-274.

Gray, W.B., & Deily, M.E. (1996). Compliance and enforcement: Air pollution regulation in the U.S. steel industry. *Journal of Environmental Economics and Management*, 31(1), 96-111.

Gray, W. B., & Shadbegian, R. J. (2005). When and why do plants comply? Paper mills in the 1980s. *Law and Policy*, 27(2), 238-261.

Gray, W. B., & Shadbegian, R. J. (2007). The environmental performance of polluting plants: A spatial analysis. *Journal of Regional Science*, 47(1), 63 – 84.

Gunningham, N., Thornton, D. & Kagan, R. A. (2005). Motivating management: Corporate compliance in environmental protection. *Law & Policy*, 27(2), 289-316.

King, A. A. & Shaver J. M. (2001). Are aliens green? Assessing foreign establishments. *Strategic Management Journal*, 22(11), 1069-1085. DOI: 10.1002/smj.201

Konar, S., & Cohen, M. A. (March 2000). Why do firms pollute (and reduce) toxic emissions? *Working Paper*. Retrieved September 21, 2006 http://sitemason.vanderbilt.edu/files/fnddVm/why%20do%20firms%20pollute.pdf. Also available from Social Science Research Network eLibrary Working Paper Series found at http://ssrn.com/abstract=922491

Laplante, B., & Rilstone, P. (1996). Environmental inspections and emissions of the pulp and paper industry in Quebec. *Journal of Environmental Economics and Management*, 31(1), 19-36.

Long, S. J. (1997). *Regression models for categorical and limited dependent variables*. Thousand Oaks, CA: Sage Publication.

Long, S. J. & Freese, J. (2006). *Regression Models for Categorical Dependent Variables Using Stata.* 2nd Ed. College Station, TX: Stata Press.

Organisation for Economic Co-Operation and Development (December, 2004). Economic aspects of environmental compliance assurance, proceedings from the OECD global forum on sustainable development. Paris, France. Retrieved from http://www.oecd.org/dataoecd/15/16/37719119.pdf

Nadeau, L. W. (1997). EPA effectiveness at reducing the duration of plant level non-compliance. *Journal of Environmental Economics and Management*, 34(1), 54-78.

Magat, W. A., & Viscusi, K. W. (1990). Effectiveness of the EPA's regulatory enforcement: The case of industrial effluent standards. *The Journal of Law and Economics,* 33(2), 331-360.

McCabe, M. W. (December 18, 1996). *Validation of Air Enforcement Data Reported to EPA by Pennsylvania, Appendix B: EPA's response to draft report.* U.S. Environmental Protection Agency. Retrieved May 20, 2008 from http://www.epa.gov/oig/reports/1997/pennapp2.htm

Pampel, F. C. (2000). *Logistic regression: A primer*. Thousand Oaks, California: Sage Publication.

Princeton University, Data and Statistical Services (2007). Panel Data. Retrieved May 15, 2008 from http://dss.princeton.edu/online_help/analysis/panel.htm#models

Protection of the Environment, Air Programs, 40 C.F.R. §50-88.17 (2008 ed.).

Schaeffer, E. (1998). Memo: Issuance of policy on timely and appropriate enforcement response to high priority violations. Washington D.C. U.S. Environmental Protection Agency. Retrieved December 2005 from http://www.epa.gov/compliance/resources/policies/civil/caa/stationary/issue-ta-rpt.pdf

Simonoff, J.S. (2003). *Analyzing Categorical Data*. New York: Springer-Verlag.

Stafford, S. L. (2002). The effect of punishment on firm compliance with hazardous waste regulations. *Journal of Environmental Economics and Management*, 44(2), 290-308.

UCLA: Academic Technology Services, Statistical Consulting Group (n.d.) Ordered logistic regression, Retrieved November 24, 2008 from http://www.ats.ucla.edu/stat/SPSS/dae/ologit.htm

U.S. General Accounting Office (June, 2000). *More consistency needed among EPA regions in approach to enforcement.* (Report number GAO/RCED-00-108). Washington, D.C. Retrieved from http://www.gao.gov/new.items/rc00108.pdf

U.S. Government Accountability Office (June, 2006). *Environmental compliance and enforcement: EPA's effort to improve and make more consistent its compliance and enforcement activities.* (Report Number GAO-06-840T). Washington D.C. Retrieved from http://www.gao.gov/new.items/d06840t.pdf

Vandenbergh, M. P. (2005). The private life of public law. *Columbia Law Review*, 105(7), 2029-2076.

Weber, J.M. & Crew, R.E. (2000). Deterrence theory and marine oil spills: Do coast guard civil penalties deter pollution? *Journal of Environmental Management*, 58(3), 161–168. doi:10.1006/jema.1999.0326,

Yaffee, Robert, (Fall, 2003). A primer for panel data analysis. *Connect Information technology at NYU*. Services Retrieved October 2007 from http://www.nyu.edu/its/pubs/connect/fall03/yaffee_primer.html