Choices in Regulatory Program Design and Enforcement

Christopher Carrigan
George Washington University

Elise Harrington
University of Pennsylvania

Research Paper Prepared for the
Penn Program on Regulation's
Best-in-Class Regulator Initiative

June, 2015

Choices in Regulatory Program Design and Enforcement

Christopher Carrigan
Elise Harrington

Table of Contents

Choices in Regulatory Program Design and Enforcement

Christopher Carrigan
George Washington University

Elise Harrington
University of Pennsylvania

Executive Summary

Historically, regulators have relied on a select number of strategies in designing and enforcing rules to carry out their statutory authority. Yet, in reality, a wide range of options exists which regulators can use to tailor their regulatory programs to their particular circumstances.

Solving regulatory problems involves considering how to design the regulatory program as well as how to enforce that program once it is in place. This paper considers the substantial diversity that exists both in the regulatory instruments that can be used to achieve regulatory goals and the various ways that regulators might choose to ensure regulatory entities respond to regulatory demands once they are in place.

Regulatory Instruments

In practice, the vast majority of regulatory problems are addressed using either means-based regulation (which specifies the particular technologies to be used by regulated entities to satisfy the regulatory requirements) or performance regulation (which specifies the end goal in terms of a regulatory target or outcome without stipulating how the firm should achieve it). Yet, the regulator also has at its disposal a number of other instrument options:

- *Market-based mechanisms* use market signals rather than specific commands to achieve regulatory goals. Options include "corrective" taxes, fees and charges, tradable permit systems, and bonds. Market-based instruments all share the feature that they allow regulated entities the freedom to choose not only their method for achieving regulatory goals but also the extent to which they actually attain these same goals based on their costs of complying.

- *Management-based regulation* mandates that regulated entities engage in a planning process to identify hazards and describe how they will minimize and respond to the associated risks. Although management-based approaches attempt to impact firm choices before a problem is identified, they allow firms flexibility in designing their plans and sometimes do not even mandate that regulated entities implement those same plans.

- *Mandated information disclosure* requires firms to gather and provide information about their operations but does not oblige firms to respond to the data they make

publicly available. Rather, mandated information disclosure attempts to use social pressure, either through public scrutiny or changes in consumer buying behavior to encourage firms improve their regulatory performance.

- *Voluntary programs* do not compel firms to comply with any specific requirements. Instead, these programs, which can be developed by the regulator, employ rewards such as certifications, recognition, awards, educational resources, and exemptions from mandatory regulatory requirements to encourage compliance. As a type of voluntary program, self-regulation does not involve the regulator specifically but often operates to dissuade that regulatory agency from imposing stricter requirements.

Regulatory Enforcement

Just as regulators have a wide variety of instruments at their disposal to craft regulatory programs, they also have a great deal of flexibility in designing their enforcement strategies. In fact, developing an overarching enforcement program requires decisions along a number of dimensions, each of which afford the regulator choices in how they interact with their regulated entities "on the ground."

- The level of *stringency* (or the degree to which requirements mandate "significant" changes in firm behavior) that regulated firms face is determined not only by the regulatory requirements themselves, but also by how they are implemented in practice. Inspectors may rationally choose different levels of stringency over time and across geographies given the diversity in how a regulatory problem manifests itself along these dimensions.

- Given limited enforcement resources, regulators can use *targeting*, focusing their enforcement efforts on those firms with poor compliance histories by imposing larger fines and engaging in more frequent inspections. In doing so, firms that are compliant have an incentive to remain so to avoid being targeted while non-compliant firms are encouraged to improve their regulatory performance to escape the targeted group.

- All regulators seek to achieve *deterrence*, dissuading violations through the threat of punishment, but the types of deterrence emphasized by the regulator will have implications for how it designs its enforcement programs. Whereas absolute deterrence aims to prevent all violations, optimal deterrence seeks to prevent only inefficient breaches. Contrasting specific deterrence, which looks prevent repeat offenses, general deterrence uses broad threats or actions against one firm to dissuade others from noncompliance.

- A regulator also has options when it comes to its regulatory *style*. While regulatory styles are perhaps best thought of along a continuum, the research literature distinguishes between specific types, any of which can be appropriate depending on the specific characteristics of the associated regulatory environment.

- *Legalistic enforcement* is a deterrence-based approach which relies predominately on strict, rule-based application of notices, fines, and other mechanisms to sanction infractions and deter future violations.

- *Accommodative enforcement* seeks to achieve compliance (not deterrence) by cooperating with and advising regulated entities through informal mechanisms such as education, negotiation, sympathy, and persuasion.

- *Flexible or responsive enforcement* (which is sometimes considered synonymous with accommodative enforcement) approaches its interactions with firms by employing a "tit-for-tat" strategy, in which inspectors cooperate when firms do likewise but impose sanctions – often in an escalating fashion – when the regulated entities violate the rules.

The breadth of regulatory instruments and enforcement strategies covered in this paper does not lend itself easily to broad generalizations. However, the large literature examining these regulatory choices yields two general insights which underscore the advantages that those regulators which have a working knowledge of the various possibilities are afforded in designing and enforcing their regulatory programs.

The first is insight is that although all regulatory approaches can achieve some of the goals that a regulator may have for its regulatory programs, no one approach can achieve them all simultaneously. Possible criteria upon which a regulator may base its approach include its ability to reduce risk, cost-effectiveness, relative efficiency, flexibility, administrative feasibility, propensity to promote equity, and ability to mitigate the potential for regulatory capture. For example, the evidence reviewed suggests that market-based instruments offer the potential for achieving regulatory goals more cost effectively than other approaches, but they may so by imposing a greater administrative burden on both the regulator by requiring them to craft more intricate requirements and the regulated entities which might need to implement more complex monitoring equipment as a result. Similarly, although targeted enforcement can achieve compliance more cost-effectively, the discretion such an approach requires of the regulator opens it up to greater interest group pressure.

The second insight is that regulatory approaches can be fruitfully used in tandem to encourage better regulatory performance among regulated entities. Perhaps the clearest illustration of this insight can be found in responsive regulation, which blends legalistic and accommodative enforcement approaches to attempt to realize the advantages of supportively interacting with regulated firms while maintaining the ability to sanction those that attempt to take advantage of the regulator's willingness to cooperate. Research on regulatory instruments illustrates similar advantages in combining approaches. Some of the most definitive evidence that alternative approaches including voluntary programs can be successful in achieving regulatory goals has come when these instruments were paired with more traditional mechanisms. For instance, research considering the efficacy of mandated information disclosure has shown that it can affect firms' regulatory performance when used to supplement traditional regulatory mechanisms by making infractions public.

These insights point to a need for the regulator to determine, prior to selecting a regulatory approach, which criteria among the possible set for evaluating regulatory instruments and enforcement methods is most appropriate for its regulatory context. The findings simultaneously underscore the need for the regulator to consider a wide range of potential options before settling on one or more specific approaches. In so doing, the regulatory agency is not only much more likely to choose the option most suited to remedy the problem at hand, but that agency is also more apt to be able to realize the potential synergies available when regulatory instruments and enforcement techniques are used in combination.

Choices in Regulatory Program Design and Enforcement

Christopher Carrigan
Elise Harrington

Introduction

Despite the fact that government legislation and policies define regulators' missions and constrain their choices, agencies still enjoy substantial discretion with respect to how they design and implement regulatory programs. In much the same way that they face choices in communicating policies to their regulated communities – using formal rulemaking, informal guidance documents, and face-to-face interactions, among other possibilities – regulatory agencies have a broad set of tools and methods at their disposal to implement the regulatory authority they receive from lawmakers. In this paper, we review the breadth of approaches that regulatory agencies can use to design regulatory programs and the range of methods used to enforce those same programs.

To do so, we first describe the variety of regulatory instruments at the agency's disposal in designing the rules that underpin their regulatory programs ranging from precisely defined prescriptive standards to loosely defined voluntary guidelines. In addition to time-honored mechanisms that require firms to implement certain technologies or allow firms to choose their own technologies but specify targets, we also review what is known about a multitude of alternative instruments that often provide regulated firms with greater degrees of flexibility in achieving regulatory objectives. Among these alternatives, we examine and assess the available empirical evidence ascertaining the effectiveness of market-based instruments – including taxes and fees, tradable permits, and bonds – as well as management-based or enforced self-regulation and information disclosure requirements. We further discuss a substantial body of research studies considering how voluntary regulatory programs as well as self-regulatory approaches can achieve social goals.

Following our discussion of the mechanisms available to regulatory agencies to design their programs, we analyze the approaches that those institutions may take in enforcing the standards they set. As with the design of regulatory instruments, we highlight the substantial diversity that exists with respect to how regulators choose to enforce their rules. We begin by describing concepts important to developing an enforcement program including stringency, targeting, and deterrence, distinguishing between specific and general deterrence and how enforcement might affect the achievement of either conception of deterrence. We also consider differences among regulators with respect to their regulatory style. Beginning with what is broadly termed "legalistic" regulatory enforcement – relying predominately on strict, rule-based application of fines to sanction infractions – we contrast this approach with a cooperative or "accommodative" approach, which emphasizes problem solving through education, persuasion, and negotiation. Finally, we consider the "responsive" approach to regulatory enforcement, which addresses the task of implementing regulations by combining accommodative tactics with legalistic elements to yield a flexible regulatory enforcement environment. Throughout, we discuss when particular approaches may be more appropriate given the regulatory context.

To aid the reader in sifting through the volume of material presented in this paper, we augment our discussion with a series of case studies, illustrating how the regulatory instruments and enforcement styles we discuss have been used by regulators in practice. Further, we begin our discussion in the next section by defining a number of terms – including cost-effectiveness, equity, and flexibility – that reflect what the regulator might wish to accomplish with its regulatory programs. One of the broad themes that emerges from the discussion is that, because they run the gamut from more directive and adversarial to more flexible and cooperative, regulatory instruments and enforcement strategies can vary widely with respect to what outcomes they can achieve. Given that multiple criteria may be important to a regulatory agency when choosing a particular approach, we define these possible measures to cover the set of criteria that a regulator might want to consider and to provide clarity in how we use these terms in the context of this discussion. Our hope is that the following sections will convey the enormous breath of choices available to regulatory agencies in designing and implementing their programs while highlighting some of the common themes that emerge from the large body of research studying these regulatory instruments and enforcement strategies.

I. Defining the Set of Evaluative Criteria

In this section, we define a series of criteria upon which any particular regulatory instrument or enforcement strategy can be evaluated, either prospectively or retrospectively. Any evaluation, of course, of course, seeks to discern "the extent to which an activity achieves its intended objectives" (Salamon 2002:23). While the terms defined in this section cannot, by themselves, determine for an agency what their objective or objectives should be, one purpose of this discussion is to present a series of possibilities that scholars and commenters have employed to decide whether a regulatory instrument or enforcement regime is effective.

As we will show in the following sections, given the wide variety of criteria available, "one size does not fit all" (Richards 2000:226) when it comes to regulatory instruments and enforcement strategies. Rather, tradeoffs exist at every level. Nevertheless, explicitly recognizing these tradeoffs can help a regulatory agency choose tools and enforcement approaches that best address the most critical issues in its particular regulatory space. While we do not claim that the criteria considered here are exhaustive, they represent the set of metrics most widely considered by regulatory scholars in evaluating regulatory instruments and enforcement strategies.

A. *Ability to Reduce Risk*

Considering how alternatives vary with respect to their relative abilities to minimize risks is fundamental to choosing among alternative strategies for implementing and enforcing regulatory programs. At one end of the spectrum, the precautionary principle reflects an approach most closely related to the well-known saying "better safe than sorry." Application of the precautionary principle ranges from not regulating without evidence of harm to "building a margin of safety into all decision making" to "even a small amount of doubt" is enough to prevent an activity from occurring (Sunstein 2003:5). Variations of the precautionary principle have been employed in regulatory decision making since the 1970s in countries all over the

world. Such precautionary measures aim to prevent high-risk, low-probability events given ambiguous scientific evidence or a lack of scientific consensus.

Alternatives to the precautionary principle require risk-assessments based on the probability an event will occur and the resulting consequences if that event does occur (Krebs 2011). A risk-based approach demands significant agency resources and expertise to implement properly (Rothstein et al. 2006). Moreover, risk assessments can be difficult to communicate to the public (and thus implement) in comparison to the precautionary principle partly because of the difficulties people have in interpreting probabilities as noted in the behavioral economics and psychology literatures. Cognitive biases contributing to these difficulties include loss aversion (people are more sensitive to losses than they are to gains); the myth of benevolent nature (people believe that health and safety risks are generally caused by human intervention and not by nature); availability bias (people focus on risks that readily come to mind); probability neglect (people focus on the outcome rather than the probability that an outcome will occur); and system neglect (people neglect the whole picture, focusing only on a subset of the system) (Sunstein 2003).

Risk-based regulation is "an evidence based means of targeting the use of resources and of prioritizing attention to the highest risks" (Black & Baldwin 2010:181). Before undertaking risk-based regulation, regulators must consider scientific evidence, economics, and public acceptability or perception of risk in the decision-making process, as well the extent to which a regulator or government believes it should intervene and regulate risks on behalf of the public (Krebs 2011:4845). While risk-based regulation promises a rational approach to regulation, experts do not all agree on its virtues, nor even on what the concept means with any precision.

Given that the only sure-fire way to eliminate risk would be to eliminate the underlying economic activity (Carrigan & Coglianese 2012), the very act of regulating suggests that the agency will have to consider risk in making decisions. However, regulatory approaches, both in the instruments used and the enforcement strategies implemented, vary based on the degree to which they mitigate risk. They can also differ with respect to who identifies and acts to reduce those risks. For example, "where monitoring is difficult or costly, regulators have sometimes imposed rules requiring firms to identify risks posed by their own operations and develop their own set of internal policies and monitoring procedures" (Coglianese & Kagan 2008:xx). While this approach to risk management places the burden on the regulated entities, other tools can task the market or the regulatory agency with identifying and responding to risks.

B. Cost-effectiveness

A variety of costs are imposed not only on regulated firms but also on the public and the regulatory agency as well when regulatory programs are introduced. For example, for the regulator, gathering the information required to effectively implement regulatory programs can be an important implicit cost to consider in choosing among regulatory instruments. Cost-effectiveness refers to the ability to attain one unit or a given level of benefit for the least cost (Leman 2002). Cost effectiveness favors the option that minimizes cost when the level of benefits is held constant. When considering the cost effectiveness of a program, a cost-effectiveness ratio is often used, which divides the total cost by the units of effectiveness since

programs may deliver different benefit levels. One concern with using such ratios is they can mask differences in scale between different strategies. Even if one regulatory approach is not as cost effective as another, the former may still deliver a great deal more benefits, which is not something that can be ascertained by simply comparing the ratios.

C. Efficiency

Although, in practice, regulators use the term efficiency in various ways (Finkel et al. 2015), in economics, it has two specific and related meanings. The first, referred to as Pareto efficiency, considers a state efficient when no one can be made better off without someone else being made worse off (Perloff 2001). The second – which is the efficiency criterion underpinning the decision when benefit-cost analysis is used to evaluate regulatory alternatives – is labeled Kaldor-Hicks or potential Pareto efficiency. If a contemplated regulatory action is Kaldor-Hicks efficient, it means that while at least one person is made better off and at least one is made worse off, in theory, the winners can compensate the losers. Thus, a policy is Kaldor-Hicks efficient if the benefits *exceed* the costs. In choosing among regulatory programs, applying this efficiency criterion dictates that the regulatory agency should choose the option that maximizes net benefits, or the difference between benefits and costs.

Cost-effectiveness analysis is sometimes used as a substitute for benefit-cost analysis to choose among alternatives because the former eases the analytical and computational demands associated with the study. In addition to not requiring the benefits to be converted to monetary units (i.e. the benefit can be left in non-monetized units, like lives saved), cost-effectiveness also considers only one benefit rather than all benefits. Even so, the very fact that it reduces the computational burden also means that, cost-effectiveness analysis may lead to inefficient decisions. One regulatory approach may be more cost effective than another while, simultaneously, having lower net benefits. Coupled with only considering one benefit, cost-effectiveness analysis makes it difficult to compare programs which have different-sized benefits. Unlike benefit-cost analysis, cost-effectiveness analysis will lead to decisions that are inefficient in cases where costs exceed the benefits even for the most cost-effective policy (which the analyst will not be able to ascertain through the cost-effectiveness analysis alone).

D. Flexibility

The degree of flexibility is a factor both in designing and enforcing regulation. Regulatory flexibility refers to the extent to which regulated entities can make individual choices about how to meet the objectives of the regulation or comply with its terms. All things being equal, more flexible regulations are preferred because they allow for regulated entities to comply using the method at their disposal that is the least costly. Thus, the flexibility criterion is closely connected with cost-effectiveness.

Selection and design of regulatory instruments includes the specification of a "command" (i.e. what the regulator determines is either allowable or not allowable) that can be flexible with respect to its stringency, structure, or specificity (Bennear & Coglianese 2012). Stringency refers to the required abatement amount or to the extent of the activities that are required by the regulatory goal. Structure, in contrast, signifies how a regulated entity must respond in order to

achieve the mandated goal, a topic which will be discussed further in part three of this paper. For example, some commands specify what actions the regulated entities are required to take while others only specify the end states that they entities need to achieve. Finally, specificity refers to degree of accuracy that the regulation requires. Regulatory requirements can be designed to mandate a specific technology or the "installation of reasonably available control technology." Moreover, specificity can vary along the stringency dimension by requiring a specified amount reduced or a "reasonable level" of reduction. As we will describe in more detail in the next section, some regulatory instruments, such as performance standards, information disclosure, and management-based regulation, allow for a greater degree of flexibility on one or all of these dimensions – stringency, structure, or specificity – relative to other techniques.

E. Manageability or Administrative Feasibility

Like flexibility, minimizing the administrative burden that the program imposes is related to cost-effectiveness. But administrative burden specifically highlights a subset of costs that relate to the firm's efforts to implement the requirements as well as those imposed on the regulatory agency in designing and executing the program. Given that selecting and implementing regulations can be costly, regulators might consider selecting the tool or enforcement strategy "that minimizes both the regulated entities' compliance costs and the government's costs of selecting and implementing a standard that achieves a given regulatory objective" (Coglianese & Lazer 2003:704). In assessing the administrative burden imposed by a regulatory program, the analyst should consider how up-front administrative costs – which can receive a disproportionately large proportion of the attention – will impact long-run costs for the regulatory program, as some costs incurred in the short-term may reduce overall costs in the long-run (Driesen 1998).

F. Predictability

Sometimes even more than the actual cost of implementing the regulatory requirements (its administrative burden), uncertainty about the nature of future regulatory programs can present a major obstacle for regulated firms as they plan their business processes (Dixit & Pindyk 1994). Moreover, ambiguity with respect to how requirements will be interpreted and applied by the regulator once they are promulgated can impose additional costs on firms as well. Yet, this same ambiguity can also place burdens on regulatory inspectors, by asking them to discern whether specific actions by firms meet the requirements of the mandate. By limiting the amount of discretion in interpreting the requirements, some instruments lend themselves more easily to providing clarity and predictability for both the regulatory agency and its regulated firms.

G. Equity

As an evaluation criterion, equity has two related meanings, both focused on how the costs and benefits are spread across society. Equity's first meaning considers how fairly costs and benefits are distributed. If one set of individuals or groups is disproportionately harmed by a regulation, or others are disproportionately benefited, the result might be unacceptable on equity grounds. For example, some regulatory programs may specifically burden small firms as they are

less equipped to implement sophisticated requirements. Equity can also incorporate the extent to which the proposed regulatory program involves a redistributive component. In this second meaning, the agency would actively look to structure the regulatory program such that it redistributes resources to those who lack them currently, rather than focusing primarily on avoiding disproportionately affecting one group. Under this second meaning, regulation serves to address past distribution inequalities in order to better provide equal opportunities throughout society (Salamon 2002).

When benefit-cost analysis is used to select an efficient regulatory alternative (i.e. a policy in which the benefits exceed the costs), it simply means that the policy winners could fully compensate the losers so that none are made worse off. Since such compensation is rarely incorporated into a regulatory proposal in practice, efficient policies can be unequitable. As a result, equity may be a consideration in designing regulatory instruments and enforcement regimes in addition to or instead of efficiency.

H. Ability to Mitigate Capture

Regulatory capture refers to a situation or process in which a regulatory agency regulates for the benefit of its regulated entities (which are firms in most instances) rather than the public interest (Carpenter & Moss 2014). An agency can be captured by industry over a period of time if the regulator begins to sympathize with its regulated firms and, in doing so, sets aside its mission to achieve public objectives (Bernstein 1955; Lowi 1969). An example can be found in Samuel Huntington's (1952) description of the long-since disbanded U.S. Interstate Commerce Commission. Although its initial purpose was to regulate railroads for the benefit of farmers and shippers, the Commission soon began to see its interests as more aligned with the railroads themselves, such that each defended the other against competition from other government agencies and from other forms of transportation.

Others have argued, somewhat counterintuitively, that the industry itself can lobby for the creation of a regulatory regime, in which case, the regulatory agency is captured as soon as it is established (Peltzman 1976; Posner 1974; Stigler 1971). Known as the economic theory of regulation (Baldwin et al. 2012; Posner 1974), this view of capture highlights how various regulations, including rate or price controls and licensing requirements, can advantage incumbent firms by raising prices and creating barriers to entry for would-be competitors (Stigler 1971).[1]

The concept of regulatory capture can be further broadened to include any systemic influence special interest groups exert over government regulation. Applying this broader conception, evidence that an agency is captured can manifest itself in many ways including, "identification with the industry, sympathy with the particular problems that regulated firms confront in meeting standards, and absence of toughness" (Makkai & Braithwaite 1992:61). For example, experts associate regulatory capture with "revolving doors" – the practice whereby personnel may move back and forth between the regulatory agency and the regulated firms (Bó 2006). Those who come from industry to a regulatory agency may be captured because they identify with the industry. Those who hope to go out the revolving door from the agency to the

[1] For an in-depth discussion of the development of the economic theory of regulation, see Carrigan and Coglianese (2015).

industry may be sympathetic to the industry and, thus, less inclined to impose strict requirements on the regulated firms.

When the concept of capture is used as an evaluative criterion, regulatory instruments or enforcement strategies fare better when they position the regulatory agency to more easily resist becoming captured by its regulated firms. As we will describe in greater depth in our discussion of enforcement, regulators that implement strategies based on "responsiveness" or "regulatory reasonableness," may be more sympathetic to the industry's ability to comply which could potentially place them at greater peril of becoming captured. The possibility of regulatory capture has been cited as important in choosing among regulatory instruments as well, as regulatory tools such as self-regulation or voluntary regulation may be more subject to industry opportunism (Coglianese & Kagan 2008; Howard et al. 2000; King & Lenox 2000).

II. Designing Regulatory Programs

In the discussion that follows, we outline and compare the different policy instruments available to regulators. Approaches to designing regulatory programs range from traditional methods including technology and performance standards to alternative mechanisms such as market-based instruments, management-based regulation, information disclosure, and voluntary programs. As described in Table 1 which previews our discussion, different regulatory designs will be more or less appropriate depending on the situation.

In choosing among regulatory alternatives, it is useful to start by identifying what is to be regulated. Issues in regulatory design are important to economic regulation, which typically refers to regulation that sets prices or imposes rules surrounding market entry in cases where market conditions reveal the existence of or potential for a natural monopoly. A natural monopoly refers to the case where having one firm supply the entire market for a product is more cost effective than having more than one firm.

Still, questions of regulatory design are equally as important to social regulation, which often involves correcting externalities and so is associated with health and safety as well as environmental regulation. For example, the existence of negative externalities (which economists define as costs imposed on others that are not paid for by those generating them) in the form of pollution are important drivers of regulation of industrial activity. As each type – economic and social regulation – has its own distinctive mechanisms, the regulatory agency needs to consider its regulatory domain when considering possible designs. At least eight governance components of market activity exist that can be regulated: entry, exit, behavior, costs, content, preferences, technology, and performance (Levi-Faur 2011).

The selection and design of regulatory instruments "must confront a fundamental issue of how tight controls should be in seeking consistency versus how much discretion should be granted in promoting flexibility and innovation" (May 2003:382). Different regulatory mechanisms afford the targets of regulation different levels of flexibility, which correspond to requirements for large or small amounts of information about the regulated entities as well as the

Table 1: Describing the Breadth of Instruments Available to Regulators

Regulation Type	Description	Primary Benefit	Primary Cost
Means-Based (Technology, Design or Specification)	Specifies particular technologies which can be used to satisfy requirement	Provides predictability and clarity for regulator and regulated entities	Can "freeze" regulatory technology and inhibit firm's incentives to innovate
Performance-Based	Specifies end goal without identifying how firm should achieve it	Encourages firms to find cheaper ways to achieve regulatory goals	When applied uniformly, firms do not have incentives to exceed regulatory goals
Market-Based (Taxes, Fees, Charges, Tradable Permits or Bonds)	Uses market signals (not commands) including prices and quantities to alter behavior	Concentrates abatement efforts with those firms that can do it most cost-effectively	Political resistance and potential for increased rule and enforcement complexity
Management-Based	Mandates firm planning to identify, minimize, and respond to hazards	Allows firms flexibility to design plans around their specific operations	Difficult for regulator to know if firms engaging in planning process responsibly
Mandated Information Disclosure	Requires firms to publicly disclose information about operations	Can be implemented at little cost to firms and can facilitate competition	Relies on consumers to read and be able to comprehend and respond to information
Voluntary and Self-Regulation	Rewards socially desirable behavior but does not compel firms to comply	Reduces regulator's enforcement costs and provides regulatory flexibility to firms	Can intensify residual risk by falsely conveying impression firms have controlled risks

broader regulatory environment. While market-based or other alternative mechanisms may be deemed more cost effective than traditional policies, the conditions for their suitability may not always apply and, therefore, other traditional mechanisms are more frequently implemented in practice (Keohane et al. 1998; Stavins 1998).

A. Technology, Design, and Performance Standards

Traditional regulatory instruments can be split into two broad categories: means-based regulation, also referred to as technology, design, or specification regulation, and performance regulation (Carrigan & Coglianese 2011). Means-based regulatory standards specify which technologies may be used to satisfy the regulatory requirement. For example, technology standards used to reduce pollution specify both how abatement will be achieved and the total amount of pollution allowed (Richards 2000). Moreover, such standards often specify how an abatement object must be designed and constructed. The availability of accurate information on available technology is critical for means-based regulatory standards to be feasible.

In contrast, performance standards specify the end goal in terms of targets and measurements, such as rates, concentrations, and quantities, but allow firms to achieve that goal using their method of choice (Coglianese et al. 2003). Performance regulation requires either the attainment of an outcome, such as with a fuel economy standard for automobiles, or the avoidance of an outcome, such as a prohibition on the emissions of toxic substances above a specified level. Like means-based regulation, information and research are essential to establishing appropriate performance standards. A regulator uses preliminary information to develop a proposed standard and such information must provide the basis for effective monitoring and measuring of compliance (Baldwin et al. 2012; Coglianese & Lazer 2003).

In formulating either a means-based or a performance standard, the regulator must address four questions:

(1) Does the standard target the specific problem or a proxy?

(2) What level of specificity is required?

(3) Should the standard be performance or means-based?

(4) Should the standard "force" a particular technology? (Breyer 1982).

Each of these questions helps determine the level of flexibility and amount of required information needed to encourage compliance.

Theoretically, means-based standards are more enforceable and predictable than other regulatory instruments (Breyer 1982). In particular, for regulations such as pollution control, technology standards provide useful certainty for both the regulator and regulated entities. As described by Wendy Wagner (2000:101), "because the reference point is a definable technology for which numerical standards have been nationally developed, technology-based requirements are almost always clear, easy to codify, and easy to reflect in permit requirements."

Nevertheless, the specificity of means-based standards may inhibit innovation by requiring compliance with a particular technology rather than allowing firms to pursue other, more cost-effective solutions. Such standards often "freeze" technological innovation and bias firms towards using existing technology rather than trying to innovate. Still, some argue that sufficiently stringent regulations, such as "technology-forcing" regulations, can spur innovation. A technology-forcing standard is really a performance standard that requires a higher level of performance than the available technology can offer at an acceptable price. For example, Driesen (1998) notes the plethora of opportunities for innovation resulting from the ban of ozone depleting chemicals in the U.S.

To operate effectively, technology-forcing standards rely on a firm's investment in research and development to find a new technological solution to a regulatory problem. Although competition between firms provides incentive for regulated entities to develop these new technologies (Gerard & Lave 2005), enforcement also plays a key role since regulated entities

are more likely to invest in research and development when the regulatory agency has shown commitment to enforcing the regulatory standard.

Performance standards address some of the critiques of means-based standards by allowing for more flexibility in how a regulated entity attains the required end state (but when, where, and by how much are still dictated by the regulator). Firms are thus encouraged to innovate to find cheaper ways to achieve the end goal mandated by the regulation. For this reason, performance standards are often advocated as a more desirable approach to regulating relative to design standards (Bennear & Coglianese 2012). Although bound by a common approach which specifies the outcome desired, in practice, performance-based standards exhibit a great deal of variability in how they are designed. Performance regulations vary with respect to:

(1) their level of specificity;

(2) how performance levels are determined;

(3) their scope with respect to the ultimate objective and motivation; and

(4) the type of problem the regulation is trying to solve (Coglianese et al. 2003).

How variation in these areas is defined for a performance standard dictates the extent to which a firm actually has flexibility to innovate. A performance standard may blur into a means-based standard if it concerns the performance of a very specific technology or requires specific modeling procedures to determine performance (Coglianese et al. 2003). Alternatively, in some cases, the two types can be fruitfully combined into one to realize advantages of both. Hybrid options can combine a technology standard with equivalency clauses such that a regulated entity can an opt-out of the more rigid standard given equivalent performance with an alternative solution.

Still, increased flexibility may not benefit all firms equally – the costs of searching for ways to meet performance standards can be an additional burden on smaller firms for example (Coglianese et al. 2003). Additionally, performance standards miss opportunities for flexibility and cost-savings when the cost of compliance varies across firms, but the regulation is nevertheless applied uniformly across all firms. In such cases, by forcing firms with higher costs of abatement to achieve the same level of abatement as those that can do so more cheaply, the program does not minimize the cost of attaining a given level of total abatement (Ackerman & Stewart 1985). If, instead, the low-cost firms were encouraged to do more than the high-cost firms, overall costs to achieve any given reduction in pollution would be lowered.

The innovation that performance standards can spur though their flexibility may also make it difficult for the regulator to determine if, in fact, the regulated firms are in compliance with those standards. As a case in point, Peter May (2003) describes New Zealand's reform of its building regulations from one based in prescriptive standards to a more performance-based regulatory system. Although the shift encouraged innovation including experimentation with new, more cost-effective materials to build condominiums, little experience existed to guide regulators and firms to know for sure whether these new materials would adequately resist water. Ultimately, they did not, culminating in a housing crisis when it was later determined that several thousands of homes and multi-unit buildings were at risk of collapsing because water had infiltrated their structures (May 2003).

Case Study: Fuel Economy Performance Standards

The U.S. Corporate Average Fuel Economy (CAFE) Standards for new vehicles were first enacted in 1975 for the purpose of increasing fuel economy of cars and light trucks in order to reduce energy consumption. The U.S. Department of Transportation's National Highway Traffic and Safety Administration (NHTSA) enforces CAFE standards. The initial iteration of CAFE standards required that cars increase their miles per gallon (mpg) from 18 to 27.5 by 1985. Light trucks (including pickups, minivans and sport utility vehicles or SUVs) were held to a slightly lower requirement of 16 mpg by 1980 and 22.5 mpg by 2008. In 2007, with the passing of the Energy Independence and Security Act (EISA) and, in 2009, with administrative action to establish a national program,[2] fuel economy standards increased to a combined average fuel economy for new passenger vehicles of 35.5 mpg by 2016 (Anderson et al. 2011).

The experience of CAFE offers important lessons for the design of performance standards more generally. The fact that CAFE standards have a long history enables regulators to learn from past design challenges and program outcomes. The CAFE standards are binding performance targets, whereby companies are fined $55 per vehicle for every one mpg that they fall short of the performance goal (measured as a sales-weighted average of their total fleet). Historically, fuel economy has closely followed the CAFE standards. After the initial standards were established, no new requirements were established until the early 2000s. While fuel economy, measured by the amount of fuel needed to travel some distance, initially improved with the introduction of the standards, without increasingly stringent standards, innovation shifted toward improvements in vehicle performance, such as horsepower and acceleration. Advances in fuel efficiency (in the sense of energy harnessed per gallon of fuel combustion) enabled bigger and faster vehicles to comply with the CAFE standards (Anderson et al. 2011).

Additionally, the design of separate standards for cars and light trucks resulted in unintended and perverse incentives to shift large cars into the light trucks category. Although fuel economy standards are intended to "tax" less efficient vehicles and "subsidize" more efficient vehicles, the separate standards resulted in a subsidy for small trucks and a tax on larger cars. This incentive structure encouraged manufacturers to alter the design of large cars to meet the specifications of a small, light truck, vastly expanding the automakers' sales of SUVs. Such design flaws provide regulators with important information for improving future performance standards. Increasing new vehicle fuel economy from 25 to 35 mpg is expected to reduce long-run fuel consumption by about 30 percent. Economists note that in comparison to a tax on fuel, fuel economy standards may be less cost effective because taxes allow manufacturers more options for complying with reductions. In countries like the U.S., however, fuel economy standards appear more politically feasible than taxes. Future innovations in the design of fuel economy standards may include combining the current performance standard with market-based mechanisms such as tradable credits (which we discuss next).

[2] The national program combines the requirements for the U.S. National Traffic Highway Safety Administration to regulate fuel efficiency, the U.S. Environmental Protection Agency to regulate carbon dioxide (CO_2) emissions per mile, and the California Air Resources Board which established more stringent statewide emissions standards in 2002 that were finalized in 2004.

B. Market-Based Mechanisms

As a category of regulatory instruments, those that are considered market-based or incentive-based can take a wide variety of forms, including taxes, fees and charges, tradable permits, and bonds. Although we distinguish between each of them below, these mechanisms share the common feature that they each use market signals – by manipulating prices and quantities – to encourage desirable social behavior rather than specific commands (Stavins 2003). Similar to performance standards, market-based instruments specify a particular price or quantity and allow firms to respond to those signals in whatever way they see fit. At least in theory, performance standards can be made to look like some market-based mechanisms, particularly if the standard is applied at the industry level relative to the firm level (Carrigan & Coglianese 2011). Still, contrasting how most performance standards are structured in practice, market-based instruments do not require firms to meet the same targets, instead allowing each firm to abate the unwanted byproduct to different levels depending on their costs associated with complying (also known as their marginal cost of abatement). In this way, they differ from most traditional regulatory commands, most especially means-based regulation.

As a result, market-based regulatory mechanisms are commonly cited as even more cost-effective than performance standards because they allow those firms that are more able to abate at lower costs to do more of it than other higher-cost firms. Moreover, firms have incentives to continue to find ways to reduce emissions as they enjoy the benefits of doing so through, for example, fewer emissions charges or costly permits needed. In contrast, means-based and even performance instruments offer no rewards for exceeding the standards (Popp 2003). Stated differently, market-based regulation gives regulated entities that have low costs of abatement incentives to reduce harmful actions or externalities as much as possible rather than just down to the level specified by the performance standard (Baldwin et al. 2012:112).

As described by U.S. Supreme Court Justice Stephen Breyer (1982: 271), incentive-based regulatory mechanisms "direct behavior in a socially desirable direction, without freezing current technology and while preserving a degree of individual choice." Incentives are intended to shift consumer and producer behavior such that they minimize the total social costs associated with a particular activity. Market-based tools permit flexibility to account for variation in location, products, processes, and factors such as firm size, all while still responding to a specific regulatory target or goal.

Even with their advantages, similar to standards, market-based mechanisms face challenges due to a series of unknowns. These unknowns include: the possibility of establishing an efficient system; the degree to which enforcement is more or less difficult compared to means-based or performance standards; the probability that such a system will gain political traction; where the administrative responsibilities will be housed; and how the proceeds raised should be distributed (Breyer 1982). For example, even when specific market-mechanisms are known to be more efficient, other factors – including political feasibility – may result in less efficient designs (Keohane et al. 1998). Moreover, additional factors such as the relative complexity of the associated rules and the additional need for monitoring and enforcement capacity, in some cases, may increase costs to the regulator.

1. Taxes

As a market-based regulatory tool, "corrective" taxes alter market signals by manipulating the effective price of the underlying good. For example, if a regulator chooses to tax gasoline, and the tax is set at a high enough rate, it will encourage consumers to drive less and manufacturers to make more fuel-efficient vehicles. Thus, the ultimate aim of a corrective tax is to change behavior rather than to raise revenue (Coglianese & Kagan 2007). Often, a regulator's efforts to change behavior involve encouraging the producer of an unwanted side effect of an otherwise economically-beneficial activity (i.e. a negative externality) to consider its impacts on others and reduce the amount of the side effect that is created as a result. A corrective tax is an instrument appropriate for a regulator to be able to realize this objective.

In designing a corrective tax, regulators "must decide on *what to tax* (the tax base), *how much to tax* (the tax rate), and *how to tax* (the administrative mechanism for collecting the tax)" (Cordes 2002:258). Given an appropriate tax rate, regulated entities are encouraged to reduce pollution to the point where the tax is equal to their marginal cost of abatement (the cost of abating one more unit of pollution). Theoretically, the tax rate should be set equal to the marginal benefit to society of lessened pollution (Stavins 2003).

Determining what to tax will impact its ultimate effects as taxing a proxy may only imperfectly affect regulated entities' behavior in intended ways, if it affects behavior at all. For example, if the goal of a tax on gasoline is to reduce emissions, one which is simply imposed on fuel consumption will not encourage companies to develop forms of gasoline that emit less, nor will it encourage consumers to choose less polluting gasoline. Only if the tax is specifically imposed on fuel that pollutes more would firms and consumers change the types of fuels they develop and use (Cordes 2002). Even so, one reason why a tax may be imposed on a proxy instead of the pollutant itself is that the technology may not be available to monitor the pollutant (Stavins 1998). Huppes and Kagan (1989) demonstrate by contrasting two corrective tax programs in the Netherlands targeting pollutants in the water supply, one on harmful substances discharged from industrial sources and another on manure from agricultural sources. While the former reduced pollution by over 90 percent and stimulated efforts to find more cost-effective abatement methods, the latter met resistance and has been largely ineffective. In addition to the presence of a more effective regulator, the industrial tax has been more successful because of the availability of better measuring and monitoring technology as well as the limited number of industrial polluters to regulate.

Taxes can be further challenging to implement because of the uncertainty in forecasting how the regulated entities will respond. As a result, regulators may have difficulty in determining what tax rate to set to achieve the regulatory goal. In practice, regulators can use trial and error (also known as learning-by-doing) to identify the optimal tax if the risks in doing so are relatively small. However, this approach may be infeasible when dealing with potentially catastrophic risks (Baldwin et al. 2012). In fact, the choice between using a price-based mechanism like taxes and a quantity-based mechanism like tradable permits (which we discuss below) can turn on the welfare effects of setting the mechanism incorrectly initially (Weitzman 1974). When a slight error can be very harmful as is the case with nuclear energy development for example, it is best to regulate quantity. In contrast, in situations where the harm caused by

additional units of pollution stays relatively constant, as might be true with carbon dioxide emissions for example, using taxes as the regulatory tool is more efficient.

2. Fees or Charges

Like taxes, fees or charges use economic signals to spur changes in behavior. For example, as described by Joseph Cordes (2002), when the U.S. city of Seattle, Washington began approaching the limits of its municipal landfill capacity, it faced the prospect of having to pay another county in order to transfer waste to it. Instead, Seattle began charging residents based on the amount of garbage disposed at a rate of $16.10 for each large trash can. At the same time, to encourage recycling, paper and glass recycling was removed free of charge, and green or yard waste was subjected to a fee of only $4.25 per month.

In operation, the policy had both intended and unintended consequences. Although residents increased their voluntary recycling, they also began compacting trash into the smallest volume possible to fit more into each receptacle in an effort to reduce the amount that they had to pay for their waste (Cordes 2002). This example further demonstrates the similarities between taxes and fees. Like taxes, fees are sometimes able to only imperfectly channel behavior to accord with social goals because the fee can only feasibly be applied to a proxy for the behavior being targeted.

The design of corrective fees is also similar to taxes. Regulators must establish what is being charged, how much to charge, and how to collect and administer the charge. The amount of the charge should be set to the social cost of the activity being regulated or, alternatively phrased, the social benefit from the removal of the activity. Obtaining accurate information needed to judge that optimal level can be challenging, and so in practice, the fee should minimally "raise the price of the regulated activity 'enough' to produce the desired cutback" (Cordes 2002:264).

3. Tradable Permits

When the optimal level of regulatory control is known (or determined over time through trial and error), taxes and tradable permits will both yield efficient outcomes. Yet, in practice, policymakers tend to adopt tradable permits systems more frequently than taxes (Keohane et al. 1998), likely because the aversion that the public has to taxes makes them political infeasibility. In tradable permit systems, the regulator "issues a number of permits and each of these allows a specified course of behavior" (Baldwin et al. 2012:117). Alternatively, the system can be structured such that to the extent regulated entities reduce emissions beyond the legal requirements, they can use those reductions as credits which can be traded with other firms (Stavins 2003; Tietenberg 2005). In this way, tradable permits offer the advantage over taxes that the aggregate amount of emissions can be determined in advance by the regulator. The regulator can allocate entitlements (either in the form of permits or in setting limits) freely initially or, alternatively, can auction them off (Tietenberg 2005). The latter option raises tax revenue for the government which can be used to offset government spending generally or to compensate those who may be negatively impacted by the regulatory policy.

Once allocated to the regulated entities, these firms then have the incentive to trade permits or credits until the cost of compliance (i.e. the marginal cost of abatement) is equal across all regulated entities. Trading allows for the firms with lower compliance costs to sell entitlements to pollute to those with higher compliance costs. If one firm has a lower cost of abatement than another, the firm with the lower cost can sell one of its entitlements to the firm with the higher cost, such that both are made better off. Trading leads to a reduction in the total cost of achieving a desired regulatory goal in comparison to means-based or performance-based regulatory instruments because the firms that abate more are those that can abate most cost effectively (Stavins 1998).

Case Study: Acid Rain Program and SO2 Tradable Permits

The 1970 Clean Air Act gave the U.S. Environmental Protection Agency (EPA) authority to regulate sulfur dioxide, or SO_2, emissions. Subsequently, the 1977 Clean Air Act Amendments required new coal-fired power plants to use flue gas desulfurization units or "scrubbers" to remove 90 percent of the SO_2 from exhaust emissions. While this law did not encourage innovations that would enable scrubbers to remove additional pollution, it did create incentives for affected businesses to find ways to reduce the cost of installing and operating scrubbers (Popp 2003). With the 1990 Clean Air Act Amendments, regulation of SO_2 emissions shifted to a market-based structure. Rather than require a specific technology, this approach established a tradable permit system which sought to reduce SO_2 emissions by 10 million tons from 1980 levels (Stavins 1998). Emissions limits were set for the electric utility plants that emitted the most SO_2, and each generation unit was required to have a certain number of emissions allowances to cover their pollution.

Experience with this shift demonstrates that firms will take advantage of the given flexibility in market-based regulation and develop other more efficient solutions to achieve compliance. Not only did firms respond with alternative solutions for compliance, such as increased use of low-sulfur coal but they took advantage of options such as banking allowances for future use. Furthermore, technological advances in the removal efficiency of scrubber technology and in the transport of coal by rail were also a direct consequence of the flexibility provided in the permitting system (Popp 2003; Stavins 1998).

While these innovations illustrate the advantages of a market-based system over conventional standards, a set of factors unique to this program and to the time period were also instrumental in its success. These conditions included pressure from advocacy organizations (in particular, the Environmental Defense Fund), an agreed upon upfront reduction goal of 10 million tons rather than simply mandating a more cost-effective reallocation of emissions, and the absence of a precedent or status quo for acid rain regulation (Stavins 1998). The viability of the program was further enhanced by the availability (and use) of emissions monitoring technology (Stavins 1998). Through the lobbying efforts of the environmental interest group community, all utilities were mandated to install costly continuous monitoring technology (Burtraw & Swift 1996). Coupled with stiff penalties for violators, this technology contributed to very high compliance rates (Tietenberg 2005). The success of the SO_2 trading system and other systems like the European Union's emissions trading system suggests the possibilities for market-based mechanisms, all while illustrating that the "best" mechanism depends on the nuances of the particular policy problem confronting the regulator (Stavins 1998).

As with any regulatory instrument, tradable permit systems have limitations. In order to ensure that the regulatory goal is achieved, regulators must be able to monitor performance for each regulated entity to ensure they are not exceeding their allocation (Driesen 1998). Additionally, tradable permit systems may not be easily adjustable given unforeseen environmental changes. For example, if permits are used to control river pollution, and the water level drops due to a heat wave or drought, adjusting pollution accordingly becomes difficult as entitlements have already been issued (Baldwin et al. 2012). To implement a tradable permit system successfully, the market must have adequate information on the number of potential buyers, sellers, prices, and volume of trades (Cordes 2002). Lack of information or other uncertainties may reduce the value of permits and fail to create the necessary incentives to alter behavior (Baldwin et al. 2012).

4. Bonds

As a regulatory instrument, bonds provide direct incentives for regulated entities to comply with the stated regulatory goals. The most straightforward form of a bonding requirement requires a regulated entity to deposit an amount equal to the size of the obligation into a secured account, where it earns interest based on the yield of the low-risk financial assets in which that money is invested (Davis 2015). The amount plus the interest accrued are then returned to the company when the regulatory agency determines that the risky activity has been conducted without incident or that the appropriate amount of cleanup has occurred after completion of the activity.

While bonds are not as commonly discussed as other market-based instruments, they are most frequently cited as mechanisms for environmental control. In such cases, regulated entities may be required to post a bond prior to engaging in any activity that might cause pollution. If the required levels of pollution control are not met, the bond is forfeited. The goal of a performance bond is for the regulated entity to "internalize perceived social costs into its private resource allocation decisions" (Shogren et al. 1993:111). In this way, the bond encourages the firm to account for the negative externalities associated with its activities, which it would otherwise ignore.

Bonds are simultaneously a mechanism to manage moral hazard (Gerard 2000). In economic theory, moral hazard refers to opportunism whereby the more informed party takes advantage of the less informed party through unobserved actions (Perloff 2001). Such opportunism may occur when the regulated entity engages in excessively risky activities or does not employ adequate protective measures since it shares much of the damage it creates with others.

If the goal is to completely eliminate the potential for moral hazard, the amount posted for performance bonds should be set to an estimate of the worst-case scenario for possible damage given the available information. However, because bonds also impose costs on firms (which are described in detail below), the optimal level for the requirement will be less than this estimate of the maximum damage that might be created through the economic activity (Davis

2015). Ideally, the bond value will vary over time as knowledge is gained from practical experience or the worst-case scenario shifts due to innovations.

Given high levels of uncertainty with respect to the future negative impacts of firm actions, performance bonds may be an effective alternative to direct price or quantity regulation. If the bond amount is set using currently available information, bonds also provide incentives for regulated entities to invest in their own monitoring and measuring of outcomes and impacts. With a flexible bond amount, the required bond amount may be adjusted as better information on impacts becomes available. Further, bonds can reduce the need for the regulator to test the product before approving it, serving as a way for the firm to guarantee the outcome of that testing process (Bohm & Russell 1985). To encourage firms to continue to monitor and limit their impacts even after completing the economic activity, bonds can be held until enough time has lapsed to ensure that "the feared damages from their past activities were either proved unfounded or were covered by the bond" (Shogren et al. 1993:113). Thus, bonds present a viable regulatory alternative even in cases where the full extent of the impact is not realized until long after the firm's project is completed.

While offering a number of attractive features, bonding also present drawbacks that can be classified into three categories: moral hazard, liquidity constraints, and legal constraints (Shogren et al. 1993). In much the same way that bonds can be a solution to moral hazard on the part of firms, they can present opportunities for regulators to act opportunistically as well. Bonding is at risk for moral hazard if the regulator aims to maximize personal welfare over social welfare by, for example, confiscating the bond unfairly.

An important additional downside of bonds is that they impose liquidity constraints on firms. In the extreme, some otherwise viable firms may be entirely eliminated from participating in the industry because they do not possess or cannot borrow the necessary capital to post the bond. Even when they can post the required bond, bonds still impose costs on firms by the very fact that the money used will likely need to be deposited into safe investments to ensure it is available should a failure occur. Since the firm's cost of capital will be greater than the interest it earns even when it acts with an appropriate level of precaution (Davis 2015), the bonds are tying up resources that could be used more productively elsewhere.

Finally, given that they are essentially contracts between regulators and regulated entities, bonds are subject to costs connected to enforcing those contracts. A firm may challenge a regulator's decision in court based, for example, on claims that either events beyond their control impeded their ability to adequately mitigate the unwanted activity or that the regulator somehow acted inappropriately in confiscating the bond (Shogren et al. 1993). Even when such defenses fail, they still impose costs by squandering resources in legal battles.

Case Study: Bonding Requirements for Natural Gas Production

Through recent technology advances in horizontal drilling techniques and capacities to better pinpoint sources of natural gas, hydraulic fracturing is making available previously inaccessible sources of natural gas in shale rock. With shale gas production in the U.S. estimated to increase by over 70 percent from 2013 to 2040 (U.S. Energy Information Administration 2015), federal and state regulators face challenges in overseeing drilling in new places while using techniques where little experience exists to guide regulatory oversight. Particularly with respect to the potential for groundwater contamination, the long-term costs in terms of threats to human health and the environment of hydraulic fracturing are highly uncertain.

Importantly, at the point at which these costs are identified, a natural gas producer may no longer be operating or be otherwise unable to pay for the damages it has caused. As a result, a gas producer has incentives to operate in ways that do not adequately consider the potential risks of extraction because that firm does not bear the burden for the costs imposed. Here, the extraction company, which knows more about how much precaution it is taking in removing natural gas from shale rock, has an incentive to do too little to prevent long-term damage because that company is less likely to have to bear its cost (Davis 2015).

To help manage these issues, bonding requirements can be imposed. In the U.S., if drilling on federal land, producers are required to post a bond with the U.S. Bureau of Land Management in the amount of $10,000 for each lease. Alternatively, the producer can post bonds of $25,000 for all leases in a state or $150,000 for all U.S. leases (Davis 2015). The last two options are known as "blanket" bonds because they cover all activity when the driller has multiple operations occurring simultaneously. A number of U.S. states also require operators to post bonds when the extraction occurs on state-owned land. Among states, the structure of these bonds vary, both in terms of the required dollar values which range from $500 to $100,000 and whether the producer can post "blanket" bonds.

In addition to bonds that mandate the driller to deposit the required amount, the federal government also allows surety bonds (Davis 2015). Surety bonds are guarantees by an insurance company that the operator is going to perform its duties without imposing environmental costs. In the event that the driller does impose such costs, the insurer pays the government the stated amount. The extraction company pays a premium to the insurance company to secure a surety bond.

In addition to covering the unforeseen environmental and health costs imposed by hydraulic fracturing, such bonds are used to ensure the operator restores the drilling site to its former state once extraction is complete. By holding the amount in a secured account, the natural gas extraction company is required to both operate safely and restore the land in order to get its money back. If the firm does not comply, the federal or relevant state government now has the required resources to remedy the failure. Thus, the moral hazard problem is mitigated through the bond, particularly if the amount is set appropriately. Similar to bonding requirements more generally, imposing these requirements for natural gas extraction ties up company operating capital and may limit some firms from participating in the industry. In the context of U.S. natural gas production, the problem may be more that the requirements are set too low. For example, the Bureau of Land Management bonding requirements have not been adjusted – even for inflation – in over 50 years (Davis 2015).

C. Management-Based Regulation

In their implementation, performance standards and market-based mechanisms both require data measuring the outputs of the regulated business process. In contrast, with management-based regulation, data gathering occurs much earlier in the firm's operating cycle. In fact, management-based regulation attempts to influence firm choices well before the problem that the regulation is trying to solve becomes evident (Bennear & Coglianese 2012). Alternatively referred to as "enforced self-regulation" or "mandated self-regulation" (Ayres & Braithwaite 1992; Hutter 2001), management-based regulation attempts to engage regulated entities in data gathering and planning to realize the goals of the regulatory program. At the same time, this regulatory strategy allows firms flexibility in designing their plans around the intricacies of their specific operations (Coglianese & Lazer 2003). While management-based approaches mandate that firms engage in the planning process, they do not force firms to achieve certain outcomes like performance standards, nor do they require specific actions by the firm (beyond planning) as means-based approaches do (Carrigan & Coglianese 2011). In fact, management-based regulation need not mandate that firms even execute their own plans.

In implementing management-based regulatory approaches, regulators outline the elements that must be included in firm's plan, which might typically include identifying the hazards in a firm's operations, what actions the business will take to mitigate the risks connected to those hazards, and the processes in place to identify and respond to problems (Coglianese & Lazer 2003). Moreover, plans might often include information on how employees will be trained to manage the hazards as well as how management will evaluate the organization's performance in achieving social goals.

Through this planning process, management-based regulation attempts to foster improved regulatory performance both by encouraging managers to evaluate their internal processes to find ways to achieve regulatory goals while, at the same time, facilitating information sharing between the regulatory and the regulated entities (Bennear & Coglianese 2012; Karkkainen 2001). The fact that management-based regulation requires an examination of certain risks or problems may encourage firms to leverage this investment and identify opportunities for additional modifications to operations. Such opportunities can result in additional benefits for both the regulated entity and society (Coglianese & Lazer 2003). In addition to assigning the risk assessment and control to the decision makers with the most information – namely the firm's management – management-based regulation offers the advantage that it may reduce employee resistance to the associated reforms (Ayres & Braithwaite 1992). After all, the firm should be more likely to want to comply with its own processes and controls, relative to requirements imposed by the regulator.

Management-based regulation is most useful when industry outputs are difficult to monitor and measure, the operations of the regulated entities are heterogeneous, and industry technology is changing rapidly (Bennear 2006; Coglianese & Lazer 2003). Mean-based standards are frequently selected as the regulatory tool of choice when firms are more homogeneous, but as heterogeneity in operations and turnover in technology increases, such traditional mechanisms become more difficult and costly to implement effectively.

Case Study: Management-Based Regulation in Offshore Oil and Gas Development

The explosion and fire on the Deepwater Horizon drilling ship in the Gulf of Mexico in 2010 unleashed the largest offshore oil spill in U.S. history (National Commission 2011). In addition to providing the impetus for a temporary moratorium on drilling in deep water in the Gulf as well as other U.S. waters, the Gulf oil spill brought with it a series of reforms including the disbanding of the regulator in charge of offshore development, the Minerals Management Service (Carrigan 2014). Through the combination of an initial directive, interim final rule and a 2012 final rule, additional requirements for testing and inspecting blowout prevention equipment, preparing and maintaining wells, and auditing of well design were also implemented in the wake of the disaster (Carrigan forthcoming).

The spill also provided the impetus to make mandatory the American Petroleum Institute's recommended practices of what amounts to a management-based regulatory approach called a Safety and Environmental Management System (SEMS) (Bennear 2015). The associated rule as well as the subsequent SEMS II rule requires oil and gas companies not only to conduct assess-ments to identify risks and develop plans to respond to and minimize those risks but also to subject those plans to review by independent auditors (Carrigan forthcoming). Furthermore, operators are required to identify who has authority over operations at all phases of a project and provides all employees with the power to stop work when facing a dangerous situation.

Still, the management-based approach now used to regulate U.S. offshore oil and gas operations is not a new development. Also prompted by major accidents, both the U.K. and Norway have long utilized celebrated management-based regulatory regimes to manage risk associated with offshore oil and gas production (Bennear 2015). Following the Piper Alpha accident in 1988 in which 167 workers died, the U.K. implemented a goal-oriented approach known as the "safety case," overseen by the U.K. Health and Safety Executive. The safety case requires regulated firms to create plans which provide descriptions of each oil and gas facility's operations, mechanisms by which blowouts will be prevented, risks associated with pipelines, and emergency procedures for workers.

While Norwegian regulatory practice was already moving from a means-based regulatory regime to one more emphasizing goals, the death of 123 offshore workers following the sinking of the Alexander Kielland drilling platform in 1980 expedited the shift (Bennear 2015). As in the U.K., operators in the Norwegian Continental Shelf are required to develop plans specific to individual platforms and to evaluate and be ready respond to operational risks. Still, unlike in the U.K. where approval by the Health and Safety Executive is required, in Norway, firms develop plans with input from organized labor and Norway's offshore regulator, the Petroleum Safety Authority, and so do not need approval from the regulator after the plan is formulated (Kaasen 1991). When a problem with the safety plan arises, the three groups – government, industry, and labor – collectively reach a solution, typically without sanctions (Kringen 2013).

While the transitions occurred at different times, the offshore regulatory approaches in Norway, the U.K., and the U.S. appear to be converging toward a consistent model which emphasizes detailed safety plans typical of management-based regulation, relative to specific requirements more associated with means-based regulation. Nevertheless, it is still too early to empirically evaluate the relative effectiveness of management-based regulation in managing the risks connected to offshore drilling (Bennear 2015).

Yet, the same conditions that encourage the use of management-based regulation can also make enforcement challenging as "good management" across a heterogeneous population can be difficult to ascertain (Coglianese & Lazer 2003). Management-based regulation limits the amount of information that regulators have to gather, but the informational disadvantage can be exploited by firms that can take advantage by doing the minimum or treating the required planning as simply another box to check in their production processes (Carrigan & Coglianese 2011). The challenge for the regulator then becomes determining if the firms are engaging in the planning process responsibly (Coglianese & Lazer 2003). Thus, although firms may be more likely to comply since they are responsible for establishing the rules and controls in their plans, nevertheless, regulators may need to conduct inspections to ensure that plans are both in compliance with the regulatory requirements and are indeed being implemented (Braithwaite 2002).

Management-based regulatory strategies have been utilized in a variety of regulatory domains, including programs to reduce pollution, prevent industrial accidents, and limit the distribution of unsafe food. Although empirical evidence ascertaining the impact of management-based regulation is limited, studies of U.S. state-level programs indicate that firms find that the required planning process can be helpful in identifying cost-effective areas to improve regulatory performance (Bennear & Coglianese 2012; Keenan et al. 1997; Natan et al. 1996). These studies also suggest that larger firms may realize greater benefits than smaller companies. In her study of differences in firm toxic chemical releases in U.S. states that adopted management-based regulations relative to those that did not, Lori Bennear (2007) reports that the implemented programs reduced emissions by 30 percent. However, these differences no longer existed after the management-based regulations were implemented for six years, suggesting the effects of the instrument on toxic emissions may be temporary.

D. Information Disclosure

As consumers are faced with more and increasingly complex decisions, mandated information disclosure has become a popular regulatory tool because it aims to provide consumers with information needed to respond effectively (Graham 2002; Kleindorfer & Orts 1998). In addition to its long history as a centerpiece in the approach to regulating financial securities (Ferrell 2007), information disclosure has come to occupy a larger place in programs regulating a variety of social domains, including worker safety (Sunstein 1999), consumer products (Fung et al. 2007), and automobile safety (Graham 2002).

Like management-based regulation, information disclosure attempts to prompt socially-responsible firm actions by placing requirements on them early in their product planning processes (Bennear & Coglianese 2012). However, instead of requiring complete plans as with management-based regulation, with information disclosure, firms are mandated simply to collect and disseminate information about their operations (Carrigan & Coglianese 2011; Karkkainen 2001). Stated differently, mandated disclosure establishes specific requirements for what information must be provided, but does not require regulated entities to respond to the information they provide (Bennear & Coglianese 2012).

Rather, by forcing firms to disclose data about their operations, the hope is that public

pressure will compel firms to improve their regulatory performance. As damaging information about firm operations attracts media attention (Hamilton 2005), public scrutiny results (Graham 2002), which can prompt regulated entities to make improvements. Firms may be further compelled to act to change their behavior if consumers react to the information by no longer buying products from those organizations failing to fulfill their perceived social responsibilities (Jin & Leslie 2003). For example, in cases where information disclosure may be linked to a decrease in a company's stock price (given publicly available information) or used as a proxy to measure a firm's production efficiency, information disclosure may successfully encourage firms to comply with social goals (Konar & Cohen 1997). These organizations may alternatively improve their regulatory performance out of a fear that they might face more stringent regulations in the future if they do not change (Bennear & Olmstead 2008; Sunstein 1999). Finally, information disclosure may prompt improvements simply by "making managers more aware of and concerned about their organization's social outputs" (Coglianese & Lazer 2003:695).

From the perspective of economic theory, mandated information disclosure can correct for the existence of information asymmetries, where one party in the transaction has more information about the product or service than the other. In such cases, the market can fail to allocate the product or service efficiently, in the sense that some mutually beneficial transactions may no longer take place (see, e.g., Akerlof 1970). Thus, one goal of information disclosure is to enhance "the preconditions of a competitive marketplace," enabling buyers and sellers to make more informed choices (Breyer 1982:161). In such cases, disclosure regulation uses information as a mechanism to increase market benefits by facilitating competition (Manning 1989).

Of course, for mandated information disclosure to be effective, the information provided must actually be useful. Not only must it be indicative of firm performance on regulatory goals (Kleindorfer & Orts 1998), the information must simultaneously be disseminated in a form that the consumer can understand and react to it. As described by Breyer (1982:164), "disclosure is likely to be effective only where the public can understand the information disclosed, where it is free to choose on the basis of that information, and where it believes the information is materially relevant to the choice." It is because of these issues that critics assert that information disclosure can fall short of fulfilling its regulatory goals. Notably, Ben-Shahar and Schneider (2011; 2014) point out that, in practice, providing consumers with information does not enable them to make more informed decisions. As described by the authors, "people don't notice disclosures, don't read them if they see them, can't understand them if they try to read them, and can't use them if they read them" (Ben-Shahar & Schneider 2014:55).

Given the varied opinions regarding the value of mandated information disclosure in achieving regulatory goals, research has sought to empirically test whether information disclosure does improve firm regulatory performance. Although certainly not conclusive, evidence from the environmental domain is suggestive that disclosure can be at least moderately effective in inducing behavioral changes and improvements (Bennear & Coglianese 2012; Kraft et al. 2011). For example, in their study of the 1996 Amendments to the U.S. Safe Drinking Water Act, Bennear and Olmstead (2008) find that mandating community water suppliers to disclose violations of drinking water regulations through annual mailings reduced subsequent violations by between 30 percent and 44 percent. These results suggest that mandated

information disclosure can significantly change firm behavior, at least when used in tandem with more traditional regulatory tools such means-based or performance standards.

Still, the value of information disclosure in achieving regulatory goals is by no means a settled issue (see, e.g., Winston 2008). One reason for the ongoing debate is the difficulty in empirically teasing out the effects of disclosure specifically. Analysis of the implementation of the U.S. Environmental Protection Agency's Toxic Release Inventory (TRI) presents one example. Created in 1986, TRI houses publicly-available data on toxic chemical releases by certain facilities which are mandated to report them because they use the chemicals in large enough volumes (Bennear & Coglianese 2012). Although the creation of TRI is correlated with significant declines in chemical releases (Fung & O'Rourke 2000; Thaler & Sunstein 2008), determining what portion is attributable to TRI is difficult since at least some of the improvements are connected to other developments including, for example, traditional regulations prompted by the 1990 Clean Air Act (Hamilton 2005). Others studying the TRI

Case Study: Los Angeles County Restaurant Hygiene Cards

One place where information disclosure has been used as a mechanism to induce more socially responsible behavior by regulated entities is in the food industry. After a November 1997 news expose revealed the extent of the unsanitary practices of some restaurants in Los Angeles, California, the Board of Supervisors passed an ordinance to force Los Angeles County restaurants to display the results of inspections by the county's Department of Public Health (Jin & Leslie 2003). Although the Department had long conducted inspections of restaurants in the county, the ordinance instructed inspectors to simultaneously issue a card which showed the restaurant's grade, indicating its cleanliness on a standard scale: "A" to denote the most sanitary restaurants, a "B" to denote the next grouping, etc. Restaurants were then required to display these cards so that possible customers could see them. Although mandated at the county level, as a result of their incorporated status, some Los Angeles County cities were free to choose whether to implement the ordinance. However, even if a city decided not to force their restaurants to display the grade cards, those restaurants still received the cards from the Department of Public Health and, thus, could display them voluntarily.

Ginger Jin and Phillip Leslie (2003) conducted a statistical study of this disclosure policy in Los Angeles County in an attempt to assess the impact of the restaurant grade cards on restaurant cleanliness. Using a series of panel data regression analyses on a set of over 13,000 restaurants in Los Angeles County, Jin and Leslie demonstrate that the information disclosure ordinance had statistically significant impacts on hygiene scores, restaurant revenue, and foodborne illnesses. In fact, the program's implementation resulted in over a 5 percent increase in restaurant hygiene scores in those places where it was mandated. Restaurants that received an "A" in their inspection enjoyed revenue growth of close to $15,000 – an increase of just under 6 percent – relative to before the cards were introduced. Further, Jin and Leslie (2003) reveal that foodborne illnesses declined with the introduction of the grade cards, as measured by a 20 percent decline in hospitalizations for digestive disorders in mandated Los Angeles County areas and a 13 percent decline where disclosure was voluntary. While the decline in hospitalizations was not all driven by customers switching to good hygiene restaurants, this case supports the theory that information disclosure can impact the behavior of regulated entities in intended ways.

program have shown that changes in firm procedures as well as some misrepresentation explain a portion of the claimed reductions (de Marchi & Hamilton 2006; Poje & Horowitz 1990; Natan & Miller 1998). The TRI reporting system itself may also have a role, as those firms whose chemical use falls below certain levels no longer report, misleadingly suggesting that they are no longer releasing the toxic chemicals at all (Bennear 2008).

E. Voluntary Programs and Self-Regulation

Voluntary programs are the most flexible instrument given there is no requirement to comply. Indeed, such programs are really not a *regulatory* instrument at all. Rather, they "encourage socially desirable behavior by offering educational resources, financial assistance, awards and certifications, and exemptions from more formal requirements, especially for firms that have a history of solid regulatory performance" (Carrigan & Coglianese 2011:117). Voluntary programs come in a variety of forms, including product certifications by regulators which aim to induce firms to change by altering consumer preferences; recognition and reward programs that label an entire regulated facility or firm as exemplary; and negotiated agreements between businesses and regulators that deliver regulatory outcomes using alternative approaches that exceed existing standards (Borck & Coglianese 2009; Borck et al. 2008). Further, industry self-regulation – which does not directly involve the government regulator at all – can also be considered a type of voluntary regulatory program (Borck & Coglianese 2009; Potoski & Prakash 2005).

In addition to the direct recognition and rewards received from the regulator, the decision by a firm to participate in a voluntary program organized by a regulator can be motivated by investor pressure (Gamper-Rabindran 2006). Similarly, the organization may benefit if potential customers value socially-responsible practices and are willing to pay more for such goods and services (Segerson & Li 1999). Public recognition through such programs may allow firms to increase their market share if consumers use the information to guide purchase decisions, as some evidence on the U.S. Environmental Protection Agency's Energy Star Program suggests may be the case (Delmas & Terlaak 2001). Alternatively, voluntary regulatory programs may facilitate opportunities for innovation by encouraging collaborative exchanges between industry players. Although, generally, voluntary programs are not likely to cause widespread innovation, they can encourage diffusion of best regulatory practices and raise industry awareness of regulatory goals (Delmas & Terlaak 2001).

Participation might also be precipitated by a belief that, by doing so, the organization will be better suited to manage potentially more stringent regulatory standards in the future, dissuade lobbying efforts by environmental groups, and increase relative costs for competitors (Gamper-Rabindran 2006; Innes & Sam 2008). Participation may also be prompted by a hope that it will encourage the regulatory agency to adopt a more lenient approach to enforcing the traditional regulatory requirements also in place. Voluntary programs can, thus, provide a platform for firms to increase regulatory flexibility by encouraging the regulator to relax current regulations or delay adoption of new or tighter regulations, all while positioning the regulated entity to have a greater voice in the creation of future regulations (Borck & Coglianese 2009).

For example, firms that participated in the U.S. Environmental Protection Agency's (EPA) National Environmental Performance Track program received a variety of benefits in exchange for agreeing to exceed existing regulatory requirements. Not only did EPA highlight these companies on its website and notify relevant elected officials, participation qualified members for special rewards and provided them with access to regulatory officials (Coglianese & Nash 2014). Regulated entities participating in Performance Track were also considered lower priority for routine inspections and subject to less stringent reporting requirements.

At the same time voluntary programs provide numerous potential benefits to firms, possibilities for free riding and issues of confidentiality can limit company participation. Regulated entities that participate may face some amount of risk that competitors might access sensitive information and that regulators may use information gleaned from voluntary programs to create more stringent regulations in the future (Delmas & Terlaak 2001). If the overall industry benefits from a voluntary program but not all members of the industry are participants, non-participants may secure those benefits without investing in the regulation. Alternatively, firms may "participate" without necessarily trying to fulfilling the obligations in good-faith, thereby enjoying the reputational advantages at little cost to their organizations (Coglianese & Borck 2009). Still, public dissemination of information on violations that can potentially threaten a firm's reputation to help mitigate such free riding (Delmas & Terlaak 2001).

Relative to voluntary programs orchestrated by the regulator, industry self-regulation is even further removed from means-based or performance-based regulatory instruments. Self-regulation shares a defining characteristic of other types of voluntary regulation, in that it does not mandate any action from the regulated entities. Further, self-regulatory initiatives encourage the regulated entities to self-impose requirements. Still, unlike many voluntary programs which are organized and overseen by a government entity, self-regulation does not explicitly involve an agency at all. In fact, with self-regulatory programs, the regulator and regulated entities are both non-governmental entities and can end up being one and the same (Coglianese & Mendelson 2010).

Often, standards-setting organizations are responsible for the maintenance of self-regulatory programs (Haufler 2001; Nash 2002). One prominent example is the International Organization for Standardization (ISO) which (among its many standard series) oversees the ISO 14000 environmental management series (Potoski & Prakash 2005). Coupled with ISO 14001 where firms obtain certification in return for implementing intricate environmental management systems, the series also consists of a variety of guidelines surrounding environmentally-friendly labels, evaluating environmental performance, life-cycle assessment, and communicating environmental attributes of products (Potoski & Prakash 2005). Even in 2007, more than 150,000 certificates had been issued for ISO 14001 alone (ISO 2009).

While third-party organizations provide one mechanism for the development of self-regulatory structures, industries themselves can also band together to collectively design their own regulatory standards. Examples include the American Forest and Paper Association's Sustainable Forestry Initiative (Meidinger 2003) as well as the Institute of Nuclear Power Operations (INPO), operated by the U.S. nuclear industry (Rees 1994). In some cases, industry codes of conduct and programs are borne from a crisis in the industry (Rees 1997).

One recent example is the oil and gas industry's Center for Offshore Safety, patterned after INPO and created less than one year after the Gulf oil spill. The Center aims to help operators develop and implement workplace safety practices (Carrigan forthcoming). Another prominent example is the chemical industry's Responsible Care program. Although created by the Canadian Chemical Producers Association before the tragic 1984 Union Carbide chemical plant accident in Bhopal, India, the Responsible Care program gained traction after it because chemical companies felt they could do little to reverse the industry's negative public perception (King & Lenox 2000; Nash & Ehrenfeld 1997). The U.S. Chemical Manufacturers Association adopted the Responsible Care program in 1988 and the British and Australian chemical industries joined shortly after (Coglianese & Mendelson 2010).

Like voluntary programs, self-regulatory regimes can be implemented to stem additional regulation by government or to influence how future regulations are designed (Carrigan & Coglianese 2011). As a result, industry members may find it beneficial to be part of such a program and to comply with its principles (Segerson 1999). Much like a trade association, firms may also join to benefit from the exchange of information that occurs as a member of such a group. In some cases such as the Responsible Care program, participation in the self-regulatory regime is a precondition for membership in the industry's trade association (King & Lenox 2000). Oftentimes, as is true of the Center for Offshore Safety which is part of the American Petroleum Institute, the program is designed and run by the industry trade association.

Given that part of the motivation for firms to join government-sponsored voluntary programs or organize their own self-policing mechanisms is to stave off further government regulation, it is reasonable to be skeptical about the extent to which these regulatory instruments can achieve social goals. In fact, by giving a false impression that they have addressed the regulatory concerns, these programs may have the perverse effect of intensifying residual risk (Lyon & Maxwell 2004).

Researchers studying specific programs often reach conflicting conclusions about whether voluntary programs do improve regulatory performance. This disparity is largely the result of two related research-design difficulties associated with engaging in such research (Carrigan and Coglianese 2011). The first is ascertaining whether the regulated entities that join the government-sponsored voluntary program or self-regulatory program are the same firms that would have performed better than their counterparts regardless of participation (Gunningham et al. 2003). In other words, those that join might simply have different motivations than those that do not join, a phenomenon that researchers call selection bias. The second is that discerning whether these mechanisms achieve regulatory goals requires disentangling their effects from other environmental developments, including other types of programs imposed by the regulator simultaneously (Greenstone 2004).

Despite these challenges, some research has found that such voluntary programs can help achieve regulatory goals, even if the improvements are only modest (Ayres & Braithwaite 1992; Braithwaite 2002). This is specifically the case when these initiatives are instituted with the specter of discipline for those do that not adhere (Braithwaite 2002; King and Lennox 2000) or in the face of societal pressure on the industry to reduce negative externalities (Gunningham et al. 2003; Howard-Grenville et al. 2008). Thus, even when the regulatory standards are completely

designed by an industry, the existence of an effective and willing regulator in the background may provide the impetus for real social gains.

In studying voluntary environmental programs in particular, Borck and Coglianese (2009:321) note that such programs "may be effective alternatives to mandatory regulations for achieving small environmental improvements at relatively low cost." Morgenstern and Pizer (2007), in reviewing seven voluntary programs, note some but not dramatic energy-related and environmental improvements. Other research has shown that third-party standards specifically

Case Study: The U.S. Environmental Protection Agency's 33/50 Program

In early 1991, the U.S. EPA initiated the 33/50 Program, a voluntary regulatory program focused on reducing emissions of 17 high-priority chemicals (U.S. EPA 1999). The aim of the program was to reduce releases of targeted chemicals by 33 percent by 1992 and 50 percent by 1995 relative to 1988 baseline levels. While participation was voluntary, EPA actively invited firms to participate in the program if any of their facilities had emitted any of the targeted chemicals in 1988 or later (U.S. EPA 1999). Starting with those with the largest amounts of emissions, EPA invited companies in waves, and by the beginning of 1994, the agency had reached out to over 8,500 of the roughly 10,000 firms that emitted any of the 17 chemicals on the list between 1988 and 1995 (Khanna 2007). Of those eligible to participate, 13 percent of firms joined the program (U.S. EPA 1999).

The 33/50 Program was structured such that participating firms would submit a plan for how they intended to reduce releases of the targeted chemicals. However, neither the plans nor the reduction targets were binding (Gamper-Rabindran 2006). In addition to providing assistance to firms that decided to participate, EPA further encouraged participation by publicly recognizing firms enrolled in the program (Khanna & Damon 1999). In a retrospective assessment of the 33/50 Program, EPA noted that the program achieved its goal of reducing 33/50 chemical releases by 50 percent one year earlier than expected (U.S. EPA 1999). By the end of 1995, emissions of the 17 chemicals targeted by the program had dropped by over 800 million pounds, equivalent to a 55 percent reduction relative to 1988 levels.

However, closer examination of these data illustrates the previously described difficulties inherent in estimating the impacts of voluntary programs such as the 33/50 Program (Carrigan & Coglianese 2011). Numerous researchers have tried to ascertain how much, if any, of the reductions achieved were actually attributable to the program. Certainly, those reductions achieved between 1988 and 1990 could not be attributed to the program because 33/50 was not initiated until 1991. Further, in addition to the fact that some of the chemicals included in the program were also subject to other EPA regulatory programs, the U.S. had also agreed to phase out use of two of the chemicals on the list by 1996 through the Montreal Protocol (Gamper-Rabindran 2006). Releases of other chemicals not targeted as part of the program also fell during the period, and some of the cited reductions were achieved by transferring chemicals to recycling firms. Thus, while some such as Khanna and Damon (1999) and Innes and Sam (2008) report that the program accounted for significant declines in emissions up to 45 percent, other studies including those conducted by Gamper-Rabindran (2006) and Vidovic and Khanna (2007) find that the 33/50 Program had close to zero impact on emissions.

can improve regulatory firm performance. After controlling for differences between participants and non-participants, Potoski and Prakash (2005) found that those certified through ISO 14001 simultaneously complied to a greater extent with government regulations as well.

Although the evidence is less definitive compared to other types of voluntary regulations, self-regulatory programs designed by industry groups can help achieve social goals as well. For example, INPO, created in 1979 after the Three Mile Island nuclear accident in the U.S., appears to have been effective in reducing nuclear accident risks (Rees 1994). In contrast, participants in the chemical industry's Responsible Care program actually reduced their toxic releases more slowly than non-participants (King & Lenox 2000). However, these results were obtained before the U.S. Chemical Manufacturers Association decided to publicly reveal those lagging in their implementation of the program's principles (Nash 2002).

Based on these and other similar results, some argue that, in addition to program features, industry and regulatory agency characteristics may influence the efficacy of self-regulation (Gunningham 1995; King & Lenox 2000). The fact that the nuclear industry was more concentrated and faced a single regulator may have made it more likely that poor performers would be targeted with enforcement actions, thus motivating the industry to enact a more stringent self-regulatory program.

III. Enforcing Regulatory Programs

In addition to instrument design, the ultimate impact of regulatory policies hinges on how those policies are enforced. Regulatory enforcement broadly refers to "all activities of state structures (or structures delegated by the state) aimed at promoting compliance and reaching regulations' outcomes" (OECD 2014:11). Compliance can be encouraged through a variety of mechanisms, including monitoring firms to ensure that "they are doing (or not doing) what is required (or forbidden) of them" (Russell 1990:243). Alternatively, compliance can be encouraged through self-monitoring and reporting as well as citizen monitoring and inspections of "products of business premises, activities, documents, etc." (OECD 2014:11).

Although compliance is often discussed as a single outcome, in reality, the term can refer to a diverse set of situations. Initial compliance is most closely associated with technology requirements, as when a particular technology is required to achieve a permitted discharge level. Continuing compliance requires ongoing oversight to ensure that the desired outcome is being attained. Like compliance, monitoring can refer to a multitude of scenarios which include monitoring for either technological capability associated with initial compliance, ongoing compliance with the standard like a permitted discharge amount, and overall ambient quality to see, for example, if the sum of all the permitted discharges is achieving the overall desired outcome (Russell 1990).

Two categories of factors affect regulatory enforcement in practice. The first comprises the "technical, economic and legal problems the agency encounters" (Kagan 1994; Morgan & Yeung 2007:187). Here, the capacity of the regulator, including the ability to detect violations and the seriousness of violations, is important (Morgan & Yeung 2007). The second set of

factors reflects the larger political environment influencing regulatory enforcement strategies, including elected officials, interest groups, and the priorities of regulatory leaders. These factors will combine to affect regulators' decisions about what rules to prioritize for enforcement, which firms to target, and the manner and stringency with which violations are addressed.

A. Stringency

Stringency refers to the extent to which regulations require "significant, costly, or controversial changes in existing patterns of human behavior" (Kagan 1989:95). Very stringent rules may require the elimination of all risks or the use of novel or untested measures. Less stringent regulations may "demand only incremental changes in the plans or practices of most regulated entities" (Kagan 1989:95). Even when not very stringent, regulations still may function as a "fail-safe" mechanism should market pressures change and lead firms to cut corners with respect to compliance with good operational practices.

Just as rules can vary in their level of stringency, so too can enforcement. The number of inspections per firm, the probability of an inspection, and the frequency of fines can all vary. In a study of the U.S. Environmental Protection Agency's enforcement of its rules and regulations associated with the U.S. Clean Water Act, Helland (1998b) demonstrates that the stringency of individual inspections can exhibit a wide range even within the same regulatory program, from cursory inspections not intended to detect violations to costly, detailed inspections in which abatement technology and emissions samples are drawn.

Researchers have not identified an "ideal" level of stringency. Indeed, what has sometimes been called the "multidimensionality" of regulatory problems (Brunel & Levinson 2013) – that is, the way they can be varied and dispersed across a regulator's geographic jurisdiction, each affected by diverse set of regulated entities – will mean that enforcement stringency should also often vary. Brunel and Levinson (2013) suggest several factors that may affect the appropriate level of stringency, such as:

(1) *simultaneity* – stringency may need to be heightened in places subject to disproportionately high or concentrated levels of regulatory problems (e.g., high levels of pollution),
(2) *industrial composition* – areas with more risky or pollution intensive industries may require more stringent enforcement, and
(3) *capital vintage* – in some cases regulatory enforcement stringency may be lessened by effectively "grandfathering" down the inspection of older regulated entities out of a sense of fairness.

B. Deterrence

Deterrence results when the threat of punishment prevents regulated entities from committing future violations. While all regulators seek to achieve deterrence, using the threat of punishment to dissuade violations, the types of deterrence emphasized by a regulator will affect how it structures its enforcement programs. Enforcement strategies can pursue optimal deterrence or absolute deterrence. Absolute deterrence aims to prevent *all* violations under the

Case Study: Optimal Deterrence - U.S. Coast Guard Oil Spill Prevention Program

Since oil tankers are constantly moving, regulators cannot often observe actions taken by those manning these ships. As a result, the regulator is not able to easily detect violations or compliance. The U.S. Coast Guard's enforcement strategy to deal with these complications combines monitoring, detection, and penalties. Specific monitoring and detection efforts include vessel inspections, selective observation of oil transfer operations when ships are docked, and random patrols to locate unreported spills (Cohen 1987). Although the associated statute sets the penalty per spill at $5,000, the fines actually imposed by the Coast Guard are typically much smaller. Violators are also responsible for cleanup costs up to a maximum threshold. Failure to disclose a spill to the Coast Guard can result in a maximum $10,000 penalty and/or one-year of jail time.

Economic theory assumes that the regulator should theoretically increase monitoring and detection until the benefits and costs are equal (i.e. where the marginal social benefit equals marginal social cost). Part of determining optimal deterrence is establishing an effective penalty, such that a "firm finds it in its own interest to choose the socially optimal level of effort" (Cohen 1987:29). Under conditions of perfect information – meaning that all spills are detected – and risk neutrality among firms – meaning that firms are indifferent between compliance and noncompliance when they have equal expected costs – penalties would be set such that they are equal to the marginal damages plus the cleanup cost. In practice, regulators including the Coast Guard cannot detect all pollution and, thus, optimal penalties must account for the ability to detect infractions. For example, if current technology or regulatory capacity is likely to only detect half of the violations, then the penalty should be twice as much.

Cohen's (1987) analysis of the Coast Guard's penalties indicates that the agency uses a negligence standard rather than strict-liability standard. This means that they punish only in cases where the regulated entity's lack of care is deemed to have caused the spill, as opposed to holding companies liable for infractions regardless of the cause. In fact, when firms are risk averse, the regulator's use of a negligence standard may be necessary to keep firms from funneling too many resources into preventing spills. Relatedly, the Coast Guard punishes intentional discharges more heavily than discharges due to unknown causes. On average, intentional polluters receive two to four times the penalty.

As the spill size increases, the penalties should also increase, but the $5,000 cap on the size of the penalty is likely too low to trigger compliance for larger spills. Yet, while the penalty is too low for large spills, Cohen (1987) simultaneously demonstrates that the Coast Guard sets penalties too high for small spills, such that the benefits of cleaning up are far below the costs and pollution penalties. Although firms are required to pay for the cleanup costs associated with a spill, the existence of a cap means that often companies are not liable for the full cost of the damages. The result is that some of the burden falls to the government. Given this shared liability regime, shifts toward more optimal penalties will likely continue to have little impact on compliance. By removing shared liability and requiring insurance to cover penalties, the Coast Guard might better encourage compliance without the risk of company bankruptcy.

Empirical analysis of tanker spills demonstrates that smaller vessels, higher oil prices, more Coast Guard resources allocated to observing oil transfer operations, and greater emphasis on patrols to locate unreported spills result in smaller oil spill sizes. Interestingly, inspections do not appear to have an effect on average spill size (Cohen 1987). The U.S. Department of Transportation (1999) has shown that from 1986 to 1995 the total number of reported oil spills increased, but the volume of oil spilled in each declined. Grau and Groves (1997) attribute at least a portion of this improvement to a shift in the Coast Guard's inspection strategy during the period to a more targeted approach where the agency sorted tankers into high and low priority categories based on risk, focusing on those in the high priority category.

premise that noncompliance will not result in other social benefits, or that the resulting social benefits never outweigh the costs. In this case, penalties "should be at least equal to the net gain to the offender arising from that conduct" (Yeung 2004:66). In order to achieve absolute deterrence, regulated entities must believe that compliance will be less costly than noncompliance. Otherwise, they have minimal incentives to change their behavior (Earnhart & Glicksman 2006).

Pursuing optimal deterrence, by contrast, seeks to prevent inefficient violations, as it asserts that society gains from infractions when the benefits of the violating behavior outweigh the costs of halting that behavior (Becker 1968; Yeung 2004). In order to achieve optimal deterrence, the penalties for violations should be set to equal the cost of harm caused. If penalties are set too high, they may result in "over-deterrence," where some violations are prevented that would otherwise create social gains which exceed the harms they cause, and so would be welfare-enhancing (Yeung 2004).

In addition to considering whether to focus on achieving optimal or absolute deterrence, regulators also have the choice with respect to whether to encourage compliance through specific deterrence or general deterrence. Just as the adjectives imply, specific deterrence focuses on how the regulator can take actions to keep specific regulated entities from violating the rules while general deterrence refers to how actions a regulator takes against a specific firm might also keep other firms in compliance.

1. Specific Deterrence

Specific deterrence aims to prevent individual offenders from violating the rules again. Regulators can use a variety of mechanisms including inspections, threats, notices, and fines to deter firms and individual facilities from violating rules. Since the 1990s, a handful of studies have examined the impact on compliance of specific deterrence-based enforcement schemes, with the majority of these projects focusing on environmental regulations. A central question regulatory enforcement researchers have analyzed is whether inspections and enforcement actions by government regulators do in fact encourage firms to behave better (Earnhart & Glicksman 2011). Overall, the results on specific deterrence indicate that monitoring, inspections, and enforcement actions do result, to varying degrees, in reductions in noncompliance (Shimshack 2007).

Studying application of the U.S. Environmental Protection Agency's water pollution regulations to the pulp and paper industry specifically, Magat and Viscusi (1990) measure the influence of inspections on compliance. Unlike other regulatory standards, "water pollution standards involve permissible pollution amounts that vary across firms and over time" (Magat & Viscusi 1990:336) to account for differences across industries as well as the feasibility in meeting targets given existing technologies. Based on the requirements, polluting facilities must measure discharge levels and report them on a monthly basis. As a result, inspections are not testing initial compliance but rather "whether the DMR [Discharge Monitoring Report] discharge levels are reported accurately and honestly" (Magat & Viscusi 1990:337). The authors report that lagged inspections (or inspections related to pollution levels or compliance status for the previous four quarters) both encouraged firms to reduce pollutant discharge levels as well as to comply more closely with permitted levels. In fact, each inspection resulted in a 20 percent

reduction in mean pollutant discharges. Inspections also encouraged firms to more frequently self-report pollutant discharges.

Similar results are obtained by Laplante and Rilstone (1996), who study pulp and paper water pollution in Canada. Lagged inspections by the Quebec Ministry of the Environment reduced pollutant discharges by about 28 percent. However, in addition to actual inspections, the authors also show that the threat of inspections substantially impacts compliance as well. Moreover, unlike Magat and Viscusi (1990) who failed to find significant variation by region, Laplante and Rilstone (1996:27) demonstrate regional variation in the effects of inspections, concluding "that there might be important regional differences in the nature of the relationship that exists between the regulator and the regulatees and/or the monitoring and enforcement procedure across regions."

Other research demonstrates that firm characteristics can affect compliance decisions as well. In their study of the steel industry, Gray and Deily (1996) show not only that lagged enforcement and inspections by the U.S. Environmental Protection Agency (EPA) increased compliance, but that the costs of bringing the facility into compliance reduced the impetus to comply. Further, among the steel manufacturers studied, "corporate attitude" significantly increased compliance, as demonstrated by the fact that "firms with higher compliance rates in other plants in the previous year were more likely to comply" (Gray & Deily 1996:110). Finally, Gray and Deily (1996) demonstrate that firm characteristics impacted enforcement decisions by U.S. EPA inspectors as well. For example, steel plants were less frequently inspected if they had exhibited past compliance, resided in less polluted areas, and were more likely to close because of financial difficulties.

More recent studies have attempted to separate out the role of inspections relative to enforcement actions on firm compliance decisions. For example, Nedeau (1997) examined the effectiveness of the U.S. EPA's enforcement program by separating monitoring – including inspections and tests – from enforcement actions – such as penalties and orders – to determine if the former also impacted the amount of time facilities spend in noncompliance with air-pollution control regulations. He found that a 10 percent increase in monitoring leads to a 4.2 percent reduction in the amount of time a plant is in violation of the standards. Nedeau (1997) attributes the effect not to increased overall regulatory activity, but to the EPA's willingness to target based on compliance status.

Gray and Shadbegian (2005) study the effects of U.S. EPA efforts to enforce air quality regulations at pulp and paper mills by separating inspections from other enforcement activities including notices of violation, penalties, and phone calls. The authors demonstrate that both can impact subsequent compliance, but that notices, penalties, and phone calls have a larger effect than inspections. At the same time, plant characteristics also weigh heavily on compliance decisions. Doubling a plant size is associated with a 6 percent decline in the probability of being in compliance, and older plants are also less likely to comply. Moreover, plants which violated water pollution regulations as well as U.S. Occupational Safety and Health Administration rules were simultaneously more likely to violate air pollution regulations as well.

Finally, Earnhart (2004) studies both specific and general deterrence by analyzing discharges by municipal wastewater treatment plants in the U.S. state of Kansas from 1990 to 1998. While enforcement actions, including letters, orders, and fines, significantly reduced subsequent emissions, much like Gray and Shadbegian (2007), Earnhart's (2004) results indicate that inspections have a weaker effect on compliance.

2. General Deterrence

While specific deterrence targets individual infractions, general deterrence uses broad threats or regulatory actions against one entity to encourage others to comply. Broadly, general deterrence requires both that those being deterred will respond to information in a rational manner and, moreover, that they have knowledge of the information. When rational behavior is paired with sufficient threats, potential violators can be dissuaded from infractions even with limited realization of any threats (Kleiman et al. 2009). The New York City Police Department's experience with "squeegeeing," where people wipe others' car windshields in traffic with the hope they will be paid for it, provides one example of general deterrence. When the Department increased enforcement of their zero tolerance policy for squeegeeing, the number of arrests declined, indicating that public awareness of arrests encouraged others to desist without having to be arrested as well (Kleiman et al. 2009:14230).

Still, critics of general deterrence question the assumption that firms know about enforcement measures against other firms. Furthermore, firms may be apt to look for differences between them and those penalized and, thus, assume they are no more likely to face punishment. This results in a "'weak signal, weak threat' hypothesis – that is, the message often does not get through or send a meaningful threat" (Thornton et al. 2005:263).

In an effort to determine the effectiveness of general deterrence, Thornton et al. (2005) differentiate between *implicit* and *explicit* general deterrence. Explicit general deterrence occurs when a firm reacts to knowledge of enforcement against other firms. Implicit general deterrence, in contrast, is where a firm reacts due to their own knowledge of regulations, regardless of enforcement actions against their peers. Social, economic, and normative pressures likely influence implicit general deterrence more than the stringency of sanctions. Publicity plays an important role as well. Arguably, enforcement is not sufficient without publicity, and vice versa (Williams & Wells 2004).

To evaluate the importance of explicit general deterrence in encouraging firm compliance, Thornton et al. (2005) surveyed 233 companies, in industries ranging from wastewater management to aluminum and steel fabrication, regarding their knowledge of recent enforcement activities. While employees were often able to recall at least one fine of another company or prison sentence for an official from another company, they generally underestimated the frequency and magnitude of penalties. The results generally indicated that people are not "particularly attentive to penalty information. Nor have they made special efforts to obtain timely and accurate information" (Thornton et al. 2005:272). When provided with enforcement information, individuals often responded that they were more likely simply to review programs than to either change their behavior or alter employee training as a result. Ultimately, these

findings indicate a weak relationship between increased information on penalties and increased perception of legal risk.

While deterrence-based regulation assumes that regulated entities will only comply with regulations given legal sanctions, firms tend to comply with regulation more frequently than expected (Thornton et al. 2005). This can be attributed to factors such as the firm's sense of social or legal obligation, as social norms may instill a sense of duty to protect and care for health, safety, and the environment. Based on in-depth interviews with facility managers from two industries (electroplating and chemical manufacturing), Gunningham et al. (2005:309) identify that implicit general deterrence and facilitating a "culture of compliance" were more important than either specific or explicit general deterrence in ensuring solid regulatory performance in the electroplating industry. The regulations rather than other enforcement actions remind firms of their responsibilities, such that "the regulations had become so embedded in their culture that they exerted an almost unconscious influence on decision-making" (Gunningham et al. 2005:309). In contrast, in the chemical industry, compliance was driven more by the fear of informal sanctions from important stakeholders given any violations.

While the evidence from Gunningham et al. (2005) supports previous studies such as those conducted by Braithwaite and Makkai (1991:35) who show minimal influence of the "certainty of detection, the certainty of punishment, and the severity of punishment" on behavior, additional studies identify effective regulatory strategies that use both specific and general deterrence. In fact, some, like Shimshack (2007:14), assert that, "Inspections and enforcement actions consistently produce spillover effects on non-sanctioned facilities" (Shimshack 2007:14). For example, Shimshack and Ward (2005) study the pulp and paper mill industry and demonstrate that an additional fine has a significant impact on "reputation-based deterrence," with spillover effects reducing the statewide probability of a violation in the following year by 64 percent. Similarly, in his study of municipal wastewater treatment plants, Earnhart (2004) finds that inspections and enforcement actions at other plants as well as the probability of inspection and enforcement at the particular plant in question substantially decreased emissions.

In 2004, the State of Oregon Department of Environmental Quality (DEQ) surveyed 450 companies who either have permits or registrations with the DEQ. The survey revealed that 11 percent of the companies adjusted management, operations, or production due to direct DEQ enforcement (i.e. specific deterrence). Additionally, penalties increased the likelihood that firms would "seek permits and modifications, request technical assistance, make process changes for environmental compliance, and involve upper management in disciplining staff when environmental violations occur" (Carlough 2004:1). Moreover, in the three years leading up to the study, 38 percent of companies made changes after hearing about inspections of other facilities, 33 percent made changes after hearing about technical assistance offerings, and 14 percent made changes after hearing about penalties levied on other firms. Seventy percent of surveyed companies asserted that they believed DEQ is likely to catch significant violations, and 41 percent believed the same for smaller violations. Moreover, 70 percent of companies indicated that bad publicity or risk of exposure would outweigh any financial savings from noncompliance. Factors such as "forced shut-downs, environmental damage, criminal prosecution, corporate reputation, community pressure and customer pressure" were more important than financial penalties (Carlough 2004:2).

Similarly, Gray and Shadbegian (2007) also found support for the effectiveness of both specific and general deterrence by studying correlations in environmental performance among firms in particular geographical areas. The authors used the total number of inspections as a proxy for specific deterrence, and the total number of inspections over three years within a ten-mile radius as a proxy for general deterrence. Greater numbers of inspections at the plant in question, at nearby plants, and at plants in the same state resulted in greater compliance. Further, plants close to one another generally had similar compliance records. Additionally, Shimshack and Ward (2008) find that enforcement – both specific and general deterrence – does not just result in compliance but can result even in over-deterrence. The authors partially attribute over-compliance in their sample to the risk of random enforcement and the effects of joint pollutant removal technologies (Shimshack & Ward 2008).

Case Study: General Deterrence – Click It or Ticket Program Implementation

Wearing seatbelts has been shown to reduce death and personal injuries during accidents (Friedland et al. 1990), and while they are readily available in today's automobiles, some individuals still choose not to wear them in violation with many states' laws. Seat belt laws first took hold in Canada, initially in Ontario, then Quebec, and later in all provinces. In the U.S., the city of Elmira, New York was among the first to enact a seat belt law in 1983, using both publicity and enforcement to achieve seat belt use of close to 80 percent. Given Elmira's success, other cities and states have adopted its program as a model for monitoring and enforcing seatbelt use. The state of North Carolina established the first U.S. statewide program by using Elmira's experience to inform its *Click It or Ticket* program strategies (Williams & Wells 2004).

Elmira's initial program increased seatbelt use from 49 percent to 77 percent using a three-phase strategy. The city's program would rely on a week of advertising on television and radio, followed by a week of publicity which emphasized warnings, and finally, one week of ticketing. While initial seatbelt use increased, within four months after the first program, it dropped by over 10 percent. Learning from Elmira, North Carolina's program used eleven different enforcement campaigns and over 22,000 checkpoints from 1993 to 1998 in order to publicize enforcement activities. Given finite resources, the strategy relied heavily on general deterrence by creating "an atmosphere in which drivers anticipate they are likely to be stopped and fined if not obeying the law" (Williams & Wells 2004:178). North Carolina's seatbelt use ranged from 80 to 84 percent after intensive enforcement and dropped by a little less than 10 percent in periods with lower levels of ticketing. Signs offering "feedback" or comparisons to other drivers (e.g., percentages of drivers wearing seatbelts in the previous week) helped increase seatbelt use to a high of 90 percent.

Like other enforcement strategies, success with *Click It or Ticket* is often attributed to a combination of strategies. According to a National Highway Traffic Safety Administration (NHTSA) survey, drivers who were ticketed or received a warning for not wearing their seatbelt indicated they started wearing a seatbelt more frequently afterward. Nevertheless, their use of seatbelts remained below drivers who never received a ticket or warning. While effective publicity will increase deterrence, the threat of "meaningful penalties" increases the perceived likelihood of detection (Williams & Wells 2004:179). In addition to deterrence strategies, other studies attribute North Carolina's success to other factors as well, including coalition building, top-down support, feedback based on data, other statewide and local programs, and continuous funding (Reinfurt 2004).

C. Targeting

Given limited resources, regulators must prioritize inspections and enforcement actions. In general, regulators will have at least three objectives in designing their enforcement scheme, which cannot all be simultaneously achieved. They will want to minimize the number of inspections necessary to minimize cost but maximize compliance rates and provide incentives even for those firms with the highest compliance costs to comply (Harrington 1988). The extent to which a regulator achieves some combination of these objectives will depend in part on how that agency decides which regulated firms to target. Targeting relies primarily (but not exclusively) on deterrence to encourage compliance. Basic economic theory suggests that regulators can expect firms with compliance costs lower than the penalty for noncompliance to comply. Conversely, if compliance costs are higher than the penalty for noncompliance, the strictly rational regulated firm will choose not to comply. Additionally, if firms know when to expect inspections, they will have incentives to comply only during that time.

Despite these expectations from an economic model of deterrence, research shows that some firms always take compliance seriously, while others do not. In fact, evidence on the enforcement of U.S. air and water pollution regulations demonstrates that compliance is surprisingly high despite the fact inspections are infrequent and penalties are rarely assessed (Harrington 1988). As Harrington (1988:49) notes, "it is not clear that compliance cost is the most important determinant of a firm's behavior."

Targeting can help explain this apparent contradiction. Even in cases where compliance costs exceed the level of the fine (which is often true in traditional regulatory environments), firms can be persuaded to comply if regulators target certain firms based on past regulatory performance. By separating firms into groups – those that take compliance seriously, and those that do not – regulators can use their scarce resources more efficiently to minimize violations by concentrating enforcement efforts on the group less likely to comply (Harrington 1988). Targeting the firms less prone to compliance through more frequent inspections and larger fines provides a "carrot" for those who take compliance seriously and a "stick" for non-compliant firms (Friesen 2003). Allowing firms to escape the targeted group based on evidence of compliance during inspections provides these organizations with clear incentives to improve. Moreover, such a system can encourage firms to self-report violations in an effort to demonstrate that they are trying to cooperate with the regulator (Helland 1998a). Similarly, for those in the non-targeted group, the fear of moving into the targeted group if they are found to be violating the rules provides incentives to remain in compliance.

Building on this result, Friesen (2003) argues that instead of assigning firms to the targeted group based on findings of non-compliance, a more cost-effective strategy random assigns firms into that group. Compliant firms can then move to the non-targeted group, which receives less frequent inspections and smaller fines, by demonstrating compliance when they are randomly inspected. Through use of random assignment, the regulator can further reduce the amount of time it spends inspecting those in the non-targeted group and, thus, reduce the overall costs of enforcement.

Still, despite its promise as a cost-effective enforcement strategy, targeting is not without its shortcomings. Factors such as uncertainty about firms' compliance cost structures and variability in those structures can complicate targeting strategies (Raymond 1999). In addition, if remedying regulatory failures is costly and time-consuming, firms may be unable to demonstrate their desire to comply, leaving them in the targeted group (Helland 1998a). Further, because targeting requires the regulator to exercise discretion in determining on which firms to focus its attention, interest groups have incentives to use their resources to influence a regulator's judgement (Kagan & Scholz 1984), potentially leading to capture. In fact, Helland (1998a) shows that while U.S. Environmental Protection Agency's targeting strategy did encourage paper mills to self-report Clean Water Act infractions, political variables also influenced which mills were more likely to be inspected.

The U.S. Occupational Safety and Health Administration's (OSHA) enforcement program provides insights into how targeting practices are applied in practice as well as the potential ways in which they can be implemented more effectively. The historic evidence studying the effectiveness of OSHA's enforcement strategies presents a mixed picture. Gray and Scholz (1990; 1993) found that enforcement activities reduce injuries by up to 10 percent, and inspections that impose a penalty reduce injuries by 22 percent and lost workdays by 20 percent. In contrast, Smith (1979) shows that OSHA's enforcement had a significant impact on injury rates in only one of the two years he studied. Other older studies also found similar limited impacts of OSHA inspections (McCaffrey 1983; Ruser & Smith 1991).

In 1999, OSHA launched the Site-Specific Targeting program (SST) which targeted enforcement based on industry sites with the highest injury and illness rates, rates which are determined from responses to an annual OSHA survey. As of 2012, OSHA used measures such as inspection counts, citations issued, penalty amounts, injury and illness rates, and fatality rates to monitor program results but did not employ any outcomes-based performance metrics to help establish the causal link between SST and improvement in worker health and safety (Office of Inspector General 2012).

Recent research shows that OSHA's use of targeting based on injury rate data fails to identify non-compliant firms that have low injury rates due to inaccurate reporting or because their actions have not yet resulted in harm. Adding other variables that reveal a firm's financial status such as debt and sales growth as well as ownership changes and compliance with other regulations may result in more effective targeting (Finkel 2013). In addition to targeting based on variables more indicative of compliance status, evidence suggests that OSHA's first or second inspections are more effective in revealing infractions than inspections in the following periods (Ko et al. 2010; Weil 1996). Prioritizing first time inspections over repeated inspections may further strengthen OSHA's targeting practices.

When examining the effectiveness of targeting or other enforcement strategies, it is important to consider the difference between achieving overall compliance relative to achieving the desired results. For example, given OSHA's goal of promoting worker safety, an approach that maximizes compliance alone encourages OSHA to focus on easy enforcement targets independent of the resulting safety outcomes (Weil 1996). In contrast, actually encouraging worker safety would demand a targeting strategy that identifies firms or facilities with the

greatest "payoff per dollar of enforcement effort" (Cohen 1998:11). The key is not merely to look at compliance metrics but also to assess targeting strategies based on how well they achieve "socially desirable outcomes" (Kagan 1989:90).

D. Regulatory Styles

Regulatory inspection personnel vary a great deal in how they approach and interact with their regulated entities, variation which researchers typically identify as differences in the regulator's style. Two styles of regulatory enforcement identify the end points in a continuum of behaviors that characterize how regulators can address their regulated entities. The first style, sometimes called the deterrence model of regulation, views regulations "as authoritative legal norms whose violation demands punishment" (Coglianese & Kagan 2007:xiv). The second style, generally referred to as the cooperative or accommodative model of regulation, emphasizes "cooperative government-business problem-solving" (Coglianese & Kagan 2007:xiv). While some situations call for a more deterrence-based or legalistic approach, others are better suited for a cooperative approach. Further, many situations demand some type of balance between the two models. Selecting an enforcement strategy requires an understanding of when to punish and when to persuade.

Regulators can and do use both models of enforcement, but they may still prefer or identify with one model more frequently (Hutter 1989). Kagan (1989:112) describes enforcement less as a choice between one model or another but as varying along "a jurisprudentially-defined dimension – from legalistic to flexible to conciliatory – and along an outcome or policy-oriented dimension – from over-regulation to under-regulation." He calls the midpoint between the two models the "'welfare-maximizing' enforcement style, where agency officials seriously weigh costs and benefits in applying regulatory standards and formulating remedial requirements, listening to but skeptically probing regulated enterprises' technical and economic arguments, conscious of the risk of excessive leniency as well as excessive stringency" (Kagan 1989:93).

In order to address different problems and situations faced by regulators, it is useful to consider some of the characteristics that influence a regulator's enforcement approach. Kagan (1989) illustrates the impact of the size and sophistication of regulated entities on enforcement by comparing strategies for the western U.S. Office of Surface Mining (OSM) with the eastern OSM:

> Western Region OSM officials dealt with large corporate mining companies; their everyday interactions were with corporate environmental specialists; the regulated enterprises adopted a cooperative stance. Eastern Region inspectors, who more frequently resorted to legal sanctions, dealt with much smaller and environmentally unsophisticated mining firms, which often adopted a non-cooperative posture (102).

Although characteristics that influence firm compliance – and thus may influence the regulator's enforcement style – are difficult to identify comprehensively, Harrington (1988) describes some of the important considerations explaining why firms may comply with regulatory requirements. They include the fear of bad publicity, the costs associated with

additional testing, and the greater administrative burden associated with increased oversight by the regulator if the firm is found out of compliance. Helland (1998a) identifies a broader set of factors as helpful in explaining how often a firm is inspected, demonstrating that inspections increase with greater community per-capita pollution and affluence as well as where plants are less likely to have to close if violations are detected.

Appendix A reproduces a comprehensive list of the factors that researchers have identified as affecting the regulator's enforcement style (Kagan 1994). The extensive number of factors listed in Appendix A underscores the large breadth of influences on a regulator's choice of style. For example, regulators may face external pressure to use more stringent, legalistic enforcement given recent accidents attributed to regulatory negligence or after a journalistic exposé of enforcement failures (Kagan 1994). We now turn to a more in-depth discussion of the primary enforcement styles that regulators tend to choose in practice, and what effects those decisions about style have on firms' regulatory performance.

1. Legalistic Enforcement

Legalistic enforcement – also called "'sanction'-oriented," "a 'deterrence' model," or a "'coercive' style of regulation" (Kagan 1989:92) – responds to infractions with immediate notices, fines, or other tools to deter future violations. Particular situations may well lend themselves to a more legalistic strategy such as technology-forcing regulations and "no risk" statutes as well as firms that require continued and close monitoring (Kagan 1994:391; Scholz 1984b:182). The desired degree of enforcement stringency is a key component in deciding whether a regulator should use legalistic enforcement, as some regulations demand a higher level of stringency by pushing "significant, costly, or controversial changes in existing patterns of human behavior" (Kagan 1989:95). In short, a deterrence-based approach may be "most appropriate when legal statutes unambiguously define corporate misbehavior" (Scholz 1997:258).

Even so, legalistic enforcement is criticized on multiple grounds, the most notable of which is that it promotes "regulatory unreasonableness" (Bardach & Kagan 1982). Regulatory unreasonableness describes situations where "compliance would not yield the intended benefits" but rather "would entail costs that clearly exceed the resulting social benefits" (Bardach & Kagan 1982:6). Unreasonable requirements can be of two forms: "'rule-level unreasonableness,' which has to do with aggregate economic inefficiency, and 'site-level unreasonableness,' which has to do with particular encounters between the enforcers and the regulated" (Bardach & Kagan 1982:7). While benefit-cost analysis can be used to identify rule-level unreasonableness, site-level unreasonableness addresses more personal, qualitative interactions. Much of the dissatisfaction with regulation can be attributed to site-level interactions, as when regulated entities are subject to specific costs and requirements determined by particular regulatory agents.

Even "sensible" regulations such as requiring two fire exits can be unreasonable when they impose costs on all to benefit the select sites that actually require said regulations. Moreover, regulatory unreasonableness does not account for the fact that regulated entities exhibit differences in capacity. Given limited resources, some firms can spend more than others on compliance. Yet, in the face of these sources of site-level unreasonableness, Bardach and

Kagan (1982) note that there is limited evidence of regulators revising or bending regulations to reduce it. At its extreme, the problem of regulatory unreasonableness can be deemed "government imposed injustice," weakening "the belief that we live under a system of fair and rational laws" (Bardach & Kagan 1982:28).

In order to cope with variation, policies tend to be over-inclusive, which is a source of unreasonableness. However, the main catalyst is legalistic enforcement. Legalistic enforcement relies on a "literal application of over-inclusive rules to all sites irrespective of differences in sites" and "makes regulatory unreasonableness not an occasional nuisance but a pervasive problem" (Boyum 1983:753). Bardach and Kagan's (1982) critique of inflexible regulation and legalistic enforcement forms much of the basis for exploring other styles of regulatory enforcement which rely more on cooperation, flexibility, and responsiveness to correct for many of the inefficiencies of "going by the book."

While Gunningham et al. (2005) note that the site managers from the electroplating and chemical manufacturing firms they interviewed did not complain of "regulatory unreasonableness," other critics of legalistic enforcement cite a plethora of issues including failures to improve performance, increased cost, reduced effectiveness due to hostility or resentment, and unwanted side-effects such as unemployment (Baldwin et al. 2012). Moreover, inspectors engaging in legalistic enforcement may gravitate toward enforcing the rules that are easier to monitor (May 2003; Sparrow 2000). As a result, enforcement may do less to achieve the ultimate goals of the regulatory program, particularly if the unobservable elements of the program are central to achieving those goals.

2. Cooperative or Accommodative Enforcement

Very much a response to the perceived failures of the legalistic approach, cooperative or accommodative enforcement, also called "conciliatory" enforcement (Shover et al. 1984), represents the opposite extreme to legalistic or deterrence-based enforcement. Hawkins (1984:3) calls the conciliatory style a "characteristic of the compliance strategy," whereas the sanctioning or legalistic strategy adopts a penal style. It is "more interested in 'compliance' than in deterrence, as oriented toward seeking results through 'cooperation' rather than by coercion, as 'consultants' rather than 'cops'" (Kagan 1989:92).

While some of the regulatory enforcement literature tends to blend the cooperative or accommodative style with the flexible or responsive style which we describe in the next section, May and Winter (2000) consider accommodative enforcement at the opposite end of the spectrum from legalistic, with flexible enforcement as the midpoint between the two (see, also, Kagan 1989). They identify four groups of regulators: token enforcers, insistent enforcers, persuasive enforcers, and legalistic enforcers. Token enforcers are accommodative, insistent enforcers and persuasive enforcers are flexible, and legalistic enforcers are legalistic. May and Winter (2000) emphasize two factors that differentiate regulators based on enforcement style: formalism and coercion. Formalism is "the degree of rigidity in interactions that varies from informal conversations to rule-bound stances on the part of the inspectors" (May & Winter 2000:147). Coercion is "the willingness to issue threats that vary from a trusting inspector not issuing warnings, to a skeptical inspector threatening to report or to impose penalties for

violations" (May & Winter 2000:147). As the type most closely associated with the accommodative style, token enforcers exhibit low levels of formalism and coercion, "marked by informal interactions with occasional threats of coercion" (May & Winter 2000:161).

Thus, the accommodative model emphasizes trust between the regulator and regulated entities, providing advice and sympathy more often than punishment. More flexible approaches often emphasize "social process" (Coglianese & Kagan 2007:xvi), focusing on problem-solving and compliance through informal techniques such as education, persuasion, and negotiation (Baldwin et al. 2012). Regulators determine enforcement strategies based on interactions with officials at each facility and employ the combination of techniques that will encourage cooperation (Coglianese & Kagan 2007). The ultimate goal of cooperative enforcement is to prevent harm rather than punish for infractions, with legal action as a last resort "since compliance strategy is concerned with repair and results, not retribution" (Hawkins 1984:4). Of course, while the flexibility afforded by a more conciliatory approach allows regulators to better tailor their enforcement to individual circumstances, this discretion can also be misused, leading to "opportunities for corruption and discrimination and opens a regulatory agency to capture by the regulated community" (Sparrow 2000:238).

3. Flexible Enforcement and Responsive Regulation

While some authors call this approach flexible enforcement and others call it responsive, they both form a style that allows for a variety of forms of compliance with the intention that regulated entities will "address noncompliance pro-actively" (Earnhart & Glicksman 2006:608). In May and Winter's (2000) description, flexible enforcement differs from accommodative enforcement in that it requires high effort, moderate stringency, and targeted enforcement while accommodative enforcement requires low effort, low stringency, and haphazard enforcement.

A flexible approach is "legalistic in some cases [and] accommodative and helpful in others" (Bardach & Kagan 1982; Kagan 1989:92). This mixing of strategies can be viewed as more effective than deterrence-based methods, and as long as the "threat of punishment" exists, cooperation is often more cost-effective (Coglianese & Kagan 2007:xviii). Broadly, flexible enforcement can be categorized as insistent or persuasive (Baldwin et al. 2012; Hutter 1989; May & Winter 2000). Insistent enforcers are less flexible than persuasive ones, emphasizing set limits to flexibility, while persuasive enforcers are much more patient. Persuasion is often viewed as the preferred starting point in flexible enforcement, reserving punishment for failed attempts at persuasion or as a "back up" to persuasion (Ayres & Braithwaite 1992; Braithwaite 2002b).

Flexible enforcement adopts a tit-for-tat strategy where inspectors determine their actions in response to the actions of regulated entities. If the regulated entity cooperates, so too does the regulator. Alternatively, if the firm does not cooperate, deterrence methods may be employed (Scholz 1984b). This model builds from Axelrod (1984) which illustrates using game-theory that a tit-for-tat (where each player starts with a cooperative move and then mirrors the other player's last move) strategy can result in successful cooperation. Ultimately, the costs and benefits of the regulator, regulated entity, and arguably society are determined by the actions of the other party (Scholz 1984a).

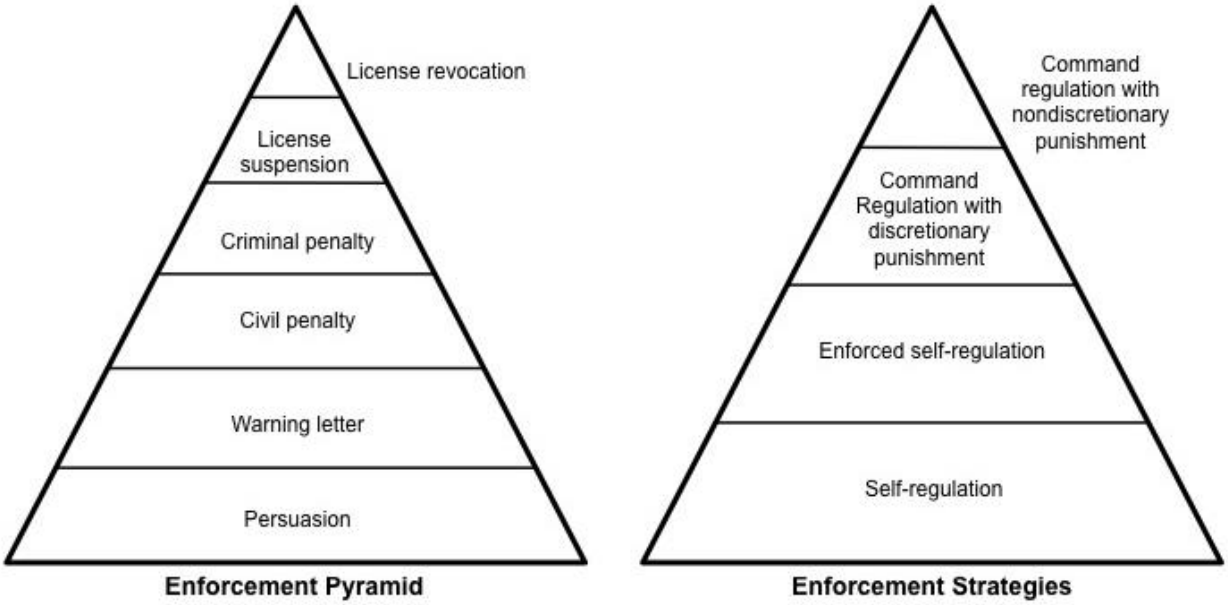
In general, firms "will always be better off cooperating when the value of future cooperative payoffs is high enough to compensate for the short term temptation to cheat" (Scholz 1984a:392). Given a violation, in a tit-for-tat approach, a regulator will respond immediately to the violation with a more punitive approach but will forgive past infractions once a firm adopts a cooperative approach. The job of a regulator is to adopt an enforcement strategy that: 1) rewards "good" firms with cooperative enforcement; 2) punishes firms that violate regulations in order to encourage future compliance; and 3) responds to "good faith efforts" in a manner that encourages firms to cooperate, even if they have previous violations (Scholz 1984a:394).

Ayres and Braithwaite (1992) term a regulator's tit-for-tat behavior under a cooperative enforcement strategy as responsive regulation. Responsive regulation is a strategy dependent "on context, regulatory culture, and history," rather than a set of prescriptive actions. "Responsiveness is rather an attitude that enables the blossoming of a wide variety of regulatory approaches" (Ayres & Braithwaite 1992:5). As a precursor to Braithwaite (1985) and Ayres and Braithwaite (1992), Hawkins (1984:5) notes that "decisions in a compliance system are graduated in character, and though in rare cases matters are ultimately settled by adjudication, they are normally controlled by parties themselves in private, intimate negotiations which rely on bargaining, not adjudication."

In order to facilitate and visualize making graduated decisions, Braithwaite (1985) introduced the use of an enforcement pyramid. Figure 1 illustrates an example of an enforcement pyramid and related enforcement strategies as presented by Morgan & Yeung (2007) and based on Ayers and Braithwaite's (1992) vision of the pyramid. The base of the pyramid represents cooperative approaches (e.g., education and persuasion) with punitive or legalistic responses escalating up the pyramid. The relative size of each section indicates the proportion of enforcement activity at that level. Mechanisms are available to the regulator to encourage better behavior, such as forbearance and information provision (Bardach & Kagan 1982:131). Generally, regulators only take formal legal action toward a small percentage of violations (at the top of the pyramid). Rather, they apply less formal tactics to address more frequent violations (at the bottom of pyramid). Nevertheless, "regulatory officials and regulatory scholars believe that governmental capacity to impose severe legal penalties, together with relatively frequent use of that capacity, is crucial to the implementation of regulatory norms" (Thornton et al. 2005:262).

Given knowledge of an escalating strategy, a firm is more likely to cooperate with an enforcement pyramid as opposed to the threat of a single deterrence strategy since, in reality, regulators may not be able to fully enforce a single extreme measure (Morgan & Yeung 2007:196–7). While the enforcement pyramid illustrates a responsive strategy toward an individual firm, the pyramid of enforcement strategies offers an illustration of escalating enforcement strategies toward the industry.

Figure 1 – Sample Enforcement and Enforcement Strategy Pyramids



Source: Morgan and Yeung (2007:196-8).

4. Empirical Evidence on Styles

In some of the first empirical work comparing a deterrence-based approach and a cooperative approach, Braithwaite et al. (1987) identified seven types of regulators based on a sample of Australian regulatory agencies: conciliators, benign big guns, diagnostic inspectorates, detached token enforcers, detached modest enforcers, token enforcers, and modest enforcers. Notably, Braithwaite et al. (1987) identify conciliators as a unique group of enforcers who primarily rely on conflict resolution to resolve conflicts rather than any kind of legalistic approach. In their sample, all conciliatory agencies were less than twenty years old, offering some support for a state-level shift toward cooperative enforcement (Braithwaite et al. 1987).

Like Braithwaite et al. (1987), May and Winter (2000) test their hypotheses of varying levels of formalism and coercion using enforcement data of agro-environmental policies in Denmark. Results indicate the formalism has a greater influence on enforcement style, but both formalism and coercion are part of enforcement. Yet, enforcement styles did not conform specifically to the three proposed styles (legalistic, accommodative, and flexible), supporting the notion that predefined "ideal" styles of enforcement are overly restrictive and a continuum may better represent how styles manifest themselves in practice. May and Winter (2000) do find support for the hypothesis that flexible enforcement requires more effort from inspectors than other enforcement styles.

While these studies offer some evidence for the use of different enforcement styles, there is limited empirical consensus supporting the impact of a cooperative regulatory strategy on outcomes. Harrison (1995) examines the Canadian pulp and paper industry in comparison to the U.S. industry in order to ascertain if cooperative or deterrence-based enforcement is more

effective. Canadian environmental enforcement is generally cooperative, typical responses to environmental violations include establishing a schedule to bring violators into compliance – accounting for financial capacity and severity of pollution. The Canadian Fisheries Act established uniform standards based on best available technology, but unlike U.S. regulations, standards were established through negotiations with the federal and provincial governments and industry. Based on discharges from 1987, Harrison (1995:236) concludes that implementation of Canadian standards was a "profound failure." Regulators spent significant effort renegotiating compliance schedules and lowering provincial and federal standards rather than requiring performance to meet water quality standards.

Additionally, compared to the results presented by Magat and Viscusi (1990) on U.S. paper and pulp compliance, Harrison (1995) determined that Canadian facilities are much less (17 to 29 percent) compliant with standards. Still, these results may be partially explained by differences in plant size and self-reported data quality between Canada and the U.S. facilities. Additionally, Canadian enforcement only partially achieves responsive regulation as explained above, since cooperative agencies rarely respond to infractions with appropriate punishment.

Hunter and Waterman (1992) examine the U.S. Environmental Protection Agency's (EPA) enforcement style based on National Pollutant Discharge Elimination System data to better understand how the Office of Water enforces the U.S. Clean Water Act. They distinguish enforcement styles as either enforced compliance or negotiated compliance. Regulatory responsibility for the Clean Water Act is relatively diffuse, separated among ten EPA regional offices, in which each has unique attributes related to water quality issues, political culture, and political environment. Additionally, in thirty-seven states, state authorities, not EPA, have primary implementation authority. Clean Water Act enforcement actions range from formal enforcement mechanisms (e.g., administrative orders, cease and desist orders, consent decrees, and penalties) to informal actions (e.g., comments, warning letters, phone calls, meetings, conferences, and determinations that no action is warranted). Hunter and Waterman (1982) divide these actions into formal and informal actions, where formal actions serve as indicators of enforced compliance and informal actions as indicators of negotiated compliance. With over 70 percent of EPA actions categorized as informal, this study indicates that Clean Water Act enforcement primarily relies on negotiated compliance with variation based on EPA region. While these results comment on the enforcement model employed by the EPA, the do not provide information on the ultimate water quality outcomes.

Contrary to earlier critiques of U.S. regulatory enforcement, increasingly studies find that agencies use a flexible approach (Rechtschaffen 2004). Earnhart and Glicksman (2006) examine the perceived enforcement approach of the Clean Water Act by surveying individuals at regulated facilities. Results indicate that enforcement is perceived to be responsive, with some interactions cooperative and some coercive. To understand the general perception of enforcement, survey respondents selected from three options: generally coercive, generally cooperative, or don't know. Overwhelmingly, respondents selected generally cooperative, with 96.2 percent answering generally cooperative while only 2.7 percent responding that

Case Study: U.S. Environmental Protection Agency's Compliance Assistance Program

The U.S. Environmental Protection Agency (EPA) has had a formal compliance assistance program since 1994 with the creation of the Office of Enforcement and Compliance Assurance. Over the past two decades, the importance of compliance enforcement has increased as indicated by the degree of focus it commands in EPA's strategic planning documents. In 1997, compliance assistance was mentioned briefly. In 2000, it increased in importance but was presented as a companion tool to more traditional enforcement. Yet, by 2003, "EPA's Strategic Plan identified compliance assistance as the first method through which EPA intended to increase compliance, followed by compliance incentives and finally, enforcement" (Stafford 2012:534). The EPA began collecting data on regional compliance assistance efforts in 2005. Compliance assistance is intended to help regulated entities "understand and meet their environmental compliance obligations or voluntarily adopt environmentally beneficial practices that go beyond compliance," through activities such as technical assistance or other support (Stafford 2012:535). Assistance is not limited to one-on-one interactions but includes print and online information, trainings, and workshops. Compliance assistance should not be confused with compliance incentives, which encourage compliance through direct rewards. Compliance assistance programs can target specific industries or regions by way of EPA regional offices. Often compliance assistance programs target small-to-medium sized business and leverage partnerships with other institutions such as the U.S. Small Business Administration.

Stafford (2012) explores two questions in her analysis. The first is whether EPA's compliance assistance program increased compliance, and the second is, as a cooperative strategy, whether compliance assistance compliments EPA's more traditional, deterrence-based enforcement. Stafford (2012) used data on the EPA's hazardous waste program to answer these questions. Her approach allows for spatial variation in compliance assistance programs, as well as variation based on firm characteristics such as size and industry. Additionally, Stafford (2012) differentiates between specific and general deterrence using state-level inspection and penalty data to assess general deterrence, and individual-level inspection data to assess specific deterrence. Given the potential influence of facility characteristics on compliance, the study further separates hazardous waste facilities into three groups: large waste quantity (greater than 1,000 kg/mo.), small waste quantity (100 to 1,000 kg/mo.), and conditionally exempt generators (less than 100 kg/mo.).

Stafford's (2012) analysis indicates that regional compliance assistance programs reduce the probability of violations in smaller waste facilities but not in those with large waste quantities, supporting other research that size is an important factor for compliance. Furthermore, facilities with larger amounts of waste are more likely to violate and facilities with multiple environmental regulations are more likely to violate. These results support the theory that managerial style or attitude may influence compliance. Lastly at the state level, facilities in states with successful environmental groups are more likely to comply. While regional compliance assistance programs influence the probability of violations, they do not reduce the probability that a facility will be inspected. As a cooperative approach, compliance assistance increases compliance with hazardous waste regulations overall, but regional programs are not a replacement for deterrence-based strategies. On the other hand, state-level compliance assistance programs do appear to function as substitutes for inspections, indicating a trade-off at the state-level between enforcement strategies.

As presenting one of the only empirical studies on both compliance levels and enforcement styles, Stafford (2012) offers important takeaways for compliance assistance programs. Size appears to be important, as compliance assistance may either not be available for or useful to larger facilities. Regulatory strategies that include larger facilities should carefully consider if compliance assistance programs address their most pressing barriers to compliance.

enforcement was generally coercive (Earnhart & Glicksman 2006:23). Beyond general perceptions, Earnhart and Glicksman (2006) compared individual responses to different questions such as their likelihood to allow the regulator into the facility unannounced, the treatment of regulated entity by the regulator, and the type of interaction with the regulaor. The results indicate that there are mutliple dimensions to the relationship between the regulator and regulated entity, which sometimes end up being cooperative and sometimes legalistic.

E. Next Generation Enforcement and Compliance

Possibly because limited consensus exists in the regulatory literature regarding which regulatory strategies – from among the large number of regulatory enforcement mechanisms available – will most often achieve the desired regulatory outcomes in the most efficient manner, strategies in practice tend to rely on a mix of the elements discussed. May and Burby (1998:158) argue that the existing regulatory enforcement scholarship is plagued by three limitations that include: (1) disagreement over what constitutes an enforcement strategy; (2) lack of consideration of the distinction between the actions of a regulatory agency versus the individual inspector; and (3) "ambiguity in the intent of a given study for which it is easy to confuse assumptions with prescriptions." Deviations from traditional enforcement approaches (in practice and in empirical research) are increasingly common. Experts, regulators, and organizations are identifying combinations of problem-solving tools and strategies that can improve regulatory programs and enforcement. The growing concept of "next generation" enforcement reinforces the idea that a regulator can select from a suite of options in order to take advantage of increases in data availability, new technology, and a variety of enforcement tactics.

The U.S. Environmental Protection Agency attributes five components to "next generation" compliance: more effective rules; advanced monitoring and technology; electronic reporting; increased transparency; and innovative enforcement (EPA 2015). While not explicitly "next generation," the Organization for Economic Cooperation and Development (OECD) offers eleven principles for improving regulatory enforcement and inspections: evidence-based enforcement; selectivity; risk focus and proportionality; responsive regulation; long-term vision; coordination and consolidation; transparent governance; information integration; clear and fair process; compliance promotion; and professionalism. Appendix B provides the OECD definitions of these principles.

The New York City Department of Health and Mental Hygiene (DOHMH) offers a promising example of "next generation" enforcement through the use of online consumer reviews posted to the website "Yelp" to improve restaurant regulation and identify unreported cases of foodborne illnesses (Harrison et al. 2014). Recognizing that consumers report food-related illnesses on Yelp, DOHMH decided to test if using these reviews could help them identify unreported instances of foodborne illnesses. By searching for terms indicative of illness including words like "sick" and "vomit," DOHMH was able to identify almost 500 reviews identifying foodborne illnesses, of which only three percent of those cases had also been reported to the agency. Cases such as this illustrate the growing number of new opportunities for regulators to combine approaches to improve enforcement strategies in order to better accomplish regulatory goals.

Conclusion

This paper has illuminated the array of options at the disposal of regulators in both designing and enforcing regulatory programs. The possibilities available to regulators to structure their regulatory programs are varied, ranging from specific means-based and performance standards to much more flexible voluntary approaches and self-regulation, with a number of options falling in between. In presenting these possibilities, we have detailed what is known about the effectiveness of each of the methods as well as the regulatory outcomes that they are better and worse at achieving. In a similar way, we have discussed the variety of issues that affect how regulators go about enforcing regulatory programs. We have contrasted legalistic and cooperative approaches as well as described how the two can be sometimes fruitfully combined in responsive regulation.

Our approach, which has been to lay out the breadth of options to consider in solving regulatory problems, coupled with the volume of material presented, does not easily lend itself to simple takeaways. Nevertheless, our review has clearly demonstrated that the appropriateness of a particular regulatory approach is contingent on the circumstances and goals of the regulator. Out of the multitude of possible regulatory goals outlined at the beginning of the paper, including cost-effectiveness, flexibility, and administrative feasibility, the theoretical and empirical evidence has clearly shown that no instrument or enforcement approach can achieve all of them simultaneously. For example, relative to traditional means-based standards, market-based mechanisms such as tradeable permits offer the potential to achieve regulatory goals more cost effectively. However, they may do so at the expense of manageability and administrative feasibility by requiring complex program designs and more intricate monitoring technologies. In the enforcement context, use of targeting can decrease the costs to regulators of encouraging compliance, but in doing, the agency may be more vulnerable to industry group pressure and, with it, the specter of becoming captured by its regulated entities.

The fact that no enforcement strategy or regulatory instrument can achieve all or even a majority of regulatory goals simultaneously suggests that regulators may have to give up some objectives to achieve others in designing their programs. Still, our review has simultaneously shown that these options can sometimes be combined to achieve regulatory goals. Perhaps the clearest illustration can be found in responsive regulation, which attempts to blend legalistic and accommodative approaches to realize the advantages of each. Yet, regulatory instruments can also be combined to better achieve regulatory objectives. For instance, research on the efficacy of mandated information disclosure has demonstrated that it is most apt to be able to achieve regulatory goals when it used as a supplement to traditional regulatory mechanisms.

A clear understanding of both the tradeoffs associated with different regulatory instruments and enforcement approaches as well as the potential for fruitful combinations underscores the need for a regulator to cast a wide net when considering regulatory alternatives. Not only can serious consideration of the great variety of options that exist give regulators the best chance of choosing those most suited for the particular circumstances at hand, but doing so can also position these same regulators to more effectively exploit the synergies that exist when the right regulatory strategies are used in tandem.

Notes

Ackerman, Bruce A., & Richard B. Stewart (1985) "Reforming Environmental Law," 37 *Stanford Law Rev.* 1333–65.

Akerlof, George A., (1970) "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism," 84 *Quarterly J. of Economics* 488–500.

Anderson, Soren T., et al. (2011) "Automobile Fuel Economy Standards: Impacts, Efficiency, and Alternatives," 9 *Rev. of Environmental Economics and Policy* 89–108.

Axelrod, Robert (1984) *The Evolution of Cooperation*. New York: Basic Books, Inc.

Ayres, Ian, & John Braithwaite (1992) *Responsive Regulation: Transcending the Deregulation Debate*. Oxford: Oxford University Press.

Baldwin, Robert, et al. (2012) *Understanding Regulation: Theory, Strategy, and Practice*. New York: Oxford University Press.

Bardach, Eugene, & Robert A. Kagan (1982) *Going by the Book: The Problem of Regulatory Unreasonableness*. Philadelphia, PA: Temple University Press.

Becker, Gary S. (1968) "Crime and Punishment: An Economic Approach," 76 *J. of Political Economy* 169–217.

Ben-Shahar, Omri, & Carl E. Schneider (2011) "The Failure of Mandated Disclosure," 159 *University of Pennsylvania Law Rev.* 647–749.

——— (2014) *More than You Wanted to Know: The Failure of Mandated Disclosure*. Princeton, NJ: Princeton University Press.

Bennear, Lori S. (2006) "Evaluating Management-Based Regulation: A Valuable Tool in the Regulatory Toolbox?" in C. Coglianese & J. Nash, eds., *Leveraging the Private Sector: Management-Based Strategies for Improving Environmental Performance*. Washington, DC: Resources for the Future Press.

——— (2007) "Are Management-Based Regulations Effective? Evidence from State Pollution Prevention Programs," 26 *J. of Policy Analysis and Management* 327–48.

——— (2008) "What Do We Really Know? The Effect of Reporting Thresholds on Inferences Using Environmental Right-to-Know Data," 2 *Regulation & Governance* 293–315.

——— (2015) "Offshore Oil and Gas Drilling: A Review of Regulatory Regimes in the United States, United Kingdom, and Norway," 9 *Rev. of Environmental Economics and Policy* 2–22.

Bennear, Lori S., & Cary Coglianese (2012) "Flexible Approaches to Environmental Regulation," in S. Kamieniecki & M. Kraft, eds., *The Oxford Handbook of U.S. Environmental Policy*. Oxford: Oxford University Press.

Bennear, Lori S., & Sheila M. Olmstead (2008) "*The Impacts of the "Right to Know": Information Disclosure and the Violation of Drinking Water Standards*," 56 *J. of Environmental Economics and Management* 117–30.

Bernstein, Marver H. (1955) *Regulating Business by Independent Commission*. Princeton, NJ: Princeton University Press.

Black, Julia, & Robert Baldwin (2010) "Really Responsive Risk-Based Regulation," 32 *Law & Policy* 181–213.

Bó, Ernesto Dal (2006) "Regulatory Capture: A Review," 22 *Oxford Rev. of Economic Policy* 203–25.

Bohm, Peter, & Clifford S. Russell (1985) "Comparative Analysis of Alternative Policy Instruments," in A. Kneese, & J. Sweeney, eds., *Handbook of Natural Resource and Energy Economics*. New York: Elsevier.

Borck, Jonathan C., & Cary Coglianese (2009) "Voluntary Environmental Programs: Assessing Their Effectiveness," 34 *Annual Rev. of Environment and Resources* 305–24.

Borck, Jonathan C., et al. (2008) "Environmental Leadership Programs: Toward an Empirical Assessment of their Performance," 35 *Ecology Law Quarterly* 771–834.

Boyum, Keith O. (1983) "Review: The Politics of 'Regulatory Unreasonableness': Bardach and Kagan's 'Going by the Book,'" 8 *American Bar Foundation Research J.* 752–60.

Braithwaite, John (1985) *To Punish or Persuade: Enforcement of Coal Mine Safety*. Albany, NY: SUNY Press.

——— (2002a) *Restorative Justice & Responsive Regulation*. Oxford: Oxford University Press.

——— (2002b) "Rewards and Regulation," 29 *J. of Law and Society* 12–26.

Braithwaite, John, & Toni Makkai (1991) "Testing an Expected Utility Model of Corporate Deterrence," 25 *Law & Society Rev.* 7–40.

Braithwaite, John, et al. (1987) "An Enforcement Taxonomy of Regulatory Agencies," 9 *Law & Policy* 323–351.

Breyer, Stephen G. (1982) *Regulation and Its Reform*. Cambridge, MA: Harvard University Press.

Brunel, Claire, & Arik Levinson (2013) "Measuring Environmental Regulatory Stringency." OECD Trade and Environment Working Papers 2013/05, Paris.

Burtraw, Dallas, & Byron Swift (1996) "A New Standard of Performance: An Analysis of the Clean Air Act's Acid Rain Program," 26 *Environmental Law Reporter News and Analysis*, 10411-23.

Carlough, Les (2004) *General Deterrence Of Environmental Violation: A Peek Into The Mind Of The Regulated Public*. State of Oregon Department of Environmental Quality. Portland, OR.

Carpenter, Daniel, & David A. Moss (2014) "Introduction," in D. Carpenter & D. Moss, eds., *Preventing Regulatory Capture: Special Interest Influence, and How to Limit It*. New York: Cambridge University Press.

Carrigan, Christopher (2014) "Captured by Disaster? Reinterpreting Regulatory Behavior in the Shadow of the Gulf Oil Spill," in D. Carpenter & D. Moss, eds., *Preventing Regulatory Capture: Special Interest Influence, and How to Limit It*. New York: Cambridge University Press.

——— (forthcoming) "Reform in Real Time: Evaluating Reorganization as a Response to the Gulf Oil Spill," in E. Balleisen, L. Bennear, K. Krawiec, & J. Wiener, eds., *Policy Shock: Regulatory Responses to Oil Spills, Nuclear Accidents, and Financial Meltdowns*. Cambridge: Cambridge University Press.

Carrigan, Christopher, & Cary Coglianese (2011) "The Politics of Regulation: From New Institutionalism to New Governance," 14 *Annual Rev. of Political Science* 107–29.

——— (2012) "Oversight in Hindsight: Assessing the U.S. Regulatory System in the Wake of Calamity," in C. Coglianese, ed., *Regulatory Breakdown: The Crisis of Confidence in U.S. Regulation*. Philadelphia, PA: University of Pennsylvania Press.

——— (2015) "George J. Stigler, 'The Theory of Economic Regulation'," in S. Balla, M. Lodge, & E. Page, ed., *The Oxford Handbook of Classics in Public Policy and Administration*. Oxford: Oxford University Press.

Coglianese, Cary, & Robert A. Kagan (2007) "Introduction," in C. Coglianese & R. Kagan, eds., *Regulation and Regulatory Processes*. Burlington, VT: Ashgate Publishing.

Coglianese, Cary, & David Lazer (2003) "Management-Based Regulation: Prescribing Private Management to Achieve Public Goals," 37 *Law & Society Rev.* 691–730.

Coglianese, Cary, & Evan Mendelson (2010) "Meta-Regulation and Self-Regulation," in M. Cave et al., eds., *The Oxford Handbook on Regulation*. Oxford: Oxford University Press.

Coglianese, Cary, & Jennifer Nash (2014) "Performance Track's Postmortem: Lessons from the Rise and Fall of EPA's 'Flagship' Voluntary Program," 38 *Harvard Environmental Law Rev.* 1–86.

Coglianese, Cary, et al. (2003) "Performance-Based Regulation: Prospects and Limitations in Health, Safety, and Environmental Protection," 55 *Administrative Law Rev.* 705–29.

Cohen, Mark A. (1987) "Optimal Enforcement Strategy to Prevent Oil Spills: An Application of a Principal-Agent Model with Moral Hazard," 30 *J. of Law and Economics* 23–51.

——— (1998) "Monitoring and Enforcement of Environmental Policy," in T. Tietenber & J. Folmer, eds. *International Yearbook of Environmental and Resource Economics, Volume III*. Cheltenham, UK: Edward Elgar.

Cordes, Joseph J. (2002) "Corrective Taxes, Charges, and Tradable Permits," in L. Salamon & O. Elliott, eds., *The Tools of Government: A Guide to the New Governance*. Oxford: Oxford University Press.

Davis, Lucas W. (2015) "Bonding Requirements for U.S. Natural Gas Producers," 9 *Rev. of Environmental Economics and Policy* 128–44.

Delmas, Magali A., & Ann K. Terlaak (2001) "A Framework for Analyzing Environmental Voluntary Agreements," 43 *California Management Rev.* 44–63.

de Marchi, Scott, & James T. Hamilton (2006) "Assessing the Accuracy of Self-Reported Data: An Evaluation of the Toxics Release Inventory," 32 *J. of Risk and Uncertainty* 57–76.

Driesen, David M. (1998) "Is Emissions Trading an Economic Incentive Program: Replacing the Command and Control/Economic Incentive Dichotomy," 55 *Washington and Lee Law Rev.* 289–350.

Earnhart, Dietrich (2004) "Regulatory Factors Shaping Environmental Performance at Publicly-Owned Treatment Plants," 48 *J. of Environmental Economics and Management* 655–81.

Earnhart, Dietrich, & Robert L. Glicksman (2006) "Depiction of the Regulator-Regulated Entity Relationship in the Chemical Industry: Deterrence-Based vs. Cooperative Enforcement," 31 *William & Mary Environmental Law & Policy Rev.* 603–60.

Earnhart, Dietrich, & Robert Glicksman (2011) *Pollution Limits and Polluters' Efforts to Comply: The Role of Government Monitoring and Enforcement*. Palo Alto, CA: Stanford University Press.

Ferrell, Allen (2007) "Mandatory Disclosure and Stock Returns: Evidence from the Over-the-Counter Market," 36 *J. of Legal Studies* 213–51.

Finkel, Adam M. (2013) “Updating OSHA Inspection Policies,” Harvard Law Petrie-Flom Center, http://blogs.law.harvard.edu/billofhealth (assessed 24 June 2015).

Finkel, Adam M., et al. (2015) “Planning for Excellence: Insights from an International Review of Regulators’ Strategic Plans,” Penn Program on Regulation Best-in-Class Regulator Research Paper, Philadelphia, PA.

Friedland, Martin L., et al. (1990) “Regulating Traffic Saftey,” in M. L. Friedland, ed., *Securing Compliance: Seven Case Studies*. Toronto: University of Toronto Press.

Friesen, Lana (2003) “Targeting Enforcement to Improve Compliance with Environmental Regulations,” 46 *J. of Environmental Economics and Management* 72–85.

Fung, Archon, & Dara O’Rourke (2000) “Reinventing Environmental Regulation from the Grassroots Up: Explaining and Expanding the Success of the Toxics Release Inventory,” 25 *Environmental Management* 115–27.

Fung Archon, et al. (2007) *Full Disclosure: The Perils and Promise of Transparency*. New York: Cambridge University Press.

Gamper-Rabindran, Shanti (2006) “Did the EPA’s Voluntary Industrial Toxics Program Reduce Emissions? A GIS Analysis of Distributional Impacts and by-Media Analysis of Substitution,” 52 *J. of Environmental Economics and Management* 391–410.

Gerard, David (2000) “The Law and Economics of Reclamation Bonds,” 26 *Resources Policy* 189–97.

Gerard, David, & Lester B. Lave (2005) “Implementing Technology-Forcing Policies: The 1970 Clean Air Act Amendments and the Introduction of Advanced Automotive Emissions Controls in the United States,” 72 *Technological Forecasting and Social Change* 761–78.

Graham, Mary (2002) *Democracy by Disclosure: The Rise of Technopopulism*. Washington, DC: Brookings Institution.

Grau, Montserrat Viladrich, & Theodore Groves (1997) “The Oil Spill Process: The Effect of Coast Guard Monitoring on Oil Spills,” 10 *Environmental and Resource Economics* 315–39.

Gray, Wayne B., & Mary E. Deily (1996) “Compliance and Enforcement: Air Pollution Regulation in the U.S. Steel Industry,” 31 *J. of Environmental Economics and Management* 96–111.

Gray, Wayne B., & John T. Scholz (1990) “OSHA Enforcement and Workplace Injuries: A Behavioral Approach to Risk Assessment,” 3 *J. of Risk and Uncertainty* 283–305.

——— (1993) “Does Regulatory Enforcement Work? A Panel Analysis of OSHA Enforcement,” 27 *Law & Society Rev.* 177–213.

Gray, Wayne B., & Ronald J. Shadbegian (2005) “When and Why Do Plants Comply? Paper Mills in the 1980s,” 27 *Law & Policy* 238–61.

——— (2007) “The Environmental Performance of Polluting Plants: A Spatial Analysis,” 47 *J. of Regional Science* 63–84.

Greenstone, Michael (2004) “Did the Clean Air Act Cause the Remarkable Decline in Sulfur Dioxide Concentrations?,” 47 *J. of Environmental Economics and Management* 585–611.

Gunningham, Neil (1995) “Environment, Self-Regulation, and the Chemical Industry: Assessing Responsible Care,” 17 *Law & Policy* 57–110.

Gunningham, Neil A., et al. (2003) *Shades of Green: Business, Regulation, and Environment*. Palo Alto, CA: Stanford University Press.

——— (2005) “Motivating Management: Corporate Compliance in Environmental Protection,” 27 *Law & Policy* 289–316.

Hamilton, James T. (2005) *Regulation through Revelation: The Origin, Politics, and Impact of the Toxics Release Inventory Program*. Cambridge, UK: Cambridge University Press.

Harrington, Winston (1988) “Enforcement Leverage When Penalties Are Restricted,” 37 *J. of Public Economics* 29–53.

Harrison, Cassandra, et al. (2014) *Using Online Reviews by Restaurant Patrons to Identify Unreported Cases of Foodborne Illness – New York City, 2012–2013* Centers for Disease Control and Prevention Morbidity and Mortality Weekly Report Vol.63, No. 20. Washington, DC: GPO.

Harrison, Kathryn (1995) “Is Cooperation the Answer? Canadian Environmental Enforcement in Comparative Context,” 14 *J. of Policy Analysis and Management* 221–44.

Haufler, Virginia (2001) *A Public Role for the Private Sector*. Washington, DC: Carnegie Endowment for International Peace.

Hawkins, Keith (1984) *Environment and Enforcement: Regulation and the Social Definition of Pollution*. Oxford University Press, Incorporated.

Helland, Eric (1998a) “The Enforcement of Pollution Control Laws: Inspections, Violations, and Self-Reporting,” 80 *Rev. of Economics and Statistics* 141–53.

——— (1998b) “The Revealed Preferences of State EPAs: Stringency, Enforcement, and Substitution,” 35 *J. of Environmental Economics and Management* 242–61.

Howard, Jennifer, et al. (2000) "Standard or Smokescreen? Implementation of a Voluntary Environmental Code," 42 *California Management Rev.* 63–82.

Howard-Grenville, Jennifer, et al. (2008) "Constructing the License to Operate: Internal Factors and Their Influence on Corporate Environmental Decisions," 30 *Law & Policy* 73–107.

Hunter, Susan, & Richard W. Waterman (1992) "Determining an Agency's Regulatory Style: How Does the EPA Water Office Enforce the Law?" 45 *Western Political Quarterly* 403–17.

Huntington, Samuel P. (1952) "The Marasmus of the ICC: The Commission, the Railroads, and the Public Interest," 61 *Yale Law J.* 467–509.

Huppes, Gjalt, & Robert A. Kagan (1989) "Market-Oriented Regulation of Environmental Problems in the Netherlands," 11 *Law & Policy* 215–39.

Hutter, Bridget M. (1989) "Variations in Regulatory Enforcement Styles," 11 *Law & Policy* 153–74.

——— (2001) *Regulation and Risk: Occupational Health and Safety on the Railways*. Oxford: Oxford University Press.

Innes, Robert, & Abdoul G. Sam (2008) "Voluntary Pollution Reductions and the Enforcement of Environmental Law: An Empirical Study of the 33/50 Program," 51 *J. of Law and Economics* 271–96.

International Organization for Standardization (ISO) (2009) *Environmental Management: The ISO 14000 Family of International Standards*. Geneva: ISO Central Secretariat.

Jin, Ginger Zhe, & Phillip Leslie (2003) "The Effect of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards," 118 *Quarterly J. of Economics* 409–51.

Kaasen, Knut (1991) "Post Piper Alpha: Some Reflections on Offshore Safety Regimes from a Norwegian Perspective," 9 *J. of Energy & Natural Resources Law* 281–9.

Kagan, Robert A. (1989) "Editor's Introduction: Understanding Regulatory Enforcement," 11 *Law & Policy* 89–119.

——— (1994) "Regulatory Enforcement," in D. Rosenbloom & R. Schwartz, eds., *Handbook of Regulation and Administrative Law*. New York: Marcel Dekker.

Karkkainen, Bradley C. (2001) "Information as Environmental Regulation: TRI and Performance Benchmarking, Precursor to a New Paradigm?" 89 *Georgetown Law J.* 257–370.

Keenan, Cheryl, et al. (1997) *Survey Evaluation of the Massachusetts Toxics Use Reduction Program.* TURI Methods and Policy Report No. 14. Lowell, MA: The Massachusetts Toxics Use Reduction Institute, University of Massachusetts, Lowell.

Keohane, Nathaniel O., et al. (1998) "The Choice of Regulatory Instruments in Environmental Policy," 22 *Harvard Environmental Law Rev.* 313–67.

Khanna, Madhu (2007) "The US 33/50 Voluntary Program: Its Design and Effectiveness," in R. D. Morgenstern and W. A. Pizer, eds., *Reality Check: The Nature and Performance of Voluntary Environmental Programs in the United States, Europe, and Japan*. Washington, DC: Resources for the Future.

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

King, Andrew A., & Michael J. Lenox (2000) "Industry Self-Regulation without Sanctions: The Chemical Industry's Responsible Care Program," 43 *Academy of Management J.* 698–716.

Kleiman, Mark, et al. (2009) "The Dynamics of Deterrence," 106 *Proceedings of the National Academy of Sciences of the United States of America* 14230–35.

Kleindorfer, Paul R., & Eric W. Orts (1998) "Informational Regulation of Environmental Risks," 18 *Risk Analysis* 155–70.

Ko, Kilkon, et al. (2010) "The Role of Inspection Sequence in Compliance with the US Occupational Safetyand Health Administration's (OSHA) Standards: Interpretations and Implications," 4 *Regulation & Governance* 48–70.

Konar, Shameek, & Mark A. Cohen (1997) "Information As Regulation: The Effect of Community Right to Know Laws on Toxic Emissions," 32 *J. of Environmental Economics and Management* 109–24.

Kraft, Michael E., et al. (2011) *Coming Clean: Information Disclosure and Environmental Performance*. Cambridge, MA: MIT Press.

Krebs, John R. (2011) "Risk, Uncertainty and Regulation," 369 *Philosophical Transactions: Mathematical, Physical and Engineering Sciences* 4842–52.

Kringen, Jacob (2013) "Contested Terrains in Risk Regulation: Robustness and Vulnerabilities in the Norwegian Petroleum Regime," in P. Lindoe et al., eds., *Risk Governance of Offshore Oil and Gas Operations*. Cambridge, UK: Cambridge University Press.

Laplante, Benoit, & Paul Rilstone (1996) "Environmental Inspections and Emissions of the Pulp and Paper Industry in Quebec," 31 *J. of Environmental Economics and Management* 19–36.

Leman, Christopher K. (2002) "Direct Government," in L. Salamon & O. Elliott, *The Tools of Government: A Guide to the New Governance*. Oxford: Oxford University Press.

Levi-Faur, David (2011) "Regulation and Regulatory Governance," in D. Levi-Faur, *Handbook on the Politics of Regulation*. Cheltenham, UK: Edward Elgar.

Lyon, Thomas P., & John W. Maxwell (2004) *Corporate Environmentalism and Public Policy*. Cambridge, UK: Cambridge University Press.

Magat, Wesley A., & W. Kip Viscusi (1990) "Effectiveness of the EPA's Regulatory Enforcement: The Case of Industrial Effluent Standards," 33 *J. of Law and Economics* 331–60.

Makkai, Toni, & John Braithwaite (1992) "In and Out of the Revolving Door: Making Sense of Regulatory Capture," 12 *J. of Public Policy* 61–78.

Manning, Peter K. (1989) "The Limits of Knowledge: The Role of Information in Regulation," in K. Hawkins & J. Thomas, eds., *Making Regulatory Policy*. Pittsburgh, PA: University of Pittsburgh Press.

May, Peter J. (2003) "Performance-Based Regulation and Regulatory Regimes: The Saga of Leaky Buildings," 25 *Law & Policy* 381–401.

May, Peter J., & Raymond Burby (1998) "Making Sense Out of Regulatory Enforcement," 20 *Law & Policy* 157–82.

May, Peter J., & Soren Winter (2000) "Reconsidering Styles of Regulatory Enforcement: Patterns in Danish Agro-Environmental Inspection," 22 *Law & Policy* 143–173.

McCaffrey, David P. (1983) "An Assessment of OSHA's Recent Effects on Injury Rates," 18 *J. of Human Resources* 131–146.

Meidinger, Errol E. (2003) "The New Environmental Law: Forest Certification," 10 *Buffalo Environmental Law J.* 211–300.

Morgan, Bronwen, & Karen Yeung (2007) *An Introduction to Law and Regulation: Text and Materials*. Cambridge, UK: Cambridge University Press.

Morgenstern, Richard D., & William A. Pizer (2007) "Concluding Observations: What Can We Learn from the Case Studies?" *Reality Check: The Nature and Performance of Voluntary Environmental Programs in the United States, Europe, and Japan*. Washington, DC: Resources for the Future.

Nadeau, Louis W. (1997) "EPA Effectiveness at Reducing the Duration of Plant-Level Noncompliance," 34 *J. of Environmental Economics and Management* 54–78.

Nash, Jennifer (2002) "Industry Codes of Practice: Emergence and Evolution," in T. Dietz & P. Stern, eds., *New Tools for Environmental Protection: Education, Information, and Voluntary Measures*. Washington, DC: National Academies Press.

Nash, Jennifer, & John Ehrenfeld (1997) "Codes of Environmental Management Practice: Assessing Their Potential as a Tool for Change," 22 *Annual Rev. of Energy and the Environment* 487–535.

Natan, Jr., Thomas E., & Catherine G. Miller (1998) "Are Toxics Release Inventory Reductions Real?" 32 *Environmental Science and Technology* 368A–74A.

Natan, Jr., Thomas E., et al. (1996) *Evaluation of the Effectiveness of Pollution Prevention Planning in New Jersey.* New Jersey Contract Number P39724. Alexandria, VA: Hampshire Research Associates, Inc.

National Commission on the BP Deepwater Horizon Oil Spill and Offshore Drilling (2011) *Deep Water: The Gulf Oil Disaster and the Future of Offshore Drilling*. Report to the President. Washington, DC.

Office of Inspector General (2012) *OSHA's Site Specific Targeting Program Has Limitations on Targeting and Inspecting High-Risk Worksites*. Washington, DC: U.S. Department of Labor.

Organization for Economic Cooperation and Development (OECD) (2014) *OECD Best Practice Principles for Regulatory Policy: Regulatory Enforcement and Inspections*. Paris: OECD Publishing.

Peltzman, Sam (1976) "Toward a More General Theory of Regulation," 19 *J. of Law and Economics* 211–40.

Perloff, Jeffrey M. (2001) *Microeconomics*. 2nd ed. Boston, MA: Addison-Wesley.

Poje, Gerald V., & Daniel M. Horowitz (1990) *Phantom Reductions: Tracking Toxic Trends*. Washington, DC: National Wildlife Federation.

Popp, David (2003) "Pollution Control Innovations and the Clean Air Act of 1990," 22 *J. of Policy Analysis and Management* 641–60.

Posner, Richard A. (1974) "Theories of Economic Regulation," 5 *Bell J. of Economics and Management Science* 335–58.

Potoski, Matthew, & Aseem Prakash (2005) "Green Clubs and Voluntary Governance: ISO 14001 and Firms' Regulatory Compliance," 49 *American J. of Political Science* 235–48.

Raymond, Mark (1999) "Enforcement Leverage When Penalties Are Restricted: A Reconsideration under Asymmetric Information," 73 *J. of Public Economics* 289–95.

Rechtschaffen, Clifford (2004) "Promoting Pragmatic Risk Regulation: Is Enforcement Discretion the Answer?" 52 *University of Kansas Law Rev.* 1327–61.

Rees, Joseph V. (1994) *Hostages of Each Other: The Transformation of Nuclear Safety since Three Mile Island*. Chicago: University of Chicago Press.

——— (1997) "Development of Communitarian Regulation in the Chemical Industry," 19 *Law & Policy* 477–528.

Reinfurt, Donald W. (2004) "Documenting the Sustainability of a Mature Click It or Ticket Program: The North Carolina Experience," 35 *J. of Safety Research* 181–88.

Richards, Kenneth (2000) "Framing Environmental Policy Instrument Choice," 10 *Duke Environmental Law & Policy Forum* 221–86.

Rothstein, Henry, et al. (2006) "The Risks of Risk-Based Regulation: Insights from the Environmental Policy Domain," 32 *Environment International* 1056–65.

Ruser, John W., & Robert S. Smith "Reestimating OSHA's Effects: Have the Data Changed?" 26 *J. of Human Resources* 212–35.

Russell, Clifford S. (1990) "Monitoring and Enforcement," in P. Portney, ed., *Environmental Regulation in the U.S.: Public Policies and Their Consequences.* Washington, DC: Resources for the Future.

Salamon, Lester M. (2002) "The New Governance and the Tools of Public Action: An Introduction," in L. Salamon, O. Elliott, *The Tools of Government: A Guide to the New Governance*. Oxford: Oxford University Press.

Scholz, John T. (1984a) "Voluntary Compliance and Regulatory Enforcement," 6 *Law & Policy* 385–404.

——— (1984b) "Cooperation, Deterrence, and the Ecology of Regulatory Enforcement," 18 *Law & Society Rev.* 179–224.

——— (1997) "Enforcement Policy and Corporate Misconduct: The Changing Perspective of Deterrence Theory," 60 *Law and Contemporary Problems* 253–68.

Segerson, Kathleen (1999) "Mandatory versus Voluntary Approaches to Food Safety," 15 *Agribusiness* 53–70.

Segerson, Kathleen, & Na Li (1999) "Voluntary Approaches to Environmental Protection," in T. Tietenberg & H. Folmer, eds., *The International Yearbook of Environmental and Resource Economics: 1999–2000*. Cheltenham, UK: Elgar.

Shimshack, Jay P. (2007) "Monitoring, Enforcement, & Environmental Compliance: Understanding Specific & General Deterrence," U.S. Environmental Protection Agency State-of-Science White Paper EPA07H001637, Washington, DC.

Shimshack, Jay P., & Michael B. Ward (2005) "Regulator Reputation, Enforcement, and Environmental Compliance," 50 *J. of Environmental Economics and Management* 519–40.

——— (2008) "Enforcement and Over-Compliance," 55 *J. of Environmental Economics and Management* 90–105.

Shogren, Jason F., et al. (1993) "Limits to Environmental Bonds," 8 *Ecological Economics* 109–33.

Shover, Neal, et al. (1984) "Regional Variation in Regulatory Law Enforcement: The Surface Mining Control and Reclamation Act of 1977," in K. Hawkins & J. Thomas, eds., *Enforcing Regulation*, Boston, MA: Kluwer-Nijhoff Publishing.

Smith, Robert Stewart (1979) "The Impact of OSHA Inspections on Manufacturing Injury Rates," 14 *J. of Human Resources* 145–170.

Sparrow, Malcolm K. (2000) *The Regulatory Craft: Controlling Risks, Solving Problems, and Managing Compliance*. Washington, DC: Brookings Institution Press.

Stafford, Sarah (2012) "Do Carrots Work? Examining the Effectiveness of EPA's Compliance Assistance Program: Effectiveness of EPA's Compliance Assistance Program," 31 *J. of Policy Analysis and Management* 533–55.

Stavins, Robert N. (1998) "What Can We Learn from the Grand Policy Experiment? Lessons from SO2 Allowance Trading," 12 *J. of Economic Perspectives* 69–88.

——— (2003) "Experience with Market-Based Environmental Policy Instruments," in K.-G. Maler and J. Vincent, eds., *Handbook of Environmental Economics, Volume 1: Environmental Degradation and Institutional Responses*. Amsterdam: Elsevier.

Stigler, George J. (1971) "The Theory of Economic Regulation," 2 *Bell J. of Economics and Management Science* 3–21.

Sunstein, Cass R. (1999) "Informational Regulation and Informational Standing: *Atkins* and Beyond," 147 *University of Pennsylvania Law Rev.* 613–75.

——— (2003) “Beyond the Precautionary Principle,” 151 *University of Pennsylvania Law Rev.* 1003–58.

Thaler, Richard H., & Cass R. Sunstein (2008) *Nudge: Improving Decisions About Health, Wealth, and Happiness*. New Haven, CT: Yale University Press.

Thornton, Dorothy, et al. (2005) “General Deterrence and Corporate Environmental Behavior,” 27 *Law & Policy* 262–88.

Tietenberg, Tom (2005) “Tradable Permits in Principle and Practice,” 14 *Penn State Environmental Law Rev.* 251-281.

U.S. Department of Transportation (1999) *Maritime Trade and Transportation 1999*. BTS99-02. Washington, DC: Bureau of Transportation Statistics.

U.S. Energy Information Administration (2015) *Annual Energy Outlook 2015 with Projections to 2040*. DOE/EIA-0383(2015). Washington, DC: U.S. Department of Energy.

U.S. Environmental Protection Agency (EPA) (1999) *33/50 Program: The Final Record*. EPA-745-99-004. Washington, DC.

——— (2015) “Next Generation Compliance,” U.S. Environmental Protection Agency, http://www2.epa.gov/compliance/next-generation-compliance (accessed 5 June 2015).

Vidovic, Martina, & Neha Khanna (2007) “Can Voluntary Pollution Prevention Programs Fulfill Their Promises? Further Evidence from the EPA’s 33/50 Program,” 53 *J. of Environmental Economics and Management* 180–95.

Wagner, Wendy E. (2000) “The Triumph of Technology-Based Standards,” 2000 *University of Illinois Law Rev.* 83–114.

Weil, David (1996) “If OSHA Is So Bad, Why Is Compliance So Good?” 27 *RAND J. of Economics* 618–40.

Weitzman, Martin L. (1974) “Prices vs. Quantities,” 41 *Rev. of Economic Studies* 477–91.

Williams, Allan F., & JoAnn K. Wells (2004) “The Role of Enforcement Programs in Increasing Seat Belt Use,” 35 *J. of Safety Research* 175–80.

Winston, Clifford (2008) “The Efficacy of Information Policy: A Review of Archon Fung, Mary Graham, and David Weil’s *Full Disclosure*: *The Perils and Promise of Transparency*,” 46 *J. of Economic Literature* 704–17.

Yeung, Karen (2004) *Securing Compliance: A Principled Approach*. Oxford: Hart Publishing.

Appendix A:
Factors that Influence a Regulator's Enforcement Style

1. Legal design factors
    a. Stringency of regulatory mission
    b. Legal powers
        i. Ex ante/ex post controls
        ii. Potency and immediacy of sanctions
        iii. Legal rights of regulated
        iv. Legal rights of complainants
    c. Specificity of legal standards and penalties
2. Task environment factors
    a. Visibility of violations
        i. Frequency of interaction with regulated entities
        ii. Visibility of violations to complainants
    b. Regulated entities' willingness to comply
        i. Size and/or sophistication of regulated entities
        ii. Cost of compliance/economic resilience
    c. Seriousness of risks to be prevented
3. Political environment factors
    a. Strength and aggressiveness of pro-regulation interests
    b. Preferences of political authorities, as influenced by:
        i. Recent catastrophes or scandals
        ii. Economically urgent projects subject to regulation
        iii. Political controversy over enforcement style
        iv. Electoral shifts/changes in regulatory leadership
        v. Budget cutbacks
        vi. Resistance by regulated government entities
4. Leadership factors
    a. Reactive vs. strong-minded regulatory leaders, If strong:
        i. Leader's policy beliefs
        ii. Beliefs concerning enforcement style
    b. Degree of staff professionalism

Source: Kagan (1994:411).

Appendix B:
Eleven Principles for Improving Enforcement and Compliance

Principle	Definition
Evidence Based Enforcement	Regulatory enforcement and inspections should be evidence-based and measurement-based: deciding what to inspection and how should be grounded on data and evidence, and results should be evaluated regularly.
Selectivity	Promoting compliance and enforcing rules should be left to market forces, private sector and civil society actions wherever possible: inspections and enforcement cannot be everywhere and address everything, and there are many other ways to achieve regulatory objectives.
Risk focus and proportionality	Enforcement needs to be risk-based and proportionate: the frequency of inspections and the resources employed should be proportional to the level of risk and enforcement actions should be aiming at reducing the actual risk posed by infractions.
Responsive regulation	Enforcement should be based on "responsive regulation" principles: inspection and enforcement actions should be modulated depending on the profile and behavior or specific businesses.
Long term vision	Governments should adopt policies and institutional mechanisms on regulatory enforcement and inspections with clear objectives and a long-term road-map.
Coordination and consolidation	Inspection functions should be coordinated and, where needed, consolidated: less duplication and overlaps will ensure better use of public resources, minimize burden on regulated subjects, and maximize effectiveness.
Transparent governance	Governance structures and human resources policies for regulatory enforcement should support transparency, professionalism, and results-oriented management. Execution of regulatory enforcement should be independent from political influence, and compliance promotion efforts should be rewarded.
Information integration	Information and communication technologies should be used to maximize risk-focus, co-ordination and information-sharing – as well as optimal sure of resources.
Clear and fair process	Governments should ensure clarity of rules and process for enforcement and inspections: coherent legislation to organize inspections and enforcement needs to be adopted and published, and clearly articulate rights and obligations and officials and of businesses.
Compliance promotion	Transparency and compliance should be promoted through the use of appropriate instruments such as guidance, toolkits, and checklists.
Professionalism	Inspectors should be trained and managed to ensure professionalism, integrity, consistency and transparency: this requires substantial training focusing not only on technical but also on generic inspection skills, and official guidelines for inspectors to help ensure consistency and fairness.

Source: OECD (2014:14).

Choices in Regulatory Program Design and Enforcement

Christopher Carrigan
Elise Harrington

June 2015

Acknowledgments

We thank Cary Coglianese for reviewing and providing very instructive comments on earlier drafts of this paper. This paper is released as part of the Penn Program on Regulation's Best-in-Class Regulator Initiative which is supported by the Alberta Energy Regulator. A subsequent version of this paper will appear as a chapter in the forthcoming volume, *What Makes a Regulator Excellent* (Cary Coglianese, ed.), to be published by the Brookings Institution Press. Additional work related to this project is available online at www.bestinclassregulator.org.

About the Authors

Christopher Carrigan is Assistant Professor of Public Policy and Public Administration at the George Washington University Trachtenberg School, a scholar at the GW Regulatory Studies Center and a member of the faculty of GW's Environmental Resource Policy Program. His research focuses on regulatory policymaking, exploring responses to disasters in regulated industries and the influence that organizational design and politics have on regulatory performance. In addition to his publications in a variety of academic journals and edited volumes, including those assessing regulatory capture and investigating regulatory failures, Professor Carrigan is co-editor of the recently published University of Pennsylvania Press book, *Does Regulation Kill Jobs?* At the Trachtenberg School, he teaches courses on microeconomics as well as research methods and applied statistics and was the recipient of the inaugural Trachtenberg School Full-Time Faculty Outstanding Teaching Award in 2014. Professor Carrigan holds a PhD in public policy from Harvard University, an MBA from the University of Chicago and a BA in economics from Davidson College. He joined the Trachtenberg School from the University of Pennsylvania where he was the Regulation Fellow at the Penn Program on Regulation. Previously, Professor Carrigan worked as a director at two economic consulting firms and a manager at a large financial services company.

After graduating from Penn in 2011 with a B.A. in Architecture and Environmental Studies, Elise Harrington took a position as a Research Associate for the School of Design on the D.O.E. Energy Efficient Buildings Hub (now Consortium for Building Energy Innovation). This fall, Elise began her Ph.D. in City and Regional Planning at PennDesign. Additionally, she joins the Kleinman Center on a part-time basis as a Research Assistant.