

Performance-Based Regulation: Concepts and Challenges

Cary Coglianese*

Around the world, one of the persistent and widespread complaints made about government regulation has been that it is too constraining. Often referred to pejoratively with the adjective “command-and-control,” regulation is frequently criticized as costly and unreasonable (e.g., Bardach & Kagan, 1982; Howard, 1994). Such objections are loudest when rules inflexibly require every regulated entity to take the same actions – or adopt the same technological fixes – even if under some circumstances, or for some entities, the required action or technology might be expensive, ineffectual, or even counterproductive. Those objecting to inflexible regulation believe that the better approach is to adopt performance-based regulation. Instead of telling businesses exactly what technologies they must use to address public problems, government should establish binding performance or outcome goals, leaving it up to regulated firms to figure out how to achieve the specified outcomes or level of performance.

Such a performance-based approach to regulation has many adherents around the world. In the United States, both Republican and Democratic administrations have expressed a preference for the use of performance standards. For example, in 1993, President William Clinton directed federal regulatory agencies to try to “specify performance objectives, rather than specifying the behavior or manner of compliance that regulated entities must adopt.”¹ Ten years later, President George W. Bush’s Office of Information and Regulatory Affairs similarly advised that performance standards are “generally superior to engineering or design standards because performance standards give the regulated parties the flexibility to achieve regulatory objectives in the most cost-effective way” (Office of Information and Regulatory Affairs, 2003). Another eight years later, President Barack Obama issued an executive order that both

* University of Pennsylvania Law School, Department of Political Science, and Penn Program on Regulation. I am grateful for helpful comments and suggestions from Robert Baldwin, Francesca Bignami, John Braithwaite, Daniel Keleman, David Levi-Faur, and David Zaring, as well as participants in sessions at the ABA Administrative Law and Regulatory Practice Section, the George Washington University Law School, and the Law and Society Association’s annual meeting where versions of this paper were presented. Matthew McCabe, Ben Meltzer, and Andrew Schlossberg provided able research assistance. I am especially grateful to Jennifer Nash and Todd Olmstead for earlier research collaborations on performance-based regulation. A number of the ideas and examples reflected in this paper initially emerged from joint research we conducted that was funded by the U.S. Department of Transportation and through a fellowship from Resources for the Future funded by the Mellon Foundation.

¹ Executive Order 12,866 (1993).

specifically affirmed the Clinton preference for “performance objectives” and added an emphasis on “flexible approaches” to regulation.²

Enthusiasm for performance-based regulation extends far outside the United States. The global trading regime operating under the World Trade Organization formally favors the use of performance standards, with Article 2.8 of the Agreement on Technical Barriers to Trade (TBT) stating that whenever appropriate countries “shall specify technical regulations based on product requirements in terms of performance rather than design or descriptive characteristics.”³ Representatives to the WTO’s Committee on Technical Barriers to Trade (TBT) “have stressed...the advantages of performance-based regulation” (WTO TBT Committee, 2011) and urged that “[w]henver possible, . . . standards should be performance based rather than based on design or descriptive characteristics” (WTO TBT Committee, 2012). Likewise, Canada, Mexico, and the United States have committed, as one of ten “common regulatory principles” in a jointly adopted Regulatory Cooperation Framework, to “[p]romote performance-based regulation” as much as generally possible.⁴ The Organization of Economic Cooperation and Development (OECD) notes that “[t]he use of performance-based regulation is rapidly developing in OECD countries” (OECD, 2002: p. 135). Performance standards’ flexibility not only offers domestic regulators an option for achieving their goals at lower cost to industry, but in the context of international trade these standards promise a simpler means of achieving the kind of regulatory coordination that can better facilitate welfare gains from trade. As a result, across a wide range of policy domains – occupational health and safety, environment and energy, building and fire protection, nuclear reactor safety, and banking and securities, to name a few – regulators and regulatory analysts around the world are increasingly considering or pursuing performance-based regulatory approaches.

Despite the broad enthusiasm performance-based regulation has garnered, remarkably little empirical research exists about how performance standards have actually worked in practice (Coglianese, Nash & Olmstead, 2003).⁵ Indeed, despite the vastness of the policy literature on regulatory instruments, consideration of the relative advantages and disadvantages of different regulatory instruments remains largely theoretical. Comparative empirical analysis, of course, faces considerable challenges. As would be the case with any governmental policy, sorting out the distinctive impacts associated with performance standards, as opposed to other types of regulatory instruments, requires overcoming the difficulties of achieving randomization or applying statistical methods to yield valid causal inferences (Coglianese, 2012; Ayres, Listokin, and Abramowicz, 2011). The empirical researcher must be able to isolate the effect of the instrument type from all the other factors that might explain regulatory outcomes.

² Executive Order 13,563 (2011).

³ Agreement on Technical Barriers to Trade, 1868 U.N.T.S. 120.

⁴ Regulatory Cooperation Framework, Common Regulatory Principles (2007), *available at* <https://www.whitehouse.gov/sites/default/files/omb/oira/irc/SPP-Common-Regulatory-Principles.pdf>. Separate recommendations emanating from Canada have also favorably singled out performance standards. *See, e.g.*, External Advisory Committee on Smart Regulation, *Smart Regulation: A Regulatory Strategy for Canada* (2004), *available at* <http://publications.gc.ca/collections/Collection/CP22-78-2004E.pdf>.

⁵ An exception is an empirical study of building code regulations in New Zealand (May, 2003).

Yet first and foremost, the development of solid empirical analysis of performance standards and other instruments depends on a clear and accepted conceptual taxonomy of regulatory design, something which is sorely lacking. Widespread variation in terminology about regulatory instruments reveals that no system yet exists by which either government officials or researchers can classify regulations by their type—performance-based or otherwise—and thus begin systematically measuring and comparing the outcomes of different instruments (Richards, 2000). Further complicating the matter, rules often come in packages, with performance standards combined together with other types of rules to impose several different types of legal obligations on specific industries or economic practices. For researchers to discover how the choice of instrument type affects the benefits and costs of regulation, we need to separate out performance based rules from other types of rules. To do that, we need, at a minimum, to develop a clearer, more widely-accepted definitional and theoretical vocabulary.⁶

In this chapter, I propose some conceptual order for understanding performance standards and their advantages as a first step toward developing better empirical evidence that can inform decisions about when such standards should—and should not—be used. I introduce a taxonomical scheme for distinguishing between performance-based instruments and other types of regulation. I also show, though, that even within the category of performance-based regulation, performance standards can vary along a number of dimensions that can affect their relative costs and benefits. In addition to offering a conceptual advance in understanding performance standards, I also draw on several examples of performance-based regulation in practice to re-consider the theoretical advantage of performance standards, namely their flexibility. I also focus on some of the potential disadvantages—what some might even consider dangers—of performance-based regulation. My purpose in highlighting the disadvantages of performance standards is not to deny their potential advantages nor to suggest that other instrument types are necessarily or always better. On the contrary, other regulatory instruments can have their problems too; indeed, sometimes they will suffer the same problems as performance standards. My aim here is simply to make the case that performance standards, as with any other instrument type, can work well and work poorly under different circumstances. Rather than continuing to treat performance standards with unbridled enthusiasm, based exclusively on an incomplete set of theoretical considerations, policy makers and researchers should take a careful look at how performance standards themselves actually perform. I do not claim here to offer definitive or categorical conclusions about how performance standards perform under all circumstances, but rather I seek to show that the *performance of* performance standards is an empirical question, one whose answer is long overdue.

In the end, if researchers and regulators do not do more to understand what performance-based regulation is and how it is structured, the disadvantages of performance standards in practice could very well come to outweigh their advantages in theory. Perhaps precisely because of the growing enthusiasm for performance-based regulation around the world, a clear-eyed account is needed, so that regulators will see that they need to approach the use of

⁶ Of course, this is not to deny that other obstacles will remain. For example, analyzing longitudinal variation in regulatory standards may be complicated for reasons of path dependence. If a technology standard is followed by a performance standard in the same jurisdiction, firms that have already invested sunk costs in the old (required) technology may well stick with it even if they would never have adopted that technology if a performance standard had been adopted in the first place.

performance standards with the same level of care that they would any instrument type – and perhaps, at times, even greater care. As with any form of regulation, if performance-based regulation is not designed and deployed well, it can prove costly, ineffectual, or even counterproductive.

What is Performance-Based Regulation?

As with other terminology in public policy, the phrase “performance-based regulation” merits definition, as it often means different things to different people. In an important respect, all regulation is or should be “performance-based”: all regulation is or should be directed at changing the behavior of regulated entities in ways that improve their performance in terms of enhancing social welfare. But such a capacious definition is obviously not very useful. To define performance-based regulation more precisely, we can distinguish four main ways that performance can be incorporated into regulation:

1. Performance as a basis for evaluating regulatory programs and agencies (evaluation or management).
2. Performance as a criterion for allocating enforcement and compliance resources (targeting).
3. Performance as a trigger for the application of differentiated or tiered regulatory standards (tracking).
4. Performance as the basis for legal or regulatory commands (standards)

Although each of these aspects of regulation can in principle be “performance-based,” my focus here is on the last of these, namely the use of performance as a basis for legal or regulatory commands.

A performance standard specifies a required outcome but leaves the means of achieving that outcome to the discretion of the regulated entity. Consider several concrete examples of this general definition from different regulatory realms and jurisdictions:

- The U.S. Consumer Product Safety Commission (CPSC) does not specify how manufacturers must package medicines to make it hard for children to open the packages and poison themselves. Instead, it imposes a performance standard for child-resistant packaging. Packaging manufacturers must test any new designs with a sample of four-year old children who are instructed to open the package. The basic standard is that 85% of the children must be unable to open the package within five minutes.⁷ The CPSC does not tell manufacturers how to construct bottle caps or other child-resistant features of packaging; it just says the packages have to resist opening by the vast majority of children in tests.

⁷ 16 CFR § 1700.20 (2015).

- The U.K. Department of Trade and Industry adopted regulations in 1999 designed to address air pollution from a variety of off-road mobile sources, such as tractors, bulldozers, fork-lifts, and other machines. Rather than specifying exactly how engines in these many different types of equipment needed to be built, the Department imposed a set of performance standards. These performance standards were stated in terms of “limit values” that a prototype engine needed to meet before the manufacturer could receive a certificate of approval, authorizing the sale of the engines in equipment. To demonstrate that an engine could meet the performance standards, a governmental authority would connect the prototype to an emissions analyzer and run the engine in accordance with a detailed testing protocol. As long as levels of carbon monoxide, hydrocarbon, nitrogen oxides, and particulates did not exceed the limit values, the engine would pass the test and receive the certificate of approval.⁸
- The National Highway Traffic Safety Administration (NHTSA) in the United States operates under a statute that requires vehicle safety standards to “be stated in objective terms,” which the agency construes as directing it to issue performance-based standards.⁹ After Congress in 1991 required automobile manufacturers to install air bags in most of their vehicles,¹⁰ NHTSA issued standards for air bags that specified the level of performance air bag systems must meet.¹¹ Measured using now-famous crash-test dummies, air bag systems are required to protect the dummies from specific levels of force related to injury on test crashes into rigid barriers (Federal Motor Vehicle Safety Standards, 1993). NHTSA did not tell auto manufacturers exactly how they needed to design, make, or install their air bags; manufacturers just needed to pass the performance test.
- Buildings in New Zealand must meet the requirements contained in a performance-based building code. As the New Zealand government explains, the code “states how a building must perform in its intended use rather than describing how the building must be designed and constructed.”¹² For example, the buildings in which nursing homes or day care centers are housed “shall have provision for maintaining the internal temperature at no less than 16°C measured at 750 mm above floor level, while the space is adequately ventilated.” To prevent fires, one of the relevant code provisions requires that the “maximum surface temperature of combustible building materials close to fixed appliances using controlled combustion and other fixed

⁸ The Non-Road Mobile Machinery Regulations, 1999, S.I. 1053, art. 2, ¶ 6.

⁹ 15 U.S.C. § 1392a (2006 & Supp. V).

¹⁰ 56 Fed. Reg. 26039 (June 6, 1991).

¹¹ 58 Fed. Reg. 46551 (Sept. 2, 1993)

¹² New Zealand Ministry of Business, Innovation & Employment, Building Code, <http://www.dbh.govt.nz/the-building-code>. For provisions of the code cited in this paragraph, see <http://www.legislation.govt.nz/regulation/public/1992/0150/latest/DLM162576.html>.

equipment when operating at their design level must not exceed 90°C.” Rather than requiring stairways be made of metal, the building code states that they “must be designed and constructed so that they remain stable during and after fire.”

Given the widespread use of performance standards, it would not be difficult to find many other examples from other areas of regulatory policy and other jurisdictions.

As these examples make clear, a performance standard stands in contrast to a “means” standard, that is, one which specifies exactly how the regulated entity must act (presumably, in order to achieve a desired level of performance). Means standards are also sometimes known as prescriptive standards, command-and-control regulation, specification standards, design standards, or technology-based standards.¹³

Means standards suffer from several disadvantages. First, for some regulated entities, the mandated means may not prove as effective as other means. Second, for some regulated entities, the mandated means may prove to be more costly than other equally effective means. Finally, by specifying how to act, means standards can inhibit innovation in finding better or cheaper ways to achieve the same outcomes. By giving firms flexibility to choose their own means to achieve the desired goals, performance standards theoretically allow firms to select the most effective or lowest-cost options. Performance standards may also place fewer obstacles in the way of innovation as well as the kind of harmonization needed to facilitate international trade.

Performance-based regulation, defined as performance standards, bears certain affinities or relationships to several other regulatory terms or concepts, including:

- *Market-based regulation.* Typically, performance standards require all regulated entities to achieve the same level of performance. For example, an emission limit in an air pollution control law will mean that all smokestacks must ensure that they emit no more than the level of chemicals specified in the limit. With market-based regulation, however, entities are not required to achieve uniform levels of performance, but instead they may average, trade, or bank performance units. Emissions trading regimes – sometimes called tradable emissions permits or just “cap and trade” – are examples of market-based air pollution control regulation (Stavins, 1998; Tietenberg, 2010). It is important to understand that market-based regulation is a type of performance standard, but just not one in which each firm must achieve the same level of performance.
- *Self-regulation.* Self-regulation refers to rules that are imposed by a private firm or a trade association on itself or its own members (Coglianese & Mendelson, 2010). The term focuses on the identity of the “regulator”—the industry, in the case of self-

¹³ At the risk of adding to nomenclature confusion, it bears noting that what some people call “technology-based standards” are actually performance standards that are *based on* what is achievable by existing technologies (e.g., Shapiro & McGarity, 1991). Furthermore, performance standards are themselves sometimes lumped under the banner of command-and-control regulation when they require a uniform level of performance for all regulated entities, as opposed to allowing different firms to meet different performance levels as under market-based instruments (e.g., Stavins, 2007:8).

regulation, as opposed to the government. Self-regulation is flexible in the sense that a firm or an industry trade group gets to choose what the regulation will be. But the nature of the rules adopted by a self-regulator may be, just as with rules adopted by any regulator, either means-based or performance-based.

- *Management-based regulation.* A requirement that regulated entities engage in their own planning and internal rulemaking efforts aimed at the achievement of specific public goals is sometimes considered to be “performance-based” but is more properly called a management-based regulation (Coglianese & Lazer, 2003).¹⁴ Under a management-based regulation, a firm can choose what actions it will take to achieve the public goal motivating the regulation. This flexibility afforded to regulated entities bears certain affinities with performance-based regulation.¹⁵ Yet, management-based regulation is in fact a type of means standard, as it requires firms to engage in a variety of specific means related to planning and other management activities. No achievement of a specified level of performance is required under a management-based regulation – and indeed, in many cases regulated entities are not even required to follow their own mandated but internally constructed plans (Taylor 1984; Benneer, 2006). It is often just the planning itself that is required.
- *Information disclosure.* Information disclosure regulation can require firms to report their levels of performance (Kleindorfer & Orts, 1998; Sunstein, 1999). But it is actually a type of means standard, as it requires firms engage in very specific activities related to the collection and reporting of information. By itself, information disclosure regulation does not require firms to achieve any specified level of performance.
- *Equivalency clauses or waiver provisions.* Equivalency clauses or waivers may be used to supplement means standards. Under these provisions, firms may choose alternative means that achieve equivalent levels of performance as the mandated means (Coglianese, Nash & Olmstead, 2003). They might be considered a type of “back-door” performance standard. Equivalency clauses or waivers exhibit a key feature of certain performance standards, discussed below, namely they typically place the burden of proof on the firm to demonstrate that the required performance level has been met. Often they also require advance governmental approval.
- *Codes of practice or guidance.* Somewhat the mirror image of equivalency clauses, codes of practice are non-binding means standards that may be used to supplement performance standards (Coglianese, Nash & Olmstead, 2003). Under such guidance, firms are given the option to choose to implement the means stated in the code in lieu of meeting the performance standard. Codes of practice are thought to be especially helpful to smaller firms.

¹⁴ Such an approach sometimes has been referred to as “mandated self-regulation” (Ayres & Braithwaite, 1992; Hutter, 2001).

¹⁵ As a result of the flexibility both types of regulation afford regulated firms, regulators have sometimes characterized management-based regulation as performance-based regulation. See, e.g., 79 Fed. Reg. 48693 (Aug. 18, 2014).

It is important to see that not every type of flexible regulatory instrument is performance-based (Bennewer & Coglianese, 2012). Management-based regulation and information disclosure give a considerable amount of ultimate discretion to the firm, but they are indeed requiring very specific means be carried out – either planning or the collection and dissemination of information. By contrast, other flexible approaches, such as market-based regulation and equivalency clauses, provide their flexibility in the same way that performance standards do: by imposing obligations to achieve outcomes without mandating the adoption of any particular means of achieving those outcomes.

Differences in Performance Standards

Just as there are differences between performance standards and other related regulatory instruments, so too are there vast differences in the types of rules that fall within the category of performance standards. Performance standards can differ along at least six different dimensions:

1. Specificity (loose vs. tight)

Performance standards can be loose, such as in an airworthiness rule requiring civil aircraft to be in an “airworthy condition.”¹⁶ Or these standards can be tightly specified, such as with motor vehicle brake standards that require motor vehicles to “decelerat[e] to a stop from 20 miles per hour at not less than” 20 feet for “vehicles with a seating capacity of 10 persons or less, including driver, and built on a passenger car chassis.”¹⁷

Note that both of these standards address the performance of some technology – airplanes or automobile brakes – but neither are technology standards that specify the use of a certain technology. Rather, they define the outcomes the technologies are supposed to achieve, either in loose or tight terms.

Sometimes loosely specified regulation is referred to as principles-based regulation, which articulates general guiding tenets. But since it is possible that some “principles” could speak to means as well as ends, I do not consider principles-based regulation as always or necessarily referring to loose performance standards.

2. Proximity between legal command and regulatory goal (close vs. distant)

Performance standards can have a close or distant connection between the legal command and the regulatory goal. For example, the general duty clause in the United Kingdom’s (UK) Health and Safety at Work Act of 1974 imposes a binding obligation on any employer to “ensure, so far as is reasonably practicable, the health, safety and welfare at work of all his employees.”¹⁸ Contrast that provision with a UK Health and Safety Executive regulation setting a “workplace exposure limit” of 2 parts per million (ppm) of formaldehyde exposure over an eight-hour period.¹⁹ The regulatory goal in both of these cases is to protect workers’ health, a goal that the first rule actually embeds in its command. With the second standard, there are one or more

¹⁶ 14 CFR § 91.7 (2015).

¹⁷ 49 C.F.R. § 393.52 (2015).

¹⁸ Health Safety and Work Act, 1974, c. 2 (U.K.).

¹⁹ Control of Substances Hazardous to Health (Amendment) Regulations, 1992, sch.2.

physiological steps between exposure at a level above the limit and any actual health problems a worker might experience. Indeed, it is at least conceivable that some workers exposed to (slightly) impermissible levels of formaldehyde might not suffer any diminution in health whatsoever. But the second standard is still based on a determination that there is a causal connection between some unacceptably high level of exposure and the public concern motivating the regulation, while the first standard is stated expressly in terms of that public concern.

3. How performance is determined (measured vs. predicted)

Methods of determining performance can vary from (a) direct observation of actual outputs or outcomes (on a continuous or periodic basis), to (b) testing under conditions similar to actual conditions, to (c) computer simulations based on models of the relationship between inputs and outputs. Often the performance in a regulatory standard can be measured, such as when an environmental regulator monitors manufacturers' compliance with motor vehicle exhaust standards by connecting emissions analyzers by a hose to a vehicle's exhaust and directly measuring the emissions that come out for a specified period of time.²⁰ Yet at times, performance cannot be measured at all but must be predicted or modeled. As Brannigan and Smidts (1998) note with respect to performance-based fire safety standards for buildings, "direct measurement of the performance of a building is not usually possible" because doing so would require that "full-scale prototypes of the building would have to be built and then burned under various scenarios." In still other circumstances, it may be possible to predict outputs simply by measuring inputs, such as in estimating how much sulfur dioxide will be emitted from a coal-powered plant based on data on the sulfur content of the coal being burned.

4. Basis for the standard (ideal vs. feasible)

Performance standards can vary in terms of the basis or justification of the required level of performance. Sometimes the required level of performance may seek to achieve an "ideal" state, such as when the U.S. Clean Air Act's risk-based national ambient air quality standards are set at levels that will "protect the public health with an adequate margin of safety."²¹ Other times, rather than basing standards on some ideal level of tolerable risk, performance standards may be set based on what is technologically feasible – in other words, at a level that achieves as many benefits as possible until the costs become unacceptable. An example would be the U.S. Clean Air Act's New Source Performance Standards, set at a level "which reflects the degree of emission limitation achievable through the application of the best system of emission reduction."²² A still further way to determine the level of a performance standard would be to conduct a benefit-cost analysis to find an economically optimal level. As should be evident, the degree of stringency of a performance standard will generally follow from the basis for the standard. That stringency matters.

5. Unit of regulation (individual vs. aggregate)

A performance standard can be applied to narrow or broad targets. The target – or the unit of regulation – is the entity or physical thing that must meet the required level of performance. For example, the unit of regulation could be at an individual level: each automobile produced by an

²⁰ 40 CFR § 86 et seq. (2015).

²¹ 42 U.S.C. § 7409(b)(1) (2006 & Supp. V 2012).

²² 42 U.S.C. § 7411(a)(1) (2006 & Supp. IV 2010).

automobile manufacturer must emit no more than a specified level of pollution. Or the unit can be an aggregation: the *average* emissions of all the cars in an automobile manufacturers' fleet must meet the specified level of pollution. The total level of pollution could well turn out to be the same under either approach, but the way performance is specified in the regulation can be either individualized or aggregated in some fashion. The U.S. EPA's "bubble policy" for stationary sources of air pollution is another example of an aggregated performance standard.²³ The EPA policy allows facilities to average emissions from across all of their point sources, so long as the aggregate level of emissions does not exceed a specified level.

6. Burden of Proof (regulator vs. regulated)

Finally, performance standards can differ in terms of how they allocate the burden of proof. Sometimes, as with equivalency clauses or waiver provisions, the burden of proof is placed on the regulated entity to show that it has achieved or can achieve the stated level of performance. For example, under fire safety code exemptions, a contractor must demonstrate to the satisfaction of the government that an alternative design or plan will provide an equivalent level of safety.²⁴ In most contexts, though, it is up to the regulator to prove that the regulated firm has failed to achieve the required level of performance. It is typically the regulator that must demonstrate that a firm emitted pollution over the applicable performance limit or that the bottle caps on a medicine bottle failed to pass the child safety performance test.

* * *

These differences within the general category of performance standards, matter for the same reason that it is important to understand variations among regulatory instruments more generally: different types of performance standards create different incentives and costs for firms. The extent to which a given performance standard will in practice allow for innovation and cost-effectiveness will depend on the specific characteristics of the standards under consideration. If a regulated firm bears the burden of proof, for example, then the more it must spend to demonstrate performance, the less innovation will take place, all other things being equal.

Performance Standards and Innovation

As noted in a recent WTO committee report, it is often thought that "[p]erformance based regulation is more easily adaptable and *encourages innovation*."²⁵ This is a common claim made about performance standards (OECD, 1997). But what exactly does it mean?

Performance standards provide flexibility that can allow for innovation because a regulated entity may be able to meet the mandated outcome in any number of ways. Even if there might only currently exist one way to meet a performance standard, human ingenuity can

²³ 46 Fed. Reg. 50766 (Oct. 14, 1981); *Chevron, U.S.A., Inc. v. Natural Res. Def. Council, Inc.*, 467 U.S. 837 (1984). See generally Tietenberg, 1985.

²⁴ For an example taken from the fire safety standards for federally-funded buildings, see 15 U.S.C. § 2227(a) (3) (2006 & Supp. V 2011) and 41 C.F.R. § 102.80-105 et seq.

²⁵ WTO Committee on Technical Barriers to Trade (2011) (emphasis added).

be applied to identify new ways. Over time, the kind of flexibility built into performance standards may permit regulated entities to find ways to meet the regulation's objectives that are even less costly than any means known at the time the regulation is adopted.

Although the innovation possibilities of performance standards mark one of their chief theoretical advantages, it is important to recognize that performance standards do not by themselves actually *encourage* innovation. That is, there is nothing in the specification of a required end-state that for that reason provides any new incentives for innovation. The incentives for innovation exist outside of the regulation, in the overall competitive business environment that drives regulated firms to look for new, less costly ways to achieve regulatory goals. While performance standards facilitate such innovation, since the regulatory scheme gives firms flexibility, the performance standards alone do not actually drive the innovation.

It is somewhat of a misconception to think that performance standards will induce innovation (as opposed merely to *allow for* innovation). There is also another corollary misconception, namely that means-based standards lock-in existing technologies and do not allow for innovation. This is not necessarily so, as means-based rules typically contain options for obtaining waivers – and of course, there is always a background option of amending a means standard to authorize or require a different kind of means. Ultimately, then, the choice between means standards and performance standards is not one between *no innovation* and *robust innovation*. The reality is more complicated.

That reality is still more complicated if the stringency of a standard is taken into account. Until now, I have focused attention on the *form* of performance standards as compared to the form of other types of regulatory instruments, explaining that the flexibility provided by performance standards does not, by itself, provide an incentive for innovation. But *form* is not the same as *stringency*. Depending on their level of stringency, performance standards can indeed encourage innovation. If a performance standard is stringent enough that it can only be met through new technologies that are not currently available, firms will definitely be encouraged to innovate (or to abandon their line of business) (Ashford, Ayers & Stone, 1985).

A further qualification can arise with types of performance standards that allow for non-uniformity of performance, so-called market-based regulation. This form of performance-based regulation (and only this form) can encourage innovation through its form, entirely separate from stringency. To understand why, recall the example of an emissions permits trading regime, a non-uniform environmental performance-based regulation. A permit trading regime is performance-based in that the regulated firm must never exceed the level of emissions authorized by the number of permits that it holds. But firms are allowed to buy and sell those permits, so that some firms may have more permits than others. Allowing firms to aggregate emissions by buying and selling permits provides an incentive for innovation, because if a firm can improve its performance it can sell its excess permits in the marketplace (Jaffe & Stavins, 1995).

Outside of stringency and marketability, the degree to which performance standards allow for innovation will vary because innovation is risky and costly for regulated entities. These costs can sometimes be affected by the standard itself. The more difficult it is for the regulated entity to demonstrate it has met the required level of performance, the less innovation will take

place, all other things being equal. Key factors affecting the difficulty of demonstrating performance comprise the *burden of proof* and the *standard of proof*. Must the regulated entity satisfy the burden of proof and demonstrate to the regulator that it meets the performance standard, say, through either (a) regular reporting or (b) preapproval of an innovative approach? Or must the regulator bring an enforcement action and bear the burden of proof? Moreover, what standard of proof (e.g., kind of evidence and level of certainty) is required? How costly is it to operationalize and measure performance? How much uncertainty will the regulator tolerate (e.g., preponderance of evidence versus beyond a reasonable doubt)? The answers to questions like these will likely affect, in practice, the resulting level of innovation permitted by performance standards.

To illustrate, compare the CPSC's standards for child-resistant packaging, described earlier,²⁶ with the Clark County (Nevada) building commission's performance-based equivalency clause (Evans, Weber & Quiter, 1998). In the former, if a manufacturer fails to meet the packaging standard, the CPSC must bring an enforcement action and meet its burden of showing that the product package is unsafe. By contrast, in the case of the Clark County building commission, the burden is on the firm to demonstrate it can achieve a level of performance equivalent to the applicable means standards. Clark County is where Las Vegas is located. When Circus Circus Enterprises proposed a new design for its 30-story Luxor Hotel and Casino in Las Vegas, the building's unique pyramidal shape and large internal atrium did not comply with the county's existing "prescriptive" code provisions addressing the ventilation of smoke. Circus Circus had to proceed under a performance-based equivalency clause, demonstrating to the county building regulators that the hotel's alternative ventilation systems would provide a level of performance in the event of fire equivalent to a traditional building following the normal, means-based code. The burden of proof was on Circus Circus, and meeting it required the use of extensive computer simulation, third-party testing, and ultimate simulations with theatrical smoke – all with multiple reviews undertaken by county officials.

We can expect that, all things being equal, the closer performance standards are to the Clark County equivalency clause process, the less innovation will take place. I also propose a corollary: the more responsive the regulator is to making exceptions or changes to means-based regulation, the more innovation can be expected to take place, all other things being equal. The key takeaway is that with performance standards, as with means standards, there will be tradeoffs between the level of certainty demanded by the regulator and the degree to which innovation will likely occur. By varying the burden and standard of proof, regulators can design performance standards in ways that increase the benefits from certainty in the achievement of the desired outcomes – but doing so can also increase costs to regulated firms and actually inhibit innovation. Not all performance standards will allow for, let alone encourage, much innovation.²⁷

²⁶ 16 C.F.R. § 1700.20 (2015).

²⁷ A similar analysis could be made of performance standards' oft-claimed advantage in regulatory harmonization and the facilitation of international trade. Means standards can certainly be harmonized to facilitate trade better, if they are made to be identical. With sufficient interest, trading partners can surely reach agreements on compatible or identical means standards. On the other hand, as the discussion in this and the previous section of this chapter should make clear, depending on how countries' performance standards are designed, they may afford little meaningful harmonization potential, such as if each requires

When Should Government Choose Performance Standards?

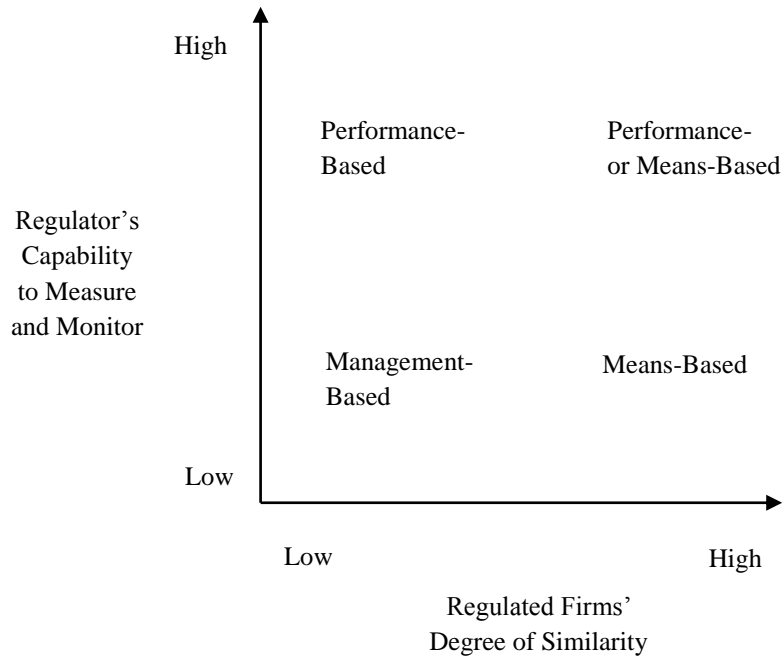
The discussion so far should make clear that performance standards allow for flexibility along two dimensions. The first is *cross-sectional* flexibility, which enables what might be called “fitness,” because performance standards allow for differently situated firms to meet the required outcomes in the ways that best fit their own circumstances. With fitness, the question is: To what degree does one size truly fit all? The second dimension is *longitudinal* flexibility, which I have been just discussing; flexibility over time allows for “innovation.”

The more heterogeneous a regulated sector, either cross-sectionally or longitudinally, the greater the cost savings that can accrue from allowing firms to choose their own means rather than requiring every firm to adopt identical means. A regulated sector will be cross-sectionally heterogeneous when firms come in different sizes, use different processes or techniques, draw upon different supply chains, face different prices for labor or capital, or exhibit any number of other differences that affect the costs of meeting regulatory objectives. For example, the industry covered by the UK’s non-road mobile machine emissions standards is quite heterogeneous, as the firms making farm tractors are quite different from those that make fork-lifts and construction trucks – and, of course, the products they make are different as well. By contrast, a sector will be heterogeneous longitudinally, if there is a high likelihood that the sector – or its products and processes – will change over time. The financial sector – as with nearly anything today having to do with digital technology – experiences considerable, even rapid changes over time.²⁸

Although performance standards can result in cost-savings under conditions of heterogeneity, these standards depend on performance being reliably measured in some way, whether through direct measurement, testing, simulation, or modeling. Thus, as I have elsewhere illustrated, the choice between performance standards and means standards requires a careful assessment of both the regulated firms’ heterogeneity and the regulator’s ability to measure the relevant outputs or outcomes (Coglianese, 2008). The resulting Figure 1, below, shows where the three primary approaches to regulation would be most suited. If performance standards are most appropriate under conditions of heterogeneity and when the regulator can assess output well, Figure 1 shows that means standards are appropriate when output assessment is difficult (or costly) and when most of the firms will be the same, both cross-sectionally and over time. When regulated firms are similar and stay that way, one size truly does fit all. Finally, as David Lazer and I have shown elsewhere, management-based regulation fits well in circumstances fitting into the lower-left quadrant of Figure 1, where heterogeneity is high but the regulator’s capacity to assess outputs is limited (Coglianese & Lazer, 2003).

different but expensive testing protocols or measurement techniques. Performance standards provide no logically inherent advantage in terms of harmonization merely because they take a performance-based form.
²⁸ Parts of the economy that touch on both the financial sector and information technology will presumably be even more prone to frequent change over time. At least that is the argument eBay has recently made in advancing a case for a performance-based approach to regulation of its PayPal business (PayPal and eBay Inc., 2013).

Figure 1: Conditions for Performance-Based and Other Forms of Regulation



While the framework in Figure 1 provides a very rough heuristic for thinking about the choice between performance-based and means-based regulation, it is helpful to consider further the conditions under which these two approaches to regulation may be appropriate. Especially given the almost universal enthusiasm for performance-based regulation – with, as already noted, strictures adopted by the last several presidential administrations in the United States seeking effectively to privilege performance standards as a default choice for government regulators – decision makers might do well to ask whether performance-based regulation is indeed usually better. Obviously the answer will depend on the circumstances under which performance standards would be applied, but it is possible to say more by considering how performance and means standards might compare in terms of the demands they place on the regulator.

Since the principal advantage of performance-based regulation lies in the cost-savings afforded by its greater flexibility, let us imagine a decision-making scenario in which benefits are held constant. In other words, imagine that a regulator faces a choice between adopting a means standard and a performance standard, each with identical social benefits because both will achieve equivalent levels of risk reduction or other social improvements. Although in reality such an assumption will obviously not always hold, for purposes of analysis it is a useful one to make. It allows us to consider the best case for performance-based regulation, namely one where it is more cost-effective, achieving identical societal benefits for lower overall cost to society (especially in the form of compliance costs to firms). Counter-intuitively, even in such circumstances where both types of standards might yield identical societal benefits, it still may sometimes be better for regulators to adopt means standards.

If benefits to society are held constant, the choice between a means standard and a performance standard in any particular regulatory context will turn on their relative costs. Here it is important to keep in mind two kinds of costs. The first, which I will call “social costs,” are the kind of costs that would typically enter into a conventional benefit-cost analysis: opportunity costs to firms (compliance costs) and any costs of negative side-effects associated with the regulation. The second kind of cost, which typically does not enter directly into the equation of a benefit-cost analysis, is the kind that the government must incur in developing and enforcing the rule. These “governmental costs” or “administrative costs” are typically much smaller than the social costs and social benefits of regulation and therefore it is typically of little consequence if they are not explicitly taken into consideration in a regulatory impact analysis. Moreover, since governmental costs are paid for by taxes, they can in theory be recouped through transfer payments; in other words, they do not represent a real opportunity cost, at least to society overall (Office of Management & Budget, 2003: 9).

But from the very real standpoint of day-to-day operations, governmental costs do matter. A regulator is constrained. It simply cannot raise taxes to pay for the costs associated with developing and enforcing a new regulation. As a result, two key practical questions arise for a regulator choosing between performance and means standards which would each deliver equal social benefits: (1) How much social cost-saving will a performance standard likely yield? (2) Are the savings in social costs sufficient to justify any additional investment of governmental resources to create and monitor performance standards? Again, because the social benefits are assumed to be the same, the only tradeoff is between cost-savings to the regulated community, which is what reduces the social costs of performance standards, and the relative differences in administrative costs to the government.

Governmental costs – the transaction costs associated with instrument choice – can turn out to be crucial. After all, in a world with no governmental costs, there would be no difference between means standards and performance standards (Coglianese & Lazer, 2003: 703). The government regulator could create highly specified means standards, so that cross-sectional fitness could be perfectly addressed. Similarly, the regulator could adapt regulations effortlessly in response to innovation and other changing circumstances over time. In a world with no governmental costs, means standards could be seamlessly tailored to fit any different or changing circumstance.

Obviously, we do not live in a world with zero governmental costs. The challenge for regulators is to use their limited resources in a way that maximizes social welfare. Consider a highly stylized example of a choice between two hypothetical standards, one means-based, one performance-based. As illustrated in Table 1, both standards would result in the same level of social benefits, but the performance standard would yield these benefits at half the costs, because firms could choose or innovate to find their own, lowest-cost means to achieve the desired output. (The numbers in the table are in common units, such as money.) As a result, the “net social benefits” (i.e., social benefits minus social costs) of the performance standard would be greater. This is the main reason why policy makers and scholars enthusiastically profess to prefer performance standards.

Yet consider what happens when we add in governmental costs. If the performance standard were to prove more costly to government – say because such a standard would impose greater costs for monitoring compliance in order to achieve the same social benefits – these costs would need to be added to the calculus. Now, if the administrative costs to government were merely treated as an additional kind of social costs, they could be subtracted from the net social benefits, with a resulting, final “overall net benefits” (see Table 1). If the governmental or administrative costs of regulating are dwarfed by the social costs, as they often are in reality by several orders of magnitude, they will rarely affect the bottom line in any comparison between means standards and performance standards – just as they do not affect the hypothetical bottom line in Table 1. Performance standards still outperform means standards in this illustration.

Yet are performance standards really the preferred regulatory approach in this example? Not necessarily. If the regulator faces a constrained budget and needs to make a choice that yields the most “bang for the buck,” it should choose the means standard under the conditions presented in Table 1. Why? For an investment of 25, the regulator yields net social benefits of 100 with the means standard, a return on investment of 4 to 1. The return on investment for the performance standard is only 3 to 1. To put the point somewhat differently, in the example in Table 1 the regulator could implement two means standards for the price of one performance standard, and doing so would result in 33 percent more net social benefits.

Table 1: A Stylized Choice Between Means Standard and Performance Standard

	Hypothetical Regulation A (Means Standard)	Hypothetical Regulation B (Performance Standard)
Social Benefits	200	200
Social Costs	100	50
Net Social Benefits	100	150
Governmental Costs	25	50
Overall Net Benefits	75	100

Of course, this stylized example should not be taken to suggest that means standards will always be preferable; only that they may sometimes be. The key in many cases will be to consider what goes into computing the governmental costs. In other words, what does it take for government to be able to ensure that a performance standard and a means standard can yield comparable social benefits? In terms of governmental costs, there might be two kinds: the cost of crafting a suitable regulatory instrument (rulemaking), and the cost of monitoring and enforcing the rule (implementation). Performance standards may well spare the government the cost of finding the one choice that fits all best, but they might well be more costly to monitor and enforce in some circumstances.

Although each policy decision will require its own analysis, the relevant questions confronting the decision maker will be:

- Is it more costly for government to identify one effective solution and ensure it is being used? (Means standards)
- Or is it more costly for government to establish a way of monitoring the outcomes of many different solutions? (Performance standards)

If there is a known means, one that is already being used effectively by many firms, then the costs of finding a means-based solution should be low. The costs of finding such a solution may also be lower for simple problems than for complex problems. But then again, it may well be easiest of all, whether the problem is simple or complex, to embed the regulatory goal into a performance standard.

But that does not mean it will be easy to ensure that regulated firms actually meet the regulatory goal. When it comes to enforcement costs, performance standards that require “in-use” or continuous monitoring are likely to be more costly than standards that can be enforced by testing samples or perhaps just by making predictions. In some cases, the costs of monitoring the required performance will be prohibitively high, such as with regulation aimed at preventing low-probability catastrophic loss. If a performance standard were simply to say, “Do not construct a building in a way that it will burn down quickly,” the regulator (and the building’s occupants) may not know if the standard has been met until the building catches fire – which will obviously be too late (Coglianese, Nash & Olmstead, 2003: 715).

In comparing governmental costs to estimated cost-savings from performance standards, it should be remembered that the actual cost-savings to regulated industry will likely vary depending on the way a performance standard is designed. Although performance standards do promise cost-savings, as I explained earlier, different types of performance standards are likely to yield different levels of savings. Assuming equivalent social benefits can be achieved, the savings in social costs from performance standards will generally be higher:

- The closer the performance standard is to the ultimate objective (for this allows the firm more flexibility than if performance is applied to only one causal branch leading to the ultimate objective).
- The easier it is for the firm to demonstrate performance or to be assured it will be found in compliance (for this makes it less costly for the firm to try some new means of meeting the required performance).
- The greater the background incentives that firms have to innovate (for as I have noted, uniform performance standards themselves do not create new incentives to innovate simply because they are performance-based).

In situations where firms have little external incentive to innovate, where it is difficult or costly for firms to be assured that trying new approaches will still meet required performance levels, or where the performance standard is set so far back on a causal chain that it does not afford much real flexibility, the cost differences between performance standards and means standards may be minimal or even nonexistent.

One final comment bears mention when it comes to the relative costs of performance standards versus means standards. It is sometimes said that a hybrid regulatory approach would be best, specifically one that combines performance standards with means-based codes of practice or guidance. Such a hybrid approach gives more sophisticated firms the ability to innovate and adapt, while at the same time it gives clear guidance to those firms that simply want to know “what do I do.” Although this may seem like an ideal solution, it is only likely to increase governmental costs. With hybrids, government faces higher costs of analysis and decision making over any alternative -- that is, against either a pure means standard or a pure performance standard -- since government must both find a recommended means and establish an appropriate performance goal. Since the monitoring costs for pure performance standards are normally thought to be greater than for pure means standards, the hybrid approach will result in monitoring costs for government that are higher than a pure means standard but lower than a pure performance standard. Furthermore, the social -- that is, nongovernmental -- costs of the hybrid will equal a pure performance standard only if the firms that follow the code of practice are those for which the code represents their lowest-cost strategy. But if some firms follow the code even though it contains, in relation to their particular circumstances, a higher-cost means (for reasons of inertia or convenience), then social costs could be higher than a pure performance standard but lower than a pure means standard (assuming that at least some firms do depart from the code when a lower-cost means is known) (Coglianese, Nash & Olmstead, 2003: 713).

The Limits of Performance-Based Regulation

Despite the theoretical attractiveness of performance standards, and their endorsement in governmental proclamations around the world, the ultimate value of any performance-based regulation will depend on what it ultimately delivers. While systematic empirical research on performance standards remains to be undertaken, experience with at least some performance standards reveals several unacknowledged or little emphasized limitations of performance-based regulation. It is, of course, possible that performance standards are still “generally superior” to technology or design standards, as stated in the U.S. government’s regulatory guidance (Office of Information and Regulatory Affairs, 2003). But the seemingly unbridled enthusiasm for performance-based regulation by regulatory commentators and officials around the world masks some of its challenges and limitations. In this final section of the paper, I offer a brief corrective account of performance standards, pointing to some of the ways that they can fail. Of course, recognizing more completely performance standards’ limitations is not to counsel against using them; however, the recitation of limitations in this section of the should make future policymakers more attentive in the future, rather than acting on the basis of generalizations about performance-based regulation that are either too sweeping or too optimistic. Ultimately, the performance of performance-based regulation is an empirical question. As a result, when regulatory officials find themselves considering performance standards, which indeed can be superior in some respects, they also should recognize that performance standards cannot do everything -- and they may even at times pose some of their own unique perils.

Performance-based regulation does not eliminate conflict or litigation over regulation.

Although it has sometimes been suggested that a shift to performance-based regulation would result in less adversarial and more collaborative regulation, the U.S. EPA's experience regulating heavy-duty diesel emissions does not support such optimism. The EPA issued standards that are similar in many respects to the standards the UK adopted for mobile machinery. The EPA's standard limited carbon monoxide emissions from heavy-duty diesel truck engines to 40 g/BHP-hr and imposed a combined standard for hydrocarbons and NO_x of 16 g/BHP-hr.²⁹ These emissions standards were expressed as grams of emissions per brake horsepower-hour – g/BHP-hr. – to focus on the relationship between the engine's emissions and the amount of work it does, as measured by a dynamometer or friction brake applied to the drive shaft. The regulation required testing emissions following a protocol known as the "steady state" test – basically running a prototype engine through 13 separate "modes" (speeds and torques) specified by the EPA and then averaging the resulting emissions captured by a machine analyzer.³⁰ This is a classic performance standard: EPA did not specify how engine manufacturers should design and build their engines to meet these tests; it just required that their engines meet the test.

At various twists and turns over the years, environmental groups, manufacturers, and government have all resorted to litigation over these EPA performance standards. Environmental groups went to court in 1984 to force EPA to issue separate NO_x emissions standards.³¹ Industry filed suit to compel the EPA to delay the implementation of the new standards in 1989. And in 1998, EPA filed an enforcement action against the engine manufacturers, securing the largest settlement (as of that date) in EPA's history (Osenga, 1998). This case shows that there is nothing inherent in performance-based regulation that can reconcile the different interests at stake over regulation nor make organizations avoid pursuing their interests through litigation.

In situations with tradeoffs, embedding just one goal in a performance standard can lead to unintended or undesired consequences for other goals.

Tradeoffs are a staple of policymaking, but the implications of such tradeoffs have previously remained unacknowledged in the literature on performance-based regulation. This is unfortunate because, if tradeoffs are not clearly identified when setting performance standards, the regulator risks adopting a standard that achieves one goal at the expense of other goals. The history of the CPSC's child resistant packaging regulation and NHTSA's air bag regulations are both compelling, albeit tragic, examples of how a myopic use of performance standards to achieve one objective can result in significant losses in terms of other important objectives or values.

The CPSC's performance standard helped ensure that drug packaging was hard for children to open; after all, it called for testing to ensure that 85% of children could not open packaging in five minutes. Yet packaging that met this standard also proved difficult for many adults to open. Subsequent studies showed that about 70% of female and 40% of male senior citizens could not open child-resistant containers – and even 22-64% of adults aged 18-45

²⁹ 37 Fed. Reg. 18,262 (Sept. 8, 1972).

³⁰ Fed. Reg. 18,262 (Sept. 8, 1972).

³¹ *Natural Res. Def. Council, Inc. v. Ruckelshaus*, 21 Env't. Rep. Cas. (BNA) 1953 (D.D.C. Sept. 14, 1984).

experienced difficulties (Sherman et al., 1979; Atkin, et al., 1994). As a result, frustrated adults took steps to subvert child-resistant designs, such as leaving medicine packages open or transferring pills into containers that were easier to open. A 1989 report found that more than 50% of accidental ingestions by young children occurred when adults left pills lying loose or transferred them into other, non-resistant packaging (Jacobsen et al., 1989). The CPSC eventually revised its performance standard to include a test both for child-resistance *and* adult ease of opening – but for more than a decade the performance test essentially hid the tradeoff between child resistance and adult access.

A similar “tunnel vision” afflicted NHTSA over its air bag performance standards. NHTSA’s prescribed performance test used crash test dummies proportioned to the size of the average-sized adult male. Air bags that met this performance test have, to NHTSA’s credit, saved many lives (Glassbrenner, 2003). But tragically, they also led to the development of air bag systems that sometimes harmed, even killed, smaller statured adults and children, even in accidents where these smaller occupants otherwise would have survived (National Center for Statistics & Analysis, 2007). The rapid deployment of airbags creates a force against smaller individuals that leads to head and face injuries – and sometimes other quite gruesome outcomes, including decapitations. With time, NHTSA and the auto industry have worked to see the development of “smart” airbag technologies that now can meet a more multi-faceted performance test. But for many years, the performance standard approach proved harmful for some drivers and passengers due to inadequate attention in initial performance standards to the tradeoffs between protection for average-sized males and the harms to smaller individuals.

Performance-based regulation may not decrease (and may even increase) the complexity and cumbersomeness of regulatory red tape.

Regulation is routinely criticized for imposing burdensome, dense requirements that consume page after page of code books. Are performance standards an answer? The U.S. EPA’s history of developing standards for heavy-duty diesel engine emissions shows that performance-based regulation does not necessarily result in the elimination of governmental red tape. In order to give manufacturers discretion to choose how to reduce NOx emissions, the EPA added layer upon layer of requirements which, over the years, have added up to an extraordinary level of prescriptiveness in how emissions testing must be done.³² Engine manufacturers now employ teams of engineers to understand and comply with these requirements. Moreover, even when industry experts think they have met the government’s tests, the EPA still can threaten to withhold a certificate of conformity and pursue enforcement actions against manufacturers – far from fostering simplicity and certainty. Not only are the costs and complexity of these emissions performance standards a burden for industry, they have not proven easy or cheap for government to design. It has taken the agency decades to refine its testing protocols.

Performance standards can still remove industry’s discretion and flexibility.

As noted, the main theoretical advantage of performance standards comes from the flexibility they provide to regulated entities. But as I have already explained, the level of discretion that a performance standard provides can vary dramatically. For example, when considering flexibility in reference to the underlying regulatory goal, the closer the binding

³² 66 Fed. Reg. 5002 (Jan. 18, 2001).

regulatory standard is to the regulation's ultimate goal, the greater the discretion the regulated entity will generally enjoy. If the ultimate goal of an environmental regulation is to reduce the health effects from human exposure to ground level ozone, then regulated firms will have more discretion the more closely tied their obligation is to this goal. After all, the health effects of ozone could be reduced by, among other things, alerting citizens to stay indoors on days with high smog levels, reducing pollution from stationary sources, reducing evaporative emissions of volatile organic compounds (VOCs) from automobiles, or reducing tailpipe hydrocarbon or NOx emissions from mobile sources.³³ But even if all four of these options were equally effective in protecting public health, only one of these four options will satisfy the regulator's command if the agency tells vehicle manufacturers to control tailpipe emissions – even if they do so using performance standards.

Of course, a performance-based tailpipe standard still allows engine and vehicle manufacturers to consider a range of options that they might not otherwise be allowed to consider if the government selected the precise means to achieve the desired reductions in tailpipe emissions. Options that remain might include the design of engine operating conditions (such as timing, fuel injection, and so forth), changes to fuel, and installation of after-treatment devices. Yet in reality, even all of these options for reducing tailpipe emissions might not be available to the manufacturer. The history of the U.S. EPA's heavy-duty diesel emissions regulation indicates that a performance-based standard is little different than a design or technology standard if only one known or available means will meet the mandated performance level. For example, once the EPA reduced its permissible 2007 NOx limits by nearly 90%,³⁴ engine manufacturers for all practical purposes were forced to adopt after-treatment – that was the only available means that would meet the new performance standard (Dollmeyer et al., 2007).

Testing and monitoring play a crucial role in the successful implementation of performance-based regulation.

Over the decades, the implementation of U.S. EPA's heavy-duty diesel emissions standards has been accompanied by the development of a variety of testing protocols: from “steady state” to multiple versions of “transient tests” to experimentation with limited in-use testing (Lloyd & Cackette, 2001). At each step along the way, as EPA tried different ways to monitor engines' performance, the agency learned that the previous version of its testing did not adequately reflect what the agency sought to measure. It may seem almost a truism to note that performance-based standards depend on the ability of government agencies to specify, measure, and monitor performance. But it is not often enough acknowledged how difficult, if not nearly impossible, it sometimes can be to obtain reliable and appropriate information about performance.

The transatlantic scandal over emissions from Volkswagen diesel automobiles has revealed even more starkly the vital role of effective monitoring in a performance-based regulatory scheme. In September, 2015, EPA accused Volkswagen of installing software in

³³ EPA, *Ozone Reduction Strategies — Tips to Reduce Ozone — Control Measures*, <http://www.epa.gov/airquality/ozonestrategy/controlmeasures.html> (last visited Apr. 2, 2015).

³⁴ 40 C.F.R. § 86.007-11 (2015).

what amounted to an estimated 11 million cars with 2-liter diesel engines that detected when the cars were being tested for nitrogen oxide emissions.³⁵ According to EPA, the company's cars met emissions standards in the laboratory or testing station, but during normal operation they emitted nitrogen oxides up to 40 times permissible levels (Isidore & Valdes-Depena, 2015). Volkswagen, which subsequently admitted the "misconduct,"³⁶ apparently evaded regulators in both the US and in Germany *for seven years*. Although the firm itself bears responsibility for circumventing the regulators' testing, the fact that the firm succeeded in evading attention for so long testifies to the inadequacy of the government's oversight (Auerbach 2015). Tellingly, Volkswagen only succumbed to detection when a small nonprofit group hired a few researchers at West Virginia University to test just three Volkswagen vehicles (Plungis & Hull 2015). It should not have been very difficult for regulators to have detected the violations themselves much earlier.

In November, 2015, EPA issued a second notice of violation, alleging that Volkswagen also installed a similar "defeat device" on 3-liter engines used in luxury sedans and sport-utility vehicles produced by the company's Volkswagen, Audi and Porsche brands (Ziegler 2015).³⁷ Volkswagen is now facing fines in the United States of up to \$18 billion in civil penalties under federal statutes (Assis 2015). Remarkably, there remains some question whether Volkswagen will ever face any similar penalties in Europe because EU regulations apparently state that "the settings of the engine and of the vehicle's controls shall be those prescribed by the *manufacturer*" (Hakim and Barthelemy 2015). Regardless of what ultimately happens in Europe, the Volkswagen scandal drives home the importance of testing and monitoring for any effective performance-based regulatory regime. Perhaps if regulators do not hear only unqualifiedly enthusiastic messages in support of performance-based regulation they might be more vigilant in the future about how they monitor compliance once regulated firms start to take advantage of the flexibility that performance standards offer (Coglianese 2015a).

Performance-based regulation may lead to a "teaching to the test" problem.

Previously unrecognized is how performance-based regulation raises a concern similar to a well-known problem in the field of education called "teaching to the test." In education, this problem arises when students learn strategies to score high on tests but nevertheless fail to

³⁵ Press Release, EPA, California Notify Volkswagen of Clean Air Act Violations (Sept. 18, 2015), *available at* <http://yosemite.epa.gov/opa/admpress.nsf/d0cf6618525a9efb85257359003fb69d/dfc8e33b5ab162b985257ec40057813b!opendocument>.

³⁶ *See, e.g.*, Volkswagen, Statement by Statement by Prof. Dr. Winterkorn (Sept. 23, 2015), *available at* https://www.volkswagen-media-services.com/en/detailpage/-/detail/Statement-by-Prof-Dr-Winterkorn/view/2721302/7a5bbec13158edd433c6630f5ac445da?p_p_auth=6jhXNw3R

³⁷ Press Release, EPA, California Notify Volkswagen of Additional Clean Air Act Violations, EPA (Nov. 2, 2015), *available at* <http://yosemite.epa.gov/opa/admpress.nsf/bd4379a92ceceecac8525735900400c27/4a45a5661216e66c85257ef10061867b!OpenDocument>.

develop the core, underlying skills that the tests are intended to measure (Popham, 2001; Crocker, 2003). As with students, regulated firms may find ways to satisfy performance standards but still use the flexibility afforded them under performance-based regulation to depart from regulators' goals.

In 1980, the U.S. EPA established a transient test and required diesel engine manufacturers to pass it.³⁸ The manufacturers built engines that met EPA's standards based on the testing protocol that the agency had established. Yet the government later claimed in a lawsuit that the manufacturers "complied" in a way incompatible with the overarching regulatory goal of reducing NOx emissions (Yandle et al., 2009). The manufacturers allegedly designed engines that would optimize fuel mix to meet the performance test for the duration of the testing protocol, but when the trucks were in use on the roads the onboard computers would automatically adjust that mix afterwards to optimize engine performance for truck drivers – diminishing emissions control.

Performance-based regulation gives industry an opportunity to achieve its own goals within the constraints set forth by the regulation. With performance-based regulation, the command "Do not exceed the standard" may well translate to "You can proceed to market your products or operate your facilities however you see fit, just as long as you do not exceed the standard." Firms seek to meet the standard and still optimize their private interests, for instance by minimizing costs. There is nothing intrinsically wrong with this logic but it does create the possibility that firms will find ways to meet the regulator's test that will satisfy the firms' goals and not the goals of the regulator or of society (May, 2003).

Even when firms fully and faithfully meet performance standards, these standards can still fail.

Even if performance standards are set based on some precursor to, or proxy for, the ultimate outcome of concern to society and even if firms fully comply with both their letter and their spirit, they may still fail. This can occur if there is a faulty or incomplete causal linkage between the performance specified in the regulation and the desired state of the world. In other words, the problem the regulation seeks to address may still persist if other factors, unaffected by the regulation, contribute to the problem or if the regulation fails to address the root causes of the problem (May, 2003). Of course, this also can be true for means standards – as can some of the other limitations reviewed in this section – but it serves as a reminder that performance standards are no panacea.

Conclusion

In calling attention to these limits of performance-based regulation, I do not mean to suggest that regulators should never use this type of regulation. On the contrary, it does hold advantages and will definitely sometimes, perhaps even often, be the best regulatory approach available. But not always. Performance standards depend on the ability of government agencies to specify, measure, and monitor performance, and yet reliable and appropriate information about performance may sometimes be difficult if not impossible to obtain. When implemented in the wrong way, or under the wrong conditions, performance-based regulation

³⁸ 45 Fed. Reg. 14496 (Mar. 5, 1980).

can perform badly, as can any regulatory instrument that is ineffectually deployed. Just as any regulatory instrument can be poorly crafted and ineptly implemented, so too can performance standards. The key question for decision makers is: Are the problems that afflict performance standards worse than the problems that can affect means standards?

If governments around the world are to rely more heavily on performance standards, regulators designing and implementing them should do so fully aware of their limitations as well as their advantages. Successful regulation requires that regulators understand their options when choosing among different types of regulation and that they make their choices fully cognizant of various features even within the same instrument type (Coglianese 2015b). Rather than relying on general claims about the superiority of some regulatory types over others, regulators need to analyze carefully the likely consequences of each relevant design choice under the conditions anticipated (Hutton 2015). They also need to review and reevaluate regulations after they are adopted to ensure that they are working as intended (Coglianese 2012). This is clearly true of performance-based regulation, which unfortunately has been implicated at times in regulatory tragedies that might have been avoided had regulators not focused so single-mindedly on the policy goal motivating the regulation. In the end, it should be clear that effective use of performance-based regulation calls for analytic sophistication: regulators must identify the major contributors to a regulatory problem, understand how these factors cause the problem and relate to each other, and determine what measures and monitoring techniques will effectively ensure compliance. At bottom, good performance-based regulation is no less, and perhaps even at times more, demanding of regulators, than any other form of regulation.

References

- Ashford, Nicholas A., Christine Ayers, and Robert F. Stone. 1985. "Using Regulation to Change the Market for Innovation," 9 *Harvard Environmental Law Review* 419-66.
- Assis, Claudia. 2015. "Volkswagen at Risk of \$18 Billion EPA Fine," *CBS MarketWatch* (Sept. 18). Available at <http://www.marketwatch.com/story/volkswagen-at-risk-of-18-billion-epa-fine-2015-09-18>.
- Atkin, Philip A., Terence P. Finnegan, Susan J. Ogle, and Gillian M. Shenfield. 1994. "Functional Ability of Patients to Manage Medication Packaging: A Survey of Geriatric Patients," 23 *Age Aging* 113-16.
- Auerbach, David. 2015. "Volkswagen's Villains," *Slate.com* (Oct. 1). Available at http://www.slate.com/articles/technology/bitwise/2015/10/volkswagen_s_emissions_scandal_has_a_villain_and_it_s_the_not_the_people.html.
- Ayres, Ian, and John Braithwaite. 1992. *Responsive Regulation: Transcending the Deregulation Debate*. New York, NY: Oxford University Press.
- Ayres, Ian, Yair J. Listokin, and Michael Abramowicz. 2011. "Randomizing Law," 159 *University of Pennsylvania Law Review* 929-1005.
- Bardach, Eugene, and Robert A. Kagan. 1982. *Going by the Book: The Problem of Regulatory Unreasonableness*. Philadelphia: Temple University Press.
- Benbear, Lori Snyder. 2006. "Evaluating Management-Based Regulation: A Valuable Tool in the Regulatory Toolbox?," in Cary Coglianese and Jennifer Nash, eds., *Leveraging the Private Sector: Management-Based Strategies for Improving Environmental Performance*. Washington, DC: RFP Press.
- Benbear, Lori Snyder, and Cary Coglianese. 2012. "Flexible Environmental Regulation" in Sheldon Kamieniecki and Michael Kraft, eds., *The Oxford Handbook of U.S. Environmental Policy*. New York, NY: Oxford University Press.
- Brannigan, Vincent, and Carol Smidts. 1998. *Performance Based Fire Safety Regulation under Intentional Uncertainty*. Available at <http://fire.nist.gov/bfrlpubs/fire98/PDF/f98056.pdf> (accessed February 9, 2015).
- Coglianese, Cary. 2015a. "What Volkswagen Reveals About the Limits of Performance-Based Regulation," *RegBlog* (Oct. 5). Available at <http://www.regblog.org/2015/10/05/coglianese-volkswagen-performance-based-regulation/>.
- Coglianese, Cary. 2015b. "Listening, Learning, and Leading: A Framework for Regulatory Excellence." Penn Program on Regulation. Available at <https://www.law.upenn.edu/live/files/4946-pprfinalconvenersreportpdf>.

- Coglianesse, Cary. 2012. *Measuring Regulatory Performance: Evaluating the Impact of Regulation and Regulatory Policy* (Expert Paper No. 1). Organization of Economic Cooperation and Development.
- Coglianesse, Cary. 2008. "Management-Based Regulation: Implications for Public Policy," OECD Working Party on Regulatory Management and Reform, April 7-8, 2008, GOV/PGC/REG(2008)(5).
- Coglianesse, Cary, and David Lazer. 2003. "Management-Based Regulation: Prescribing Private Management to Achieve Public Goals," *37 Law & Society Review* 691-730.
- Coglianesse, Cary, and Evan Mendelson. 2010. "Meta-Regulation and Self-Regulation" in Robert Baldwin, Martin Cave and Martin Lodge, eds., *The Oxford Handbook of Regulation*. Oxford: Oxford University Press.
- Coglianesse, Cary, Jennifer Nash, and Todd Olmstead. 2003. "Performance-Based Regulation: Prospects and Limitations in Health, Safety, and Environmental Protection," 55 *Administrative Law Review* 705-29.
- Crocker, Linda. 2003. "Teaching for the Test: Validity, Fairness, and Moral Action," 22 *Educational Measurement: Issues and Practice* 5-11.
- Dollmeyer, Thomas A., David A. Vittorio, Thomas A. Grana, James R. Katzenmeyer, Stephen J. Charlton, James Clerc, Robert C. Morphet, and Brian W. Schwandt. 2007. "Meeting the U.S. 2007 Heavy-Duty Diesel Emission Standards – Designing for the Customer," SAE Technical Paper # 2007-01-4170.
- Evans, Douglas H., Robert D. Weber, and James R. Quiter. 1998. "Luxor Hotel and Casino: An Application of Performance-Based Fire Safety Design Methods," Proceedings of the 1998 Pacific Rim Conference and Second International Conference on Performance-Based Codes and Fire Safety Design Methods.
- Executive Order 12,866, 58 Fed. Reg. 51,735-51,744 (October 4, 1993).
- Executive Order 13,563, 3 C.F.R. 215 (2011). Available at <http://www.whitehouse.gov/the-press-office/2011/01/18/improving-regulation-and-regulatory-review-executive-order> (accessed February 16, 2015).
- Federal Motor Vehicle Safety Standards; Occupant Crash Protection, 58 Fed. Reg. 46551 (September 2, 1993).
- Glassbrenner, Donna. 2003. "Estimating the Lives Saved by Safety Belts and Air Bags." 18th Int'l Technical Conference on the Enhanced Safety of Vehicles, Conf. Paper No. 500. Washington, DC: National Highway Traffic Safety Administration.
- Hakim, Danny and Claire Barthelemy. 2015. "VW's Emissions-Test Trickery May Not Be Illegal in Europe," *N.Y. Times* (Nov. 11). Available at http://www.nytimes.com/2015/11/12/business/international/vw-scandal-eu-emissions-tests.html?_r=0.

- Howard, Philip K. 1994. *The Death of Common Sense: How Law is Suffocating America*. New York: Random House.
- Hutter, Bridget. 2001. *Regulation and Risk: Occupational Health and Safety on the Railways*. Oxford: Oxford University Press.
- Hutton, Dame Deidre. 2015. "The Role of Stakeholder Relationships in Regulatory Excellence," Penn Program in Regulation. Available at <https://www.law.upenn.edu/live/files/4727-hutton-keynote-address-penn-law-2015pdf>.
- Isidore, Chris, and Peter Valdes-Dapena. 2015. "EPA Accuses VW of Cheating on Emission Rules," *CNN Money* (Sept. 22). Available at <http://money.cnn.com/2015/09/18/autos/epa-cheating-vw/>.
- Jacobsen, Barbara J., Amy R. Rock, Murray S. Cohn, and Toby Litovitz. 1989. "Accidental Ingestions of Oral Prescription Drugs: A Multicenter Survey," *79 American Journal of Public Health* 853-56
- Jaffe, Adam B., and Robert N. Stavins. 1995. "Dynamic Incentives of Environmental Regulations: The Effects of Alternative Policy Instruments on Technology Diffusion," *29 Journal of Environmental Economics & Management* S34-S63.
- Kleindorfer, Paul R., and Eric W. Orts. 1998. "Informational Regulation of Environmental Risks," *18 Risk Analysis* 155-70.
- Lloyd, Alan C., and Thomas A. Cackette. 2001. "Diesel Engines: Environmental Impact and Control," *51 Journal of the Air & Waste Management Association* 809-47.
- May, Peter. 2003. "Performance-Based Regulation and Regulatory Regimes: The Saga of Leaky Buildings," *25 Law & Policy* 381-401.
- National Center for Statistics & Analysis. 2007. *Counts of Frontal Air Bag Related Fatalities and Seriously Injured Persons*. National Highway Traffic Safety Administration.
- New Zealand Legislation. n.d. *Building Regulations 1992: Schedule 1: The Building Code*. Available at <http://www.legislation.govt.nz/regulation/public/1992/0150/latest/DLM162576.html> (accessed February 9, 2015).
- New Zealand Ministry of Business, Innovation, and Employment. n.d. *The Building Code*. Available at <http://www.dbh.govt.nz/the-building-code> (accessed February 9, 2015).
- Office of Management & Budget. 2003. *Circular A-4: Regulatory Analysis*. Washington, DC. 1-16.
- OECD. 2002. *Regulatory Policies in OECD Countries: From Interventionism to Regulatory Governance*. Organization for Economic Co-Operation and Development.

- OECD. 1997. *Regulatory Reform and Innovation*. Organization for Economic Co-Operation and Development.
- Office of Information and Regulatory Affairs (OIRA). 2003. *Circular A-4: Regulatory Analysis* (September 17). Available at <http://www.whitehouse.gov/omb/circulars/a004/a-4.pdf>.
- Osenga, Mike. 1998. "Diesel Industry Confronts the Emissions Settlement," *Diesel Progress* 42 (North American Edition, Dec. 1998 issue).
- PayPal, and eBay Inc. 2013. *21st Century Regulation: Putting Innovation at the Heart of Payments Regulation*. Available at http://www.ebaymainstreet.com/sites/default/files/PayPal-Payment-Regulations-Booklet_US.pdf (accessed February 9, 2015).
- Plungis, Jeff and Dana Hull. 2015. "VW's Emissions Cheating Found by Curious Clean Air-Group," *BloombergBusiness* (Sept. 19). Available at <http://www.bloomberg.com/news/articles/2015-09-19/volkswagen-emissions-cheating-found-by-curious-clean-air-group>.
- Popham, W. James. 2001. "Teaching to the Test?," 58 *Educational Leadership* 16-20.
- Richards, Kenneth. 2000. "Framing Environmental Policy Instrument Choice," 10 *Duke Environmental Law and Policy Forum* 221-82.
- Shapiro, Sidney A., and Thomas O. McGarity. 1991. "Not So Paradoxical: The Rationale for Technology-Based Regulation," 1991 *Duke Law Journal* 729-752.
- Sherman, Frederick T., Joshua D. Warach, and Leslie S. Libow. 1979. "Child-Resistant Containers for the Elderly?," 241 *Journal of the American Medical Association* 1001-02 (Mar. 9).
- Stavins, Robert N. 2007. A U.S. Cap-and-Trade System to Address Global Climate Change (The Hamilton Project Discussion Paper 2007-13, October). Washington, D.C.: The Brookings Institution.
- Stavins, Robert N. 1998. "What Can We Learn from the Grand Policy Experiment? Lessons from SO2 Allowance Trading," 12 *The Journal of Economic Perspectives* 69-88.
- Sunstein, Cass R. 1999. "Informational Regulation and Informational Standing: Akins and Beyond," 147 *University of Pennsylvania Law Review* 613-75.
- Taylor, Serge. 1984. *Making Bureaucracies Think: The Environmental Impact Statement Strategy of Administrative Reform*, Palo Alto, CA: Stanford University Press.
- Tietenberg, Thomas H. 1985. *Emissions Trading: An Exercise in Reforming Pollution Policy*. Washington, DC: Resources for the Future Press.
- Tietenberg, Tom. 2010. "Cap-and-Trade: The Evolution of an Economic Idea," 39 *Agricultural & Resource Economics Review* 359-67.

Traffic Safety Act of 1966, S. Rep. No. 1301, 89th Cong., 2d Sess. (1966).

WTO Committee on Technical Barriers to Trade (G/TBT/W/341) (September 13, 2011).

WTO Committee on Technical Barriers to Trade (G/TBT/1/Rev.12) (January 21, 2015).

Yandle, Bruce, Andrew P. Morriss, and Lea-Rachel Kosnik. 2009. "Regulation by Litigation: The Diesel Engine Episode" in Andrew P. Morriss, Bruce Yandle, and Andrew Dorchak, eds., *Regulation by Litigation*. New Haven, CT: Yale University Press.

Ziegler, Chris. 2015. "Another Volkswagen Diesel Engine Is Cheating Emissions Tests, EPA Claims," *The Verge* (Nov. 2). Available at <http://www.theverge.com/2015/11/2/9659382/volkswagen-diesel-audi-porsche-3-liter-emissions-cheat-defeat-device-epa>.