Administrative Conference of the United States

REGULATORY EXPERIMENTATION

Final Report: November 17, 2017

Zachary J. Gubler
ASU Sandra Day O'Connor
College of Law

REGULATORY EXPERIMENTATION

Zachary J. Gubler[*]

---

[*]Professor of Law, ASU Sandra Day O'Connor College of Law

When making decisions under conditions of uncertainty, the optimal approach is often trial and error. We see this across diverse fields and activities. In the world of venture capital, start-up companies only qualify for the next stage of financing if they can show success by hitting certain pre-specified milestones.[1] In drug development, FDA approval is premised on the successful use of the drug in clinical trials.[2] In corporate research and development, companies often test a product in local markets before a national or global rollout.[3] In all of these cases, and more, decision-makers rely on an experimental approach because of the useful information that it generates, which can then be used to inform more long-term decisions.

Like venture capitalists, clinical researchers and product specialists, administrative agencies face considerable uncertainty in confronting their regulatory tasks. For example, how does a rule regulating the sale of borrowed stock affect market efficiency?[4] What effect will raising limits on amounts banks can lend to individuals have on the riskiness of loan portfolios?[5] How costly will it be from an administrative perspective to regulate all producers of greenhouse gases of a certain size?[6]

These are but a few of the types of questions that administrative agencies must address every day. Agencies can try to answer them, offering speculation and educated guesses but the accuracy of their responses turn on the information available to them. How should agencies go about getting this information? They might commission studies, where the researcher goes out and canvases the scholarly literature in the area and talks to experts. This is a fairly common occurrence. This report, however, focuses on a different alternative: regulatory experimentation. This is where the agency itself takes some temporary regulatory action in an effort to generate valuable information that can then be used to make more informed, long-term decisions.

This report consists of three parts. In Part I, the report lays out the theoretical and conceptual underpinnings of regulatory experimentation. It discusses how an agency should think about whether or not to undertake a regulatory experiment and how to structure the experiment when it does. It also considers relevant legal issues and how public stakeholders are likely

---

[1] *See* 1 JOSEPH W. BARTLETT, EQUITY FINANCE: VENTURE CAPITAL, BUYOUTS, RESTRUCTURINGS AND REORGANIZATIONS §9.2 (1995).

[2] *See* Richard Simon, *Optimal Two-Stage Designs for Phase II Clinical Trials*, 10 CONTROLLED CLINICAL TRIALS 1, 1 (1989).

[3] *See* ROBERT G. COOPER, WINNING AT NEW PRODUCTS: CREATING VALUE THROUGH INNOVATION 316 (4TH ED. 2011).

[4] *See* Pilot Adopting Release, 69 Fed. Reg. 48,008, 48,024 (Aug. 6, 2004).

[5] *See* Proposed Extension of Lending Limits Pilot Program, 69 Fed. Reg. 21,978 (2004).

[6] *See* Proposed Tailoring Rule, 74 Fed. Reg. 55,292 (2009).

to react to regulatory experiments. In Part II, the report considers various examples of real-world regulatory experiments and compares and contrasts them in search of generalizable lessons. Part III presents these lessons.

I. THEORY AND CONCEPTS

With the emphasis on the states as laboratories of experimentation, one could be forgiven for thinking that the federal government lacks any experimental resources.[7] However, federal agencies have the means at their disposal to also act as powerful laboratories,[8] especially if one understands the   term   "experimentation" fairly broadly.   In this report, an agency engages in experimentation anytime it takes a regulatory action with the express purpose, from the outset, of generating information that would be expected to inform a more permanent decision down the road. Similarly, an "experimental rule" is any rule adopted with that purpose.

Under this definition, all regulatory experimentation is intended from the beginning to be a temporary measure meant to inform more permanent action. In many, if not most cases, the agency will actually structure the rule as a temporary rule, typically subject to an automatic sunset provision that marks the end of the experiment.[9] However, as we shall see, regulatory experiments, although temporary, can also be structured as permanent rules, which can make it somewhat difficult to distinguish a regulatory experiment from a non-experimental agency action.[10]   Putting those structuring concerns aside for the moment, let us first begin with two more fundamental questions: why should agencies engage in regulatory experimentation and how should they make the decision to undertake a regulatory experiment in the first place?

---

[7] *See, e.g.*, Barry Friedman, *Valuing Federalism*, 82 MINN. L. REV. 317 (1997).

[8] *See, e.g.*, Zachary J. Gubler, *Experimental Rules*, 55 B.C. L. REV. 129 (2014).

[9] Many commentators have suggested sunset rules as a means of experimentation. *See, e.g.*, Jacob E. Gersen, *Temporary Legislation*, 74 U. CHI. L. REV. 247, 261 (2007) ("In contexts where initial policy judgments are likely to be inaccurate, temporary legislation has certain advantages over permanent legislation."); Matthew Spitzer & Eric Talley, *On Experimentation and Real Options in Financial Regulation*, 43 J. LEGAL STUD. 121 (2014); Sofia Ranchordas, *Innovation-Friendly Regulation: The Sunset of Regulation, the Sunrise of Innovation*, 55 JURIMETRICS J. 201 (2015); Roberta Romano, *Regulating in the Dark and a Postscript Assessment of the Iron Law of Financial Regulation*, 43 HOFSTRA L. REV. 25 (2014); Alex Lee, *An Options Approach to Agency Rulemaking*, 65 ADMIN. L. REV. 881 (2013).

[10] *See infra* notes 47-67 and accompanying text.

A. *The What, Why and When of Regulatory Experimentation*

Imagine a world without experimentation. It is difficult to do, in no small part because our world is so saturated with it. But if it could be imagined, the result would be radically different from the world we have now. In the absence of experimentation, many of modern civilization's best thoughts and ideas would still be just that: mere thoughts and ideas. Or else, they would be embodied in physical objects that are awkward and unrefined and possibly completely unworkable. Experimentation is valuable because it provides information that is almost impossible to acquire otherwise. Will consumers like a new flavor of cola? Will people be willing to transfer money over the internet? Will the drug actually have the desired effect? These types of questions cannot be answered definitively simply through logic and reasoning. They require a test in real or simulated conditions in an attempt to acquire the necessary information.

Administrative agencies face similar questions all of the time, questions that ultimately cannot be answered through reason and logic alone. Agencies should engage in experimentation then for the reason that other important decision-makers do, because it is the best process that we have for minimizing the uncertainty involved in certain types of decision-making.

However, as with most things, experimentation is not free. There are out-of-pocket costs: you need researchers, subjects, possibly new or newly configured equipment, among other things. There are also important opportunity costs: every minute spent experimenting is a minute not spent doing something else. And clearly not every question is in the same need of an experiment to get at the right answer. While it might take an experiment to figure out whether a drug that has only been used in mice will work well in humans, it would probably be overkill to run an experiment to determine the effects of a generic drug with the same active ingredients as a currently manufactured patented drug.

Given the real costs associated with experimentation, how should agencies think about whether to engage in a regulatory experiment? At a very high level of generality, the agency must weigh the expected benefits of the information that will result from the regulatory experiment against the costs of undertaking the experiment. Let us consider an example before examining these elements in greater detail.

1. Preliminary Example: The SEC's Proxy Access Rule

Imagine that the Securities and Exchange Commission (the "SEC") is trying to decide whether to adopt a rule that would give shareholders of

public companies more of a say in electing the members of their companies' boards of directors.[11]   Under current law, shareholders have the right to nominate and campaign for their own director candidates.[12]   But these "proxy contests" are very costly, and therefore they do not happen very often.[13]   Consequently, the company's (in other words, management's) candidates end up having a huge incumbent advantage.[14] Yet, there are potentially significant benefits from giving shareholder more of a say, particularly if the threat of ouster will cause boards to adopt more shareholder-friendly policies that increase the value of the company.[15] However, there are potential concerns as well. In particular, how will such newly discovered power be used, particularly by shareholders, like pension funds, that might not have the same interests as the average shareholder?[16]

Let us say that the SEC determines that if the rule ends up "working" — in this case, if shareholder empowerment ends up disciplining slack management without empowering special interests — corporate constituencies will gain significantly over the status quo. But if the rule does not work, then the status quo would be preferable.

Let us put some numbers on these possibilities just to make it more concrete.[17]   Let us assume that, if it is successful, the "proxy access rule" — the rule giving shareholders a more significant say in director elections

---

[11] This example is based on an actual rulemaking that the SEC undertook in 2010. More about this below. *See infra* notes 86-106 and accompanying text.

[12] Leo E. Strine, Jr., *The Dangers of Denial: The Need for A Clear-Eyed Understanding of the Power and Accountability Structure Established by the Delaware General Corporation Law*, 50 WAKE FOREST L. REV. 761, 777 (2015)

[13] Andrew M. Schwartz, *Financing Corporate Elections*, 41 J. Corp. L. 863, 876 (2016) ("Even as far back as the 1950s, a proxy campaign could cost tens of thousands of dollars on each side, or even $100,000 or more in a heated contest. By the turn of the 21st century, proxy contests routinely cost hundreds of thousands of dollars, and million-dollar expenses were not unheard of. Today, in the heyday of the activist investor, a full-on proxy contest87 can easily run several million dollars or even more.").

[14] *See, e.g.*, Leo E. Strine, Jr., *The Professorial Bear Hug: The ESB Proposal as a Conscious Effort to Make the Delaware Courts Confront the Basic "Just Say No" Question*, 55 STAN. L. REV. 863, 864 n.6 (2002) ("[I]t must be remembered that incumbent directors have a large advantage in proxy contests because they can spend the company's money to tell their story and retain important, if not unconstrained, discretion regarding timing.").

[15] *See, e.g.*, Lucian Arye Bebchuk, *The Case for Increasing Shareholder Power*, 118 HARV. L. REV. 833 (2005).

[16] *See, e.g.*, Christine Hall, Dodd-Corker Fed Bill May Contain Left-Wing "Shareholder" Power Grab, Competitive Enterprise Inst. (Mar. 3, 2010), http://cei.org/news-releases/dodd-corker-fed-bill-may-contain-left-wing-shareholder-power-grab (reporting that seventeen groups expressed these concerns in opposition to the Dodd-Frank Act's proposed proxy access provision).

[17] The following illustration is adapted from Yair Listokin, *Learning Through Policy Variation*, 118 YALE L.J. 480 (2008).

— is worth $100 million more in value per year than the status quo. Assuming that the rule does not end up serving as a mere vehicle for special interest shareholders, the rule will increase shareholder value considerably by creating better managed boards of directors. However, this is not a sure thing. This only happens if the rule is successful, and let us assume that the probability of this happening is low, maybe only 20%. If, on the other hand, the rule ends up being a failure — a vehicle for special interest shareholders to meddle in otherwise relatively well managed boards — the value of the rule is $50 million less than the status quo. In other words, corporate shareholders, and probably other stake holders as well, would actually be worse off if that happened, and let us assume the probability of failure in that case is 80% (100%, minus the 20% probability of success).

Given these numbers, let us think about the value of regulatory experimentation in this context. How should the agency — the SEC in this case — think about whether to adopt this risky rule as opposed to just sticking with the status quo? That depends on whether the rule is adopted as a regulatory experiment or not. Let us first assume that it is just adopted as a regular permanent rule as part of an informal rulemaking. In that case, the agency must subject the rule to the notice and comment procedures required by the Administrative Procedure Act (the "APA"), the result of which is a final rule.[18] If it turns out that the rule is a failure, the agency could always repeal it, but not unilaterally.[19] They would have to initiate notice and comment procedures again, and by then, there might be stakeholder groups that favor the rule, despite its failure. In other words, final rules like these, although technically reversible by subsequent agency action, tend to have a lot of inertia built into them.[20] Despite their susceptibility to repeal, they have a tendency to be permanent.

Assuming this, should the SEC adopt the proxy access rule or not? Surely not, at least not if the rule is likely to be permanent. After all, the expected value of the proxy access rule is actually negative when compared to the status quo. It is worth $100 million with 20% probability (an expected value of $20 million) and -$50 million with 80% probability (an expected value of -$40 million) for a total expected value of -$20 million ($20 million + -$40 million) less than the status quo. In that case, the SEC would be better off simply sticking with the status quo.

But what if the SEC can conduct a regulatory experiment? In other words, what if they can adopt the proxy access rule on a temporary basis to

---

[18] *See* 5 U.S.C. §555 (2017).

[19] Perez v. Mortgage Bankers Ass'n, 135 S. Ct. 1199, 1206 (2015) ("The D.C. Circuit correctly read § 1 of the APA to mandate that agencies use the same procedures when they amend or repeal a rule as they used to issue the rule in the first instance.").

[20] *See* Gubler, *supra* note 8, at 139-40.

generate information relevant to the decision about whether to adopt it on a more permanent basis later on? (Presumably the relevant information in question has to do with the extent to which the rule will be used as a vehicle for special interest shareholders to gain board representation.) In that case, the calculation changes dramatically. If the rule works, and the SEC decides to adopt it on a permanent basis, then it generates $100 million per year. And if, at the end of the experiment, the rule is deemed not to have worked, the SEC can simply revert to the status quo. In other words, the experiment would yield an 20% chance of an infinite stream of $100 million per year and a 80% chance of a mere few years' worth of payouts of -$50 million during the experimental period until the agency deems it a failure and reverts to the status quo. In that case, it is preferable to choose the experiment.[21]

2. Variables Influencing the Regulatory Experimentation Decision

The proxy access rule is an example of informal rulemaking, and this report will focus on informal rules because of their widespread use among agencies.[22] However, many of the lessons this report uncovers can be generalized to other forms of regulatory action. To uncover these lessons, let us consider the definition of regulatory experimentation that has just been introduced. Specifically, that definition contains three main variables: the benefits of experimentation, the risk of experimentation and the costs of experimentation. Let us now consider each one of these in greater detail:

a. Expected Net Benefits

The net benefits of a rule consist of the rule's benefits, minus its costs. So, for example in the proxy access rule example, the benefit of the rule is enhanced monitoring of the public company board, which increases shareholder value. The cost of the rule has to do with empowering shareholders that have special interests, which might decrease shareholder value overall. The difference between these two values is the rule's net benefit.[23]

---

[21] For a mathematical illustration of this example, with different values, see Listokin, *supra* note 17, at 494.

[22] Vanessa K. Burrows & Todd Garvey, Cong. Research Serv., R41546, A Brief Overview of Rulemaking and Judicial Review 1-2 (2011).

[23] We also see these net benefits at work in the various examples of regulatory experimentation discussed in greater detail below. In the SEC's short sale experiment, for example, the benefit of deregulation is increased market efficiency whereas the cost is market volatility and manipulation. *See infra* notes 120-150 and accompanying text. In the OCC's lending limits example, the benefit is improved competition among smaller regional

However, as is the case with most things in life, lawmaking rarely deals with certainties. It would be the rare case for an agency to be able to calculate the net benefits of a rule without resorting to probabilities. More often than not, the net benefits must be discounted by the likelihood of different scenarios or states of the world. The result is the *expected* net benefits of the rule. In the proxy access example, we envisioned two different states of the world, one where the net benefit was $100 million (because the benefits of proxy access end up being significantly greater than its costs) and another state of the world where the net benefit was -$50 million (because the benefits of proxy access end up being significantly less than its costs). We then calculated the expected net benefit of the rule by discounting these two states of the world by our estimate of the probability that each state of the world would materialize. We said that there was an 80% chance the rule would yield a net benefit of -$50 million per year and a 20% chance that it would yield a net benefit of $100 million per year. The expected net benefit on an annual basis is therefore -$20 million.

Of course, all regulatory actions involve expected net benefits, not just regulatory experiments. However, as illustrated by the proxy access example, the expected net benefits need to be thought about differently in the regulatory experimentation context. In the proxy access example, the rule could produce a net benefit of $100 million per year, but only with a 20% likelihood. Although the rule is potentially quite valuable, the expected net benefits of the rule gets weighed down because of the low probability of this scenario materializing. If we were to make a regulatory decision based solely on the expected net benefit of the proxy access rule, which is -$20 million, undoubtedly we would choose to pass on the rule.

But we need to think about this calculation differently when dealing with a regulatory experiment. This is because a regulatory experiment effectively minimizes the bad state of the world — in our example, the state of the world where the rule produces -$50 million in value — by limiting the time period during which the rule is in force. In the case of the proxy access rule, when the rule is structured as an experiment, the agency anticipates from the outset the possibility that the rule will fail, or in other words that the bad state of the world will materialize, and plans for that contingency. If the experiment reveals that bad state of the world, then the agency can shut down the experiment, thereby minimizing the effects of that bad state of the world and restoring the status quo or otherwise reverting to the next best alternative. The result is that the experiment

---

banks and the cost is increased loan portfolio risk. *See infra* notes 190-219 and accompanying text. In the EPA's greenhouse gas experiment, the benefit is improved air quality and the cost is largely the administrative burden resulting from an increased regulatory scope. *See infra* notes 219-229 and accompanying text.

allows the agency to capture, on a probabilistic basis, the good state of the world — an infinite stream of payments of $100 million per year — while at the same time reducing the bad state of the world — an infinite stream of payments of -$50 million per year — to the few years that the experiment is in force. Thus, although the concept of net benefits isn't unique to the experimental regulation context, the way net benefits are calculated are different in experimental regulation context.

One might point out that this discussion overlooks another type of benefit unique to regulatory experiments: the informational benefits generated by the experiment, even if the rule at the heart of the experiment is not ultimately adopted on a permanent basis following the experiment.[24] One might refer to these as "experimental benefits." A common catchphrase among entrepreneurs is to "fail fast, fail often." The idea is that even failed experiments are valuable because they convey important information. For example, in the proxy access context, maybe the experiment proves that in fact the proxy access rule does result in an inordinate shift of power to special interests, and consequently the SEC decides not to adopt the rule on a permanent basis. But perhaps the experiment implies that a different rule — perhaps one that exempts certain types of special interest investors from the rule's applicability — would be a success, and the SEC could consider adopting that type of rule. In that case, even though the rule that is the subject of the experiment is a failure, the experiment itself ends up being valuable.

While these experimental benefits are undoubtedly real, they are not always going to be susceptible of quantification. Indeed, as a general matter, it is likely that the experimental benefits are going to be more speculative than the net benefits of the rule discussed previously. What one ends up learning from a failure depends on the precise nature of the failure.

---

[24] For example, the SEC's recent tick-size pilot program sought to increase the price increment at which certain small company stocks trade from $0.01 to $0.05. The hypothesis was that the greater increment would increase the "spread" – in other words, the difference between the bid and ask price – and increase the broker's commission, which would in turn incentivize brokers to spend more time marketing these small companies. The pilot was adopted against the backdrop of a decades-long slump in IPO's and therefore an attempt to reverse this trend. *See* Order Directing the Exchanges and the Financial Industry Regulatory Authority To Submit a Tick Size Pilot Plan, 79 Fed. Reg. 36,840 (June 30, 2014). However, early reports suggest that the SEC's hypothesis has not been borne out by the pilot, which suggests that it is unlikely the SEC will adopt the rule change on a permanent basis. *See, e.g.*, Rick Baert, *Tick-Size Pilot Disaapointment Has Experts Searching for Alternatives*, Pensions & Investments (June 21, 2017), *available at* http://www.pionline.com/article/20170621/ONLINE/170629944/tick-size-pilot-disappointment-has-experts-searching-for-alternatives. Nevertheless, this information is valuable in that it allows one to eliminate a leading hypothesis from the list of possible reasons for the fall in IPOs.

And therefore, in order to calculate the experimental benefits, one needs to have an estimate not simply of the likelihood of failure but of the different ways in which the experiment might fail. By contrast, it is not necessary to get into these types of fine-grained estimates with respect to the more general net benefits discussed previously, which turns on the probability and magnitude in the good state of the world ("success") and the bad state of the world ("failure").

Given the speculative nature of experimental benefits, it will often, although not always, be less costly for the agency to simply not include them in the net benefit calculation. This is the approach taken in the example above. In that example, the proxy access rule is assumed to yield a $100 million benefit in the event it is successful and a -$50 million loss in the event it is a failure. However, importantly, that -$50 million loss does not reflect the benefits that might accrue if the experiment yields some valuable information, which then leads the agency to take some other regulatory action that is itself very valuable. In cases like this where it is not reasonable to include estimates of experimental benefits in the calculation, it is important to recognize that this is a conservative approach that will understate the benefits of experimentation.

b. The costs of experimentation

Once one has determined that the expected net benefits of a regulatory experiment are greater than those of the status quo or the next best alternative, the question then is whether those expected net benefits from experimentation outweigh the costs of the experiment. There are two principal types of costs associated with a regulatory experiment, implementation costs and disruption costs.

i. Implementation costs

The primary source of implementation costs is that the agency must design a process that is consistent with the various procedures established by Congress and the judiciary and applicable to informal rulemaking. The bulk of these costs have to do with creating an administrative record that will survive "hard look review," which "require[s] agencies to offer detailed explanations for their decisions, to provide strong justifications for any departures from past decisions, to permit widespread public participation in the rulemaking process, and to consider alternative regulatory measures to those proposed."[25] Of course, all informal rulemaking requires this process.

---

[25] Jason Webb Yackee & Susan Webb Yackee, *Testing the Ossification Thesis: An Empirical Examination of Federal Regulatory Volume and Speed, 1950-1990*, 80 GEO.

Why should an experimental rulemaking be different? One possibility is if judicial review is stricter for regulatory experiments than non-experimental rules. However, this seems unlikely to be the case.[26]

Rather, the reason why experimental rules might entail greater implementation costs than their non-experimental counterparts is because adopting a rule on an experimental basis will often require the agency to go through these procedures twice, once when they adopt the experiment and a second time when they adopt whatever permanent rule is justified in light of the experimental results.[27] To be sure, there are actions the agency can take to structure the experiment so as to minimize these double procedural costs.[28] But it is still a cost of experimentation.

ii. Disruption costs

Disruption costs consist of the transaction and information costs associated with regulated entities having to understand and comply with new rules. Rule changes require regulated entities to hire lawyers to help them interpret and apply the new rules. They might also have to hire more employees to help them comply on a going-forward basis.

These disruption costs are particularly high in the regulatory experimentation context for the same reason that implementation costs are particularly high in that same context: regulatory experiments are temporary and therefore by definition entail more regulatory action than permanent rules. To the extent that increased regulatory action involves an increased number of rule changes, the costs of understanding and complying with these changes (the disruption costs) increase as well.[29]

In light of these various costs — implementation costs and disruption costs — there will be a desire on the part of the agency to minimize them. There are several approaches agencies might take. First, they might minimize costs by paying attention to the structure of the experiment, as discussed in greater detail below.[30] The optimal structure will be one that is consistent with the most likely outcome of the experiment: adoption of the experimental rule on a permanent basis or reversion to the status quo. This will minimize implementation and disruption costs by avoiding unnecessary regulatory actions, which would occur for example if the likelihood is that

---

WASH. L. REV. 1414, 1426 (2012).

[26] *See infra* notes 72-106 and accompanying text.

[27] *See supra* notes 18-19 and accompanying text.

[28] *See infra* notes 47-72 and accompanying text.

[29] *See, e.g.*, Rebecca Kysar, *Lasting Legislation*, at 1064 (describing how temporary tax legislation might cause taxpayers to "obtain costly tax advice to shift income and deductions between years in avoidance of sunset date").

[30] *See infra* notes 48-67 and accompanying text.

an experimental rule would be temporary and yet the agency adopts it as a permanent rule. Second, they can reduce disruption costs by extending the expiration of experiments where agencies have not yet determined the subsequent regulatory action.[31]   Third, they can reduce implementation costs by using interim final rules, where justified, when a regulatory action follows a regulatory experiment.[32]

c. Summary

To summarize, an experiment is a regulatory action where failure is anticipated and planned for from the beginning, which allows the agency to minimize the downside risk of the rule that is the subject of the experiment. This suggests two observations: First, regulatory experiments might require agencies to become more comfortable with the risk of failure than they currently are. This might require something of a culture change, at least with respect to the independent agencies that are not subject to OIRA review and the type of formal cost benefit analysis required by Executive Order 12,866.[33]   With respect to the independent agencies, the only place where their efforts at reasoned decision-making are memorialized is the concise statement of basis and purpose.[34]   This document typically reads more like a legal brief than anything else: responding to plausible arguments about the potential effects of a given agency action with other plausible arguments about why those predicted effects are not that big of a deal or are not that likely.[35]

In other words, many independent agencies have a tendency to justify their informal rulemakings by explaining why the risks are minimal. Sometimes, it truly is the case that the risks are minimal. But other times, this is not exactly what agencies mean. Instead, they mean that the risks are potentially significant but, in light of the potential benefits, worth accepting, at least on a temporary basis. Even though the agency might characterize risk as minimal in each case, an experimental approach is probably only justified in the latter case but not in the former. The experimental approach to regulation would require a greater willingness on the part of agencies to embrace, and frankly, acknowledge, risk.

The second observation to be made emerges from the important role

---

[31] For an example of this, see the description of the SEC's short sale experiment, discussed below. *See infra* notes 120-150 and accompanying text.

[32] *See infra* notes 113-117 and accompanying text.

[33] Exec. Order No. 12,866, 3 C.F.R. 638 (1993), reprinted as amended in 5 U.S.C. § 601 (2017).

[34] *See* 5 U.S.C. § 553(c) (2017).

[35] *See* Zachary J. Gubler, *Making Experimental Rules Work*, 3 ADMIN. L. REV. 551 (2015).

that risk plays in regulatory experimentation. Regulatory experiments only make sense if a given rule is risky. Thus, it becomes crucial for agencies faced with the decision whether to engage in a regulatory experiment to get a handle on the risk involved. This likely will require the use of experts. To be sure, the notice and comment process gives a sense of risks. But those submissions are often, although not always, written by lawyers. The agency would be well served by polling experts in the field — economists, scientists and others — about the risk presented by a given rule.

3. A More Formal Model of Regulatory Experimentation

Now that we have considered in greater detail the two elements of regulatory experimentation, we can now revisit our example from before, this time in a more generalized version.

Figure 1: Value of the Status Quo
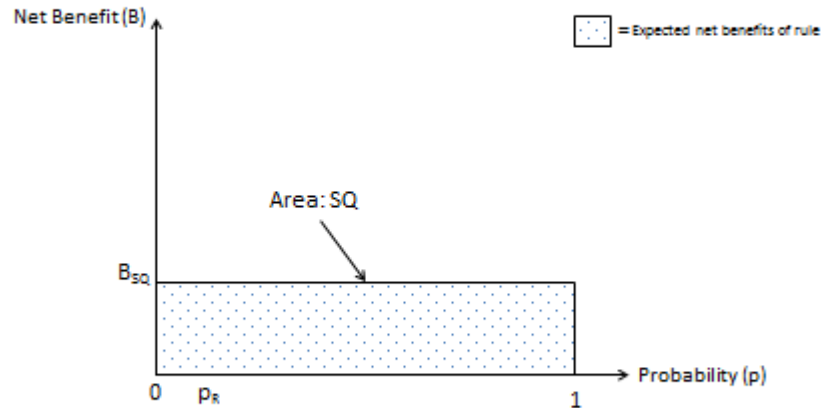
Figure 1 depicts the expected net benefits of the non-risky rule (call it the "status quo"), which is illustrated by the rectangle formed by the points B_{SQ} on the Y axis and the probability of 1 on the Y axis. That is the expected value of all the future annual net benefits generated by the status quo, which is "discounted" by a 1 since there is really no risk here: we already know what the effect the rule will have on things since we're already living with it. Let us call this value "Area SQ."

Now, let us consider Figure 2 below.

Figure 2: Value of Risky Rule R Without Experimentation
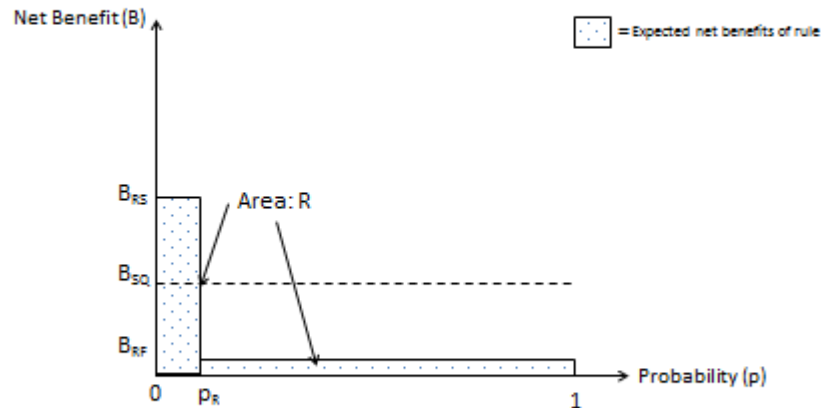
Let us say that there is some rule "R". This rule is like the proxy access rule in the example above. In other words, it is riskier than the status quo, because we do not know what effect it will have on markets, firms, consumers, suppliers and other economic actors. If the effect is positive, then the net benefit is B_{RS}, "RS" to designate the case where the risky rule is "successful." Otherwise, if the risky rule is a failure, "RF," then the net benefit is B_{RF}, which is much less than the net benefit of the status quo, B_{SQ}. The probability that the net benefit of the risky rule R will be the higher value, B_{RS}, is p_R. Therefore the expected net benefit of rule F is the sum of B_{RS}, discounted by probability p_R, and B_{RF}, discounted by $1-p_R$. This expected value is represented by the two shaded rectangles. Let us call this combined value "Area R." This represents the expected value of adopting Rule R on a non-experimental basis. If an agency were to make this decision, then clearly it is sub-optimal since the two rectangles comprising Area R in Figure 2 is less than Area SQ in Figure 1.

However, what if the agency can experiment with Rule R, adopting it on an experimental basis in order to generate information to be used to assess the desirability of maintaining the rule on a permanent basis? In that case,

the agency would capture all of Area R and more. Here's why: As in our proxy access example previously, the net benefit values in these figures reflect the net benefit of adopting a rule on a permanent basis. It is the sum of the infinite stream of annual net benefits the rule generates. With respect to the risky rule, the net benefits are discounted by the probability that different states of the world will arise, reflecting the risky nature of the rule. But if one of those states of the world arises, we are assuming that we are stuck with that result permanently. However, this is not the case with the regulatory experiment illustrated in Figure 3 below.

Figure 3: Value of Experimenting with Risky Rule R
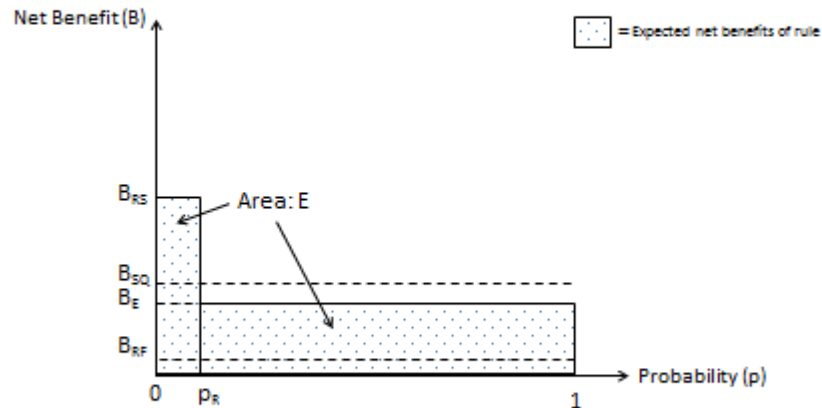
In the case of the regulatory experiment depicted in Figure 3, in the event that that the rule is a "failure," meaning the bad state of the world materializes, we are not limited to the pitifully small net benefits represented by the narrow, horizontal shaded rectangle we saw in Figure 2. Why? Because in the case of the regulatory experiment in Figure 3, if the rule is a failure, we will simply revert to the status quo once the experiment has run its course.   For this reason, the area of the horizontal shaded rectangle in Figure 3 is much greater than in Figure 2. Note that even though we revert to the status quo at the end of the experimental period in the event that the risky rule R is a failure, we don't capture all of the value of the status quo, as illustrated by the narrow unshaded space between the horizontal shaded rectangle and the dashed line indicating the net benefit of

the status quo, B$_{SQ}$. To see why, consider Figure 4 below.

Figure 4: Choosing Whether to Experiment

Rule: Choose the experiment if (x-y) > (experimental costs)
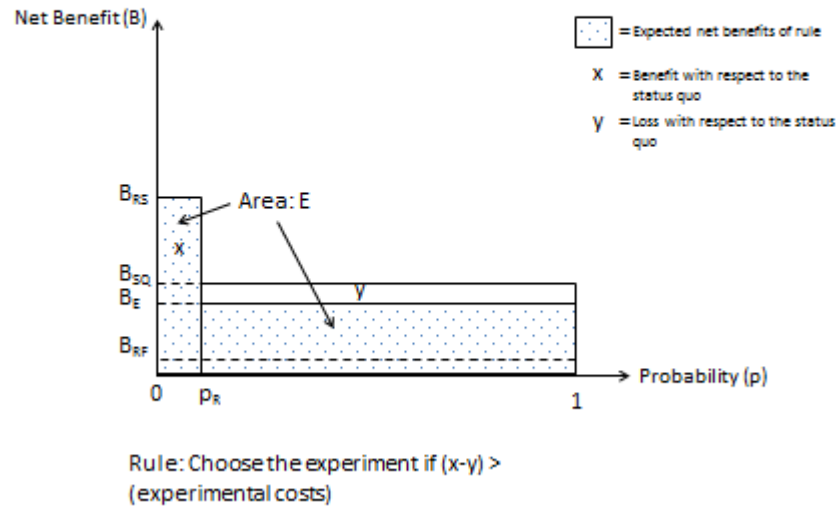
In Figure 4, we see the unshaded horizontal rectangle designated "y." Under the status quo, we would have captured that value, which we can confirm by referring back to Figure 1. However, we don't capture that value under the regulatory experiment. Why not? The answer is because that represents the loss in value, relative to the status quo, that results if the risky rule turns out to be a failure. To be sure, in that event, the failed rule is only in place for a short time, only the duration of the experiment. Nevertheless, that is still some amount of time living with a rule that is less preferable than the status quo. However, note that the experiment also allows us to gain the area designated "x" above the status quo line. That area represents the expected net benefits of the risky rule in the event that it turns out to be a success. Figure 4 therefore reflects the fundamental trade-off of a regulatory experiment: A regulatory experiment allows one to capture the expected net benefits of a potentially highly valuable rule at the cost of only having to live with that rule for a short period of time (during the experimental period) in the event that it proves to be a failure. An agency should undertake the regulatory experiment as long as that expected benefit relative to the status quo (area "x") exceeds the expected loss relative to the status quo (area "y") and that that excess is greater than the costs of the experiment itself.

4. Practical Mental Model for Determining Justifiability of
   Experimentation

It is of course one thing to explain these concepts theoretically within the context of a highly stylized example. It is quite another thing to actually apply them in the real-life act of rulemaking. So far, this discussion suggests that agencies need to be able to quantify the value of a proposed rule in different states of the world and be able to compare that value to the status quo or the next best alternative. How should agencies go about doing this? There is little question that quantification will sometimes, maybe even often, be difficult. The good news is that it is also not always going to be necessary.

As a general matter, the above discussion suggests that it probably makes sense to run a regulatory experiment if (1) the net benefits if the rule is an unlikely success are significant relative to the next best alternative, (2) the net benefits of the risky rule, in the more likely case that the rule is a failure, are not catastrophically low, even though they might be low relative to the next best alternative, and (3) the costs of experimentation are relatively minimal.[36]   In other words, to re-purpose a famous investing adage, it probably makes sense to run a regulatory experiment if one can say about the rule in question, "Heads we win big; tails, we might lose, possibly even significantly, but not catastrophically."[37]

The great benefit of experimentation is that it places a constraint on the downside risk of whatever happens to be the decision at hand. For example, if some new car engine has the potential of extreme fuel efficiency but also runs the risk of blowing up, it hardly makes sense to make the engine standard in all new production vehicles. But that downside risk can be minimized through an experiment, a limited trial involving the engine, where, if the downside risk ends up materializing, development of the engine can be easily reversed.

B. *Structuring a Regulatory Experiment*

Regulatory experiments can be structured in different ways. To understand the available alternatives, let us revisit our definition of regulatory experimentation from before. Under that definition, a regulatory

---

[36] Costs can be minimized through the structure of the experimentation. *See infra* notes 50-67 and accompanying text.

[37] *See* MOHNISH PABRAI, THE DHANDHO INVESTOR: THE LOW-RISK VALUE METHOD TO HIGH RETURNS 12 (2007) (summing up the value approach to investing as, "Heads I win; tails, I don't lose much!").

experiment is a regulatory action designed with the express purpose, from the outset, of generating information that would be expected to inform a more permanent decision down the road. There are three elements to this definition: (1) ex ante specification of some hypothesis to be tested, (2) a rule change that would allow for such a test, and (3) reversibility of that change in the event that it is determined that the rule change is not cost-effective.

1. The Elements of Regulatory Experimentation

    a. Ex ante specification

The Scientific Approach. Our definition of regulatory experimentation includes the requirement that there be some plan from the outset about what it is the agency is actually trying to test with its experiment. This "ex ante specification" requirement could take a number of different forms. On one end of the spectrum is the approach taken by the hard sciences, which requires, on an ex ante basis, the identification of a hypothesis and the specification of a theoretically-driven model to be tested.[38] Depending on the results of the test, the hypothesis is then rejected or it fails to be rejected.[39] This is how clinical drug trials are typically conducted.[40] The design of the experiment contains the instructions for the subsequent action. If the results are positive, then that necessitates one action (perhaps expanding the size of the trial or even approval to bring the drug to market), and if negative, that necessitates another action (perhaps resulting in the elimination of the research altogether).[41]

The Exploratory Approach. On the other end of the spectrum is a much looser, more exploratory form of experimentation. Rather than identifying a particular hypothesis to be tested, this approach instead identifies some type of data set, in this case a data set that will be generated by the experiment.[42] Armed with this data, the person running the experiment then

---

[38] *See, e.g.*, DAVID S. MOORE & GEORGE P. MCCABE, INTRODUCTION TO THE PRACTICE OF STATISTICS 427 (5TH ED. 2006).

[39] *See, e.g.*, *id.*

[40] *See, e.g.*, Suzanne White Junod, FDA and Clinical Drug Trials, https://www.fda.gov/aboutfda/whatwedo/history/overviews/ucm304485.htm (last visited August 17, 2017).

[41] *See id.* (explaining how the scientific community came to the conclusion that "[clinical] trials should have a prospectively defined and identified endpoint, a real hypothesis and an actual analytical plan").

[42] *See, e.g.*, Liane Colonna, *A Taxonomy and Classification of Data Mining*, 16 SMU SCI. & TECH. L. REV. 316-17 (2013).

uses the data to find some interesting statistical inferences to be drawn.[43]

The In-Between Approach. One can also imagine other approaches to the ex ante specification requirement that fall somewhere in between these two ends of the spectrum. For example, one might identify a hypothesis to be tested prior to gathering the data, in contrast to the exploratory approach. However, unlike the scientific approach, one might not pre-specify what will count as a "successful" or "unsuccessful" experiment, for example by failing to identify the confidence intervals required to determine whether to reject or fail to reject the hypothesis.

Probably anything other than the scientific approach would be viewed by scientists as a form of "data mining," which is to say the "practice of examining the data after they have been collected for statistically significant differences in outcomes that were not pre-specified in the hope of finding statistical significance somewhere."[44]   This is certainly true of the exploratory approach.   But it is also probably true of the in-between approach, since that approach, while pre-specifying a hypothesis, does not identify the circumstances under which the hypothesis should be rejected or not.

As a general matter, agencies are probably reluctant to adopt the scientific approach to the ex-ante specification requirement, and, indeed, this generalization is supported by the examples of regulatory experimentation discussed below.[45]   While this practice might seem less than optimal, it is understandable in light of the nature of regulatory decision-making and the reality of budget constraints. Regulatory decision-making typically presents highly complex problems with multiple variables. Rarely does the entire regulatory problem boil down to something that could easily be tested in a single experiment with a single hypothesis. For this reason, it is understandable that an agency would not pre-commit to particular regulatory actions depending on the results of the experiment, which disqualifies the scientific approach. Furthermore, given the budget

---

[43] *See id.*

[44] D. James Greiner et al., *The Limits of Unbundled Legal Assistance : A Randomized Study in a Massachusetts District Court and Prospects for the Future*, 126 HARV. L. REV. 901, 984 (2013).

[45] Not one of them adopts the scientific approach. For example, in the short sale pilot, the SEC really adopted something resembling the exploratory approach. While it pre-specified a hypothesis (de-regulating short sales will lead to bear raids), in the end, this was not the hypothesis that they tested. *See supra* notes 11-21 and accompanying text. In the lending limits experiment, the OCC seems to have adopted the in-between approach, since they pre-specified a hypothesis (relaxing lending limits will increase the riskiness of loan portfolios among small regional banks) but they never identified the criteria they would use for evaluating whether the hypothesis should be rejected or not. *See infra* notes 190-219 and accompanying text.

constraints faced by agencies, there would be significant tradeoffs involved in taking a strictly scientific approach, tradeoffs that might look very different and yield different conclusions than those faced by clinical drug trials, for example.

b. A rule change

An experiment requires some intervention in a sample population in an effort to see how that intervention affects some relevant variable. In the case of regulatory experimentation, the intervention is not the administration of a new drug as in a clinical trial but rather the introduction of a new rule.

c. Reversibility

The value of an experiment in the regulatory context derives from the fact that the rule being adopted is only temporary. The experiment allows one to capture, as a statistical matter, the value of the rule in a world where the rule is a success and it does so without having to endure the costs of the world where the rule is a failure.[46] However, this only works if the experiment can be ended and the rule reversed.

To be sure, all administrative rules can be reversed. However, some are less costly to reverse than others. In particular, there is a distinction to be made between temporary or sunset rules, which are rules that automatically expire upon the occurrence of some date, and permanent rules, which do not.[47] Because of the procedural costs of informal rules, including the costs of establishing an administrative record that can survive hard look review,[48] it is more costly to reverse a permanent rule than a temporary one.[49]

2. Different Structures

---

[46] For a concrete example, think back to the proxy access rule example. *See supra* notes 11-21 and accompanying text. The rule there is worth a lot in one state of the world and very little in another. The experiment allows you to capture the statistical benefits of the first state but not the second, since the experiment can always be ended and one can revert to the status quo.

[47] *See* Gersen, *supra* note 9, at 247 (defining a "sunset" as a clause included in a law limiting the duration of that law's validity).

[48] *See id.*

[49] Of course, it is also more costly to make a temporary rule permanent than to leave a permanent rule in place. For this reason, in structuring regulatory experiment, agencies should choose a structure that is consistent with the probabilities of the experimental outcomes. *See infra* notes 50-67 and accompanying text.

These three aspects of the definition of regulatory experimentation give rise to four different ways to structure a regulatory experiment. These different structures depend on whether the experiment is randomized or not and whether the rule being tested is adopted as a temporary (or "sunset") rule or a permanent rule. By "permanent rule," we mean a rule that must be modified or repealed through some regulatory action as opposed to automatic termination pursuant to a sunset provision. The four different structures are as follows: (1) randomized trial + sunset; (2) randomized trial + permanent; (3) non-randomized + sunset and (4) non-randomized + permanent rule.

3. Choosing Among These Structures

    a. Randomization v. Non-Randomization

A randomized trial solves a problem that is inherent in all cases of experimentation, the problem of creating a control group.[50] Imagine that you want to test whether a new drug actually achieves its intended effect, which is lowering one's blood pressure. The idea is to create a treatment group, whose members will be administered the blood pressure drug, and a control group, whose members will be administered a placebo. Then, the blood pressure of the members in the various groups will be measured to see if there is any statistically significant difference between them.

Of course, you realize that in order to test the effect of the drug, and only the drug, on blood pressure, you need to eliminate all other variables that might affect blood pressure, including things like weight and age. How do you go about controlling for these factors? One possibility is to make sure that the control and treatment groups are identical in every relevant respect other than whether the subjects are administered the drug. One could do this in the real world by actually assembling a bunch of people who are virtual clones along the relevant factual dimensions, or, failing that unlikely scenario, one could accomplish the same thing through statistics.[51] This is what regression analysis does: it is a statistical method that holds certain variables of a sample of the population fixed (like weight and age) in an effort to examine how the variable of interest (in this case, blood pressure) changes with the isolated independent variable (in this case, whether the drug is administered or not).[52]

But here's the problem: it is very difficult to make sure that the model that you are testing has not omitted some important variable, thereby

---

[50] *See, e.g.*, Ian Ayres et al., *Randomizing Law*, 159 U. PA. L. REV. 929. 934-36 (2011).
[51] *See id.*
[52] *See, e.g.*, PETER KENNEDY, A GUIDE TO ECONOMETRICS (2008)

biasing the results.[53] For example, what if it turns out, as is likely the case, that not just weight but diet is an important determinant of blood pressure, and yet you fail to control for that variable? In that case, your results are not going to be reliable.

An alternative approach would be to gather a random group of subjects and randomly decide who gets the blood pressure drug and who gets the placebo. The genius of randomization is that it creates the necessary controlled environment without the experiment designer having to identify all of the various factors that could affect the dependent variable (in this case blood pressure).[54] It does so by introducing chance variation into the mix.[55] "If interventions for patients are chosen by chance, then the law of large numbers implies that the average values of patient characteristics should be roughly equal in the intervention groups."[56] In other words, randomization ensures not that the control and treatment groups will be identical but that the distribution of relevant independent variables in each group (those variables like weight, age and diet that might affect blood pressure) will be the same, and it does so without requiring the experimenter to be able to even identify, let alone measure, those factors.[57]

For this reason, randomized trials are the gold standard of experimentation.[58] However, it will not always be possible to engage in randomization with respect to regulatory experimentation. Depending on the context, the costs of experimentation might be greater for a randomized trial than a non-randomized one.[59] There might also be concerns raised among regulated entities about fairness and justice, particularly if the costs of experimentation are borne in part by the regulated entities.[60] Finally, there might be judicial obstacles to randomized regulatory experimentation. After all, from a certain perspective, it would appear that a rule that applies randomly to one group of regulated entities and not to another is the example *par excellence* of arbitrary and capricious agency action, which is prohibited under the APA.[61] It is argued below that this is actually an

---

[53] *See id.*

[54] *See, e.g.*, Ayres et al., *supra* note 50, at 936.

[55] *See id.*

[56] David P. Harrington, *The Randomized Clinical Trial*, 95 J. Am. Stat. Ass'n 312, 312 (2000).

[57] *See, e.g.*, Ayres et al., *supra* note 50, at 936.

[58] *See, e.g.*, D. James Greiner & Cassandra Wolos Pattanayak, *Randomized Evaluation in Legal Assistance: What Difference Does Representation (Offer and Actual Use) Make?*, 121 Yale L.J. 2118, 2122 (2012).

[59] This might have to do with the disruption costs associated with randomization.

[60] *See, e.g.*, Ayres et al., *supra* note 50, at 968-75.

[61] After all, where one must choose among a fixed set of things, there really is no difference in saying that one chooses arbitrarily or one chooses randomly.

incorrect way of considering the legal issue, and that randomized regulatory experiments should not pose any greater legal challenge than non-randomized ones.[62] However, there is no guarantee that courts will agree with this assessment, especially in light of the intuitive nature of the argument that randomized rules violate the arbitrary and capricious standard.   For these reasons, an agency might decide to forego a randomized regulatory experiment in favor of a non-randomized one.

b. Sunset Rules v. Permanent Rules

As discussed previously, a rule promulgated through an informal rulemaking can automatically expire pursuant to a sunset provision. Otherwise, reversal of the rule requires agency action, which typically will trigger the notice and comment procedures. The former type of rule we've referred to here as a temporary rule and the latter rule as a permanent rule. Although the permanent rule is more costly to reverse, either type of rule is technically reversible, and therefore either type can be used as a vehicle for a regulatory experiment. How should an agency choose between the two?

The decision should turn on what the agency thinks the probability is that the new rule will be "successful"— in other words, that the rule that is the subject of the experiment will be adopted on a permanent basis following the experiment. If the agency thinks that it is sufficiently likely[63] that the new rule will ultimately be adopted on a permanent basis, then it should structure the experiment as a permanent rule.[64] Otherwise, it should structure it as a sunset rule.

Making sure that the structure selected is optimal is important because it can actually lower the costs of experimentation.   These costs include implementation costs, which consist not only of the cost of designing and conducting the experiment but of taking whatever subsequent action is supported by the results of the experiment.[65] The agency minimizes these

---

[62] *See infra* notes 73-106 and accompanying text.

[63] I use this phrase to allow for the fact that it should be up to the agency to decide the correct threshold for determining when a permanent rule is justified. This might be a likelihood of 50%, but it might be something else depending on the agency's budget constraint and other priorities.

[64] To be sure, a permanent rule alone, with nothing more, does not qualify as a regulatory experiment pursuant to the definition used in this report because it does not exhibit the type of ex ante specification characteristic of regulatory experiments. Nevertheless, some permanent rules do take on an experimental cast when accompanied, for example, with a commitment by the agency to conduct and issue a report pertaining to the performance of the rule after some period of time following the rule's adoption. *See infra* notes 219-229 and accompanying text.

[65] *See supra* notes 24-29 and accompanying text.

costs by structuring the experiment in a way that anticipates the outcome of the experiment — through a sunset, if the agency thinks that the experiment will ultimately favor maintaining the status quo — and through a permanent rule if the agency thinks the experiment will ultimately favor adopting the new rule on a permanent basis. It seems likely that in most cases, the probability that the new rule will be successful is less than 50%, otherwise it presumably would have already been adopted. And for this reason, it seems likely that in most cases, regulatory experiments should be structured as sunset rules.

Figure 5: Structuring Regulatory Experiments
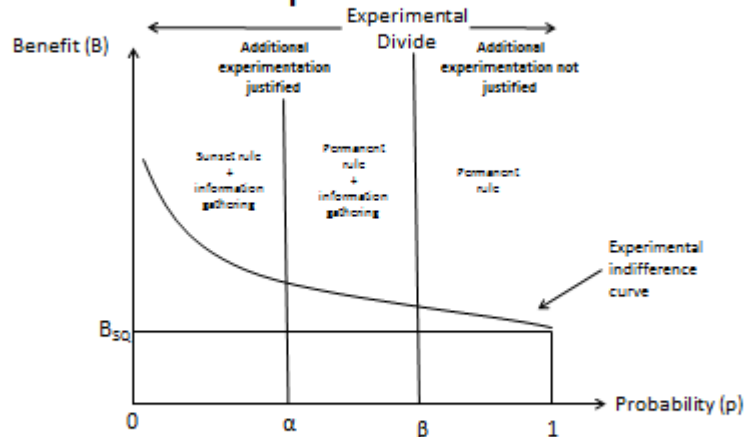
Figure 5 helps capture the considerations that should be taken into account in deciding on the structure of the regulatory experiment. As a preliminary matter, if agencies had unlimited budgets, maybe every rule should be the subject of a regulatory experiment.[66] However, this of course is not the case, and therefore, there is probably some set of rules whose probability of success is sufficiently high that they should not be structured as regulatory experiments at all. In other words, for these relatively non-risky rules, there is no reason to structure them as experiments from the

---

[66] This is one reading of the proposal made by Michael Greenstone. *See* Michael Greenstone, *Toward a Culture of Persistent Regulatory Experimentation and Evaluation*, *in* NEW PERSPECTIVES ON REGULATION 113, 114-15, 118-19 (David Moss & John Cisternino eds., 2009).

outset. If they turn out to be less beneficial than originally thought, then by all means the agency should re-consider the desirability of maintaining them on the books. But the agency should not have to consider the various structural aspects — ex ante specification, control variables and reversibility — associated with regulatory experimentation. In Figure 5, these rules would fall to the right of the line labeled "experimental divide" that is located at probability β on the X-axis. What this means is that every rule to the left of this line (and above the experimental indifference curve) is a candidate for regulatory experimentation.

What about the decision between a sunset rule and a permanent rule? We can see this decision also illustrated in Figure 5. Pursuant to the discussion previously, there is some probability of success, labeled α in Figure 5, such that any rule whose probability of success is greater than α should probably be structured as a permanent rule since it is more likely than some baseline, determined by the agency, that the rule will be successful and therefore adopted on a permanent basis. Although the rules that fall within the interval established by α and β are permanent rules, they are nevertheless experimental and therefore different from the permanent rules that fall to the right of β. The difference of course is that the regulatory experiments structured as permanent rules are designed from the outset with an eye toward the failure of the rule and how to benefit the most from that eventuality.

That leaves everything to the left of α on Figure 5. These are rules where the probability of success is less than α, and these should be structured as a sunset rule. How does an agency decide on the value of α? This will depend on the agency and the type of risk it faces in its regulatory task. Maybe some agencies are faced with particularly risky rules all of the time and therefore a 30% chance of success, for example, might represent fairly good odds. For other agencies, α should probably be much higher, possibly even higher than 50%.

4. Learning Reinforcement

The goal of regulatory experimentation is to generate information about regulatory options. For this reason, it is crucial that the agency design a process through which the correct lessons of the experiment can be drawn. The shape this process takes depends in part on the agency's approach to meeting the ex-ante specification requirement of regulatory experimentation, as discussed above.[67] Under the strict scientific approach, analysis of the data will be fairly straightforward since the hypothesis, as

---

[67] *See supra* notes 2-46 and accompanying text.

well as the statistical criteria for determining whether the hypothesis is to be rejected or not, will be specified in advance. However, there are more options available when the agency takes the looser, exploratory approach.

One can think of the range of alternatives here in terms of the openness or closedness of the process, meaning the degree to which the public is invited in to the actual process of analyzing and evaluating the experimental data. Obviously any informal rulemaking has to be done pursuant to the notice and comment process, which requires, among other things, disclosure of the data that the agency relies on the rulemaking.[68] But to just understand the results of the experiment, the agency has a number of options at its disposal. On the closed end of the spectrum are commissioned studies, in-house studies and expert panels. On the open end of the spectrum are public forums and public calls for research papers. In between these two endpoints might be more constrained third-party commissioned work, including for example, where the agency identifies the questions that it wants answered but then opens up the process to allow anyone to research them.

The examples discussed below reflect the wide variety of approaches taken to this question of learning reinforcement. In the short sale pilot, the SEC adopted a very open process, inviting outside researchers to analyze the data while maintaining flexibility as to the types of studies or even the types of answers it was looking for.[69] In its shelf offering experiment, the SEC took an approach that was slightly more closed, identifying specific questions that it wanted answered but soliciting help from the public at large in answering them.[70] The OCC's approach was quite closed: the agency itself analyzed the data regarding a bank's portfolio risk and did not solicit public help.[71] In the EPA's experiment, which was sidelined by the Supreme Court, it is anybody's guess, but the study was to be completed by the agency itself and did not appear to contemplate a significant amount of public input.[72]

C. Legal Considerations

Regulatory experiments raise a number of legal questions. This part addresses three such issues: the proper judicial review for experimental

---

[68] *See* Administrative Procedure Act, 5 U.S.C. § 553 (2017) (setting out the APA's rulemaking requirements); United States v. N.S. Food Prods. Corp., 568 F.2d 240, 251-53 (2d Cir. 1977) (interpreting those requirements as requiring the agency to disclose all the data upon which it relied and to respond to all "vital questions").

[69] *See infra* notes 120-150 and accompanying text.

[70] *See infra* notes 151-191 and accompanying text.

[71] *See infra* notes 190-219 and accompanying text.

[72] *See infra* notes 219-229 and accompanying text.

rules; whether regulatory experiments trigger OIRA review and whether regulatory experiments trigger notice and comment rulemaking under the APA.

1. Judicial Review of Experimental Rules

How is regulatory experimentation viewed by the courts? As a preliminary matter, one might wonder whether agencies must be able to point to explicit statutory authorization in order to engage in regulatory experimentation. This is unlikely to be the case. After all, a regulatory experiment, for the purposes of this report, is simply a temporary regulatory action taken with an eye toward more permanent action in the future, and agencies are generally thought to enjoy wide discretion in determining the duration of regulatory actions in the absence of congressional direction to the contrary.

The question then about judicial review of experimental regulation is really just one about judicial review of final rules resulting from informal rulemaking. The standard for judicial review in that case is whether the rule is "arbitrary and capricious".[73] Initially, this standard was interpreted as being purely procedural.[74] In other words, the focus was on whether the agency "ha[s] responded to significant points made during the public comment period, ha[s] examined all relevant factors and ha[s] considered significant alternative to the course of action ultimately chosen."[75] In implementing this standard, it is often said that the court is concerned with whether the agency has taken a "hard look" at the relevant question as a procedural matter.[76] However, more modern courts often add a substantive component to their review, ensuring not only that the agency has taken a "hard look" but taking a hard look themselves.[77] In other words, the court wants to ensure that the agency's conclusions are satisfactory and logically follow from the facts that were before it.[78]

With respect to the specific question of judicial review of regulatory experimentation, there are potentially two strands of precedent, which point in different directions, what I'll call the traditional strand and the emergent strand. The traditional strand is associated with *Hüls Am. Inc. v. Browner*.[79] Under that case and its progeny, a court would be expected to apply a

---

[73] *See* 5 U.S.C. 706(2)(a) (2017).

[74] Indeed, the language of the statute does not suggest anything more.

[75] Merrick B. Garland, *Deregulation and Judicial Review*, 98 HARV. L. REV. 505, 527 (1985).

[76] *See id.* at 527.

[77] *See id.* at 545-56.

[78] *See id.* at 545.

[79] 83 F.3d 445, 452 (D.C. Cir. 1996)

deferential standard of review to an agency's decision to conduct a regulatory experiment, provided that the agency does so in a way that exhibits reasoned decision-making.

This conclusion follows from the fact that courts have generally taken the view that "in informal rulemaking, it is "desirable" that agencies "independently amass [and] verify the accuracy of" data."[80] Given this view, the D.C. Circuit in particular has tended to apply "[an] extreme degree of deference" to an agency's evaluation of the extant empirical studies within the agency's technical field.[81] It has also tended to defer to an agency's decision as to whether the available data is sufficient to take some regulatory action[82] or whether additional empirical studies are necessary.[83]

In short, agencies tend to enjoy wide latitude when it comes to evaluating the sufficiency and meaning of available data and whether the agency should undertake additional studies to respond to potential gaps in that data. These cases might then stand for the proposition that an agency should also enjoy considerable deference if, after evaluating the state of the empirical data, it concludes that a reasonable way of generating necessary information is through a regulatory experiment. If this were how a court were to analyze the question of judicial review for regulatory experiments, it would not mean that an agency could simply experiment with reckless abandon, adopting rules willy-nilly on an "experimental" basis, regardless of whether such rules are justified. The agency would still need to engage in reasoned decision-making.[84] In the regulatory experimentation context, that would require the agency, at a minimum, to identify the data that it lacks and hopes to generate; to explain why a regulatory experiment is a reasonable way of generating the desired data; and to specify a time frame by which it hopes to have collected the information necessary to make a

---

[80] Nat'l Ass'n of Regulatory Util. Comm'rs v. FCC, 737 F.2d 1095, 1124 (D.C.Cir.1984).

[81] Hüls Am. Inc. v. Browner, 83 F.3d 445, 452 (D.C. Cir. 1996).

[82] *See, e.g.*, Chamber of Commerce, 412 F.3d at 142 (rejecting the plaintiff's argument that the agency's "[failure] to develop new . . . empirical data" was arbitrary and capricious.).

[83] *See, e.g.*, *See* Center for Biological Diversity v. E.P.A., 749 F.3d 1079, 1087 (D.C. Cir. 2014). Here, the D.C. Circuit refused to find the EPA in violation of the Clear Air Act despite the agency's admitted failure to revise certain air quality standards required by the Act. The court deferred to the agency's determination that "the available information was insufficient to permit a reasoned judgment about whether any proposed standard" satisfied the statutory requirement and held that in such a case, it would be arbitrary and capricious not to gather additional data through a regulatory experiment.

[84] *See* Motor Vehicle Mfr.'s Ass'n v. State Farm Mut. Auto. Ins. Co., 463 U.S. 29, 53 (1983) (explaining that the agency must show that its action was the result of "reasoned decisionmaking").

more permanent decision. Where the agency failed to do this, a court is would be unlikely to defer to the experiment.[85]

While *Browner* and related cases might argue in favor of judicial deference for regulatory experiments, it is important to point out that none of these cases actually involved a regulatory action that would qualify as "experimental" in the sense that term is used in this report. And deferring to the agency's evaluation and assessment of the empirical literature supporting a final rule is quite different from allowing a rule to proceed in the absence of empirical support on the premise that the agency will develop such data over time through a regulatory experiment. For this reason, when drawing inferences from this strand of the case law, it is advisable to do so with some caution. This is all the more true, in light of another, more emergent, strand of the case law. The relevant case here is *Business Roundtable v. SEC*,[86] which seemed to ignore the lessons of *Browner* and applied a less deferential standard to the agency's evaluations of the available store of empirical data.[87]

The case concerned an SEC rule, Rule 14a-11, which was an example of the type of proxy access rule described above.[88] In particular, this rule would have required public companies to include on the corporate ballot board of director candidates that are nominated by certain significant

---

[85] *See, e.g.*, Crowley's Yacht Yard, Inc. v. Pena , 886 F.Supp,. 98 (D.D.C. 1995). In this case, the plaintiff, a company that runs recreational vessels down the Chicago River, challenged the validity of a temporary rule adopted by the Coast Guard limiting when drawbridges along the river could be opened to allow for recreational traffic. The Coast Guard did not help its case by failing to submit to the court the administrative record which would have allowed the court to evaluate the evidence supporting the rule. However, the court said that the rule itself "indicated that it was not supported by substantial evidence." Apparently, the Coast Guard had previously found that the comments and data "at this point are insufficient to provide a basis for a permanent regulatory change," which the court interpreted as undercutting the subsequent regulatory change that was the subject of the case. However, that regulatory change was a temporary one and completely consistent with the Coast Guard's prior finding, if the purpose was to generate relevant data to enable it to assess permanent changes. I suspect that this case would have come out differently if the Coast Guard had explained its action by indicating that it was adopting the temporary rule in an effort to fill the gap in available data so that it could in the future make a permanent regulatory change.

[86] 647 F.3d 1144 (D.C. Cir. 2011).

[87] *See, e.g.*, James D. Cox & Benjamin J.C. Baucom, *The Emperor Has No Clothes: Confronting the D.C. Circuit's Usurpation of SEC Rulemaking Authority*, 90 TEX. L. REV. 1811, 1828 (2012) ("[T]he ultimate effect of the Chamber of Commerce and Business Roundtable decisions appears to be nothing less than establishing a new review standard."); Catherine M. Sharkey, *State Farm "With Teeth": Heightened Judicial Review in the Absence of Executive Oversight*, 89 N.Y.U. L. REV. 1589, 1630 (2016) ("The Business Roundtable court was also revolutionary in its willingness to upend the typical deference to agency cost-benefit analysis.").

[88] *See supra* notes 11-21 and accompanying text.

investors.[89] This rule represented a significant change in how the country's largest companies conducted elections to their governing body. As a matter of law, public companies are required to hold an annual meeting to elect their board members, in advance of which, they are to distribute to shareholders a so-called proxy statement containing a list of nominees to the board, including a significant amount of relevant information, and a ballot allowing the shareholders to vote for the nominees. Importantly, however, historically, companies have only been required to include on this corporate ballot the company's (or in other words, management's) nominees to the board. Thus, if a significant shareholder or group of shareholders wanted a change in board composition because of mismanagement, for example, they would have to distribute to shareholders their own proxy statement at a significant cost. Although such a shareholder would shoulder all of the costs of the election contest himself, he would be required to share the benefits of any management change with all of the other shareholders. And for this reason, there has long been concern that in public companies with a dispersed shareholder base, shareholders do not have sufficient incentives to act as a check on board misfeasance. In response to these longstanding concerns, the SEC adopted its proxy access rule,[90] which sought to eliminate the costs of running a proxy contest by allowing certain investors to simply use the corporate ballot system to propose and promote its nominees to the board.

The rule was immediately challenged in the courts by the *Business Roundtable,* a "big business" interest group, and the D.C. Circuit ended up vacating it.[91] The court's decision was largely predicated on its view that the SEC's conclusions regarding the costs and benefits of the rule were premised on evidence that the court found to be either insufficient or unpersuasive.[92] These conclusions included that the proxy access rule would increase shareholder value, that it would not significantly increase companies' costs, that the rule would not be co-opted by special interest shareholders (like public pension funds) and that it would generate proxy contests that would not otherwise have happened without it.[93] The court took issue with each one of these conclusions.[94] In each case, the court determined that there either was not sufficient evidence to support the

---

[89] *See* Facilitating Shareholder Director Nominations, 75 Fed. Reg. 56,668 (Sept. 16, 2010).

[90] *See, e.g.*, Jill E. Fisch, *The Destructive Ambiguity of Federal Proxy Access*, 61 Emory L.J. 435, 440-47 (2012) (providing a history of the proxy access rule).

[91] Business Roundtable v. SEC, 647 F.3d 1144, 1149-54 (D.C. Cir. 2011).

[92] *See id.*

[93] *See id.*

[94] *See id.*

conclusion or that what evidence did exist was "(at best) mixed."[95]

On the one hand, *Business Roundtable* might simply illustrate the substantive type of "hard look" review, where the court takes an independent look at the agency's conclusions and finds them wanting. However, even under substantive hard look review, courts had tended in the past to defer to the agency's evaluations of empirical data within its expertise.[96] For this reason, some commentators have suggested that *Business Roundtable* represents such a departure from the D.C. Circuit's prior case law that it might signal a new standard,[97] what some have referred to as "State Farm with teeth."[98] If this is true, then it would be important to explain when this new standard applies. Some have suggested that it should and does apply only to independent agencies that do not voluntarily submit their cost-benefit analyses to OIRA.[99]

However, even if *Business Roundtable* announces a new, less deferential standard with respect to an agency's evaluation of relevant empirical data, it does not necessarily follow that this same standard would also apply to the agency's decision to conduct a regulatory experiment. Indeed, the best defense of the *Business Roundtable* court's enhanced scrutiny is that it is information-forcing: it causes the agency to create a more complete administrative record with more complete analysis that is conducted by more economists (and fewer lawyers).[100] This defense has the benefit of viewing the case as consistent with the D.C. Circuit's stated preference that agencies have more, rather than less, information in making regulatory determinations.[101]

But if this is the defense of *Business Roundtable*, then it would make little sense for a court to rely on that case to put a stop to a regulatory experiment. After all, the decision to undertake a regulatory experiment is itself a decision to produce more information, which is the purported reason for *Business Roundtable*-enhanced scrutiny of the administrative record. For these reasons, courts might decide to defer to experimental regulation, provided of course that the agency engages in reasoned decision-making in explaining why it believes the experiment is justified.

If courts were to adopt a deferential posture with respect to experimental regulation, that would be significant because it would suggest that informal agency rules that risk being vacated on appeal might have a much better

---

[95] *See id.* at 1151.
[96] *See supra* notes 79-84 and accompanying text.
[97] *See, e.g.*, Cox & Baucom, *supra* note 87, at 1828; Sharkey, *supra* note 87, at 1630.
[98] *See* Sharkey, *supra* note 87, at 1589.
[99] *See id.*
[100] *See id.* at 1632.
[101] *See supra* notes 77-81 and accompanying text.

shot at surviving judicial review if they are structured as regulatory experiments. To illustrate this possibility, one could consider a counter-factual involving *Business Roundtable* itself. There, the court determined that the agency's conclusions were, for the most part,[102] based on insufficient evidentiary support. What if the SEC had recognized this insufficient evidentiary support and, for that reason, adopted the proxy access rule as a temporary regulatory experiment designed with the purpose of generating much-needed additional data?

As a *practical* matter, if the court's analysis in *Business Roundtable* was really motivated by concerns over information production, then the regulatory experiment would probably have been met with open arms at the D.C. Circuit. However, what is the likely outcome as a *legal* matter? In other words, is there reason to believe that the court would defer to the agency's decision to experiment regardless of the court's view of the information environment surrounding the policy in question? Specifically, what about a case where the court actually believes that the evidence is sufficient for the agency to take a permanent regulatory action but the agency disagrees and instead decides to conduct an experiment? Even then, the underlying logic of *Business Roundtable*, as an information-forcing decision, combined with the longstanding view that more information is a good thing, might weigh in favor of a court deferring to the agency on the decision whether to conduct a regulatory experiment.

As is hopefully evidenced by this discussion, the issue of judicial review of regulatory experiments is a matter that is far from settled. But even if courts end up deferring to the decision whether to experiment, there is another issue lurking here: what about judicial review of the design of the experiment? Could a court defer to the decision to experiment but then apply a higher level of scrutiny to the question of experimental design? It certainly could, but it seems unlikely in light of precedent, which would view the question of experimental design as uniquely within the agency's area of expertise.[103]

Nor does *Business Roundtable* seem to disrupt this conclusion, to the extent that that case is really about incentivizing the agency to create fuller, more information-rich administrative records. Of course, what this means is that even a randomized regulatory experiment would get a deferential standard of review. This might seem odd, since a rule that applies literally to a random set of regulated entities would seem to run afoul of the "arbitrary and capricious" standard that underlies hard look review.

---

[102] There were some other grounds for the court's decision, other than lack of sufficient evidentiary support. *See* Business Roundtable v. SEC, 647 F.3d 1144, 1153-54 (D.C. Cir. 2011).

[103] *See supra* notes 79-85 and accompanying text.

However, that standard applies to the reasons for adopting the rule in the first place, not to the effect that the rule will have on a given population. And in the case of experimental regulation, the reasons for adopting the rule are anything but arbitrary: regulatory experiments generate information that is needed in order to make more permanent regulatory decisions.[104]

To be sure, there are no cases that make this point. However, the SEC's short sale regulation pilot might point to this conclusion. As discussed in greater detail below, in that case, the SEC adopted a randomized experiment, where a rule applied to a random set of public companies.[105] Despite opposition to the experiment by various groups, including most notably the New York Stock Exchange, there was no lawsuit brought challenging the experiment. It is possible, although admittedly speculative, that this was the case because no plaintiff thought they could prevail on the claim given the level of deference the court would likely apply to the experiment, including its design.

2. OIRA Review of Experimental Rules

Under Executive Order 12,866, all executive agencies must conduct a formal cost-benefit analysis of any "significant regulatory action" and submit that analysis to the Office of Information and Regulatory Affairs ("OIRA") for review.[106] If OIRA objects to the agency's analysis, the agency is prohibited from publishing the proposed rule until the agency has consulted with OIRA.[107] In this context, a "significant regulatory action" means any regulatory action that is "likely to result" in a rule that may have an annual effect on the economy of at least $100 million or certain other actions, including those that raise novel legal or policy issues.[108]

Are regulatory experiments likely to fall under OIRA review? Certainly if they have an annual economic impact of $100 million or more. The more difficult question is where the regulatory experiment itself does not meet or surpass this threshold, but the experiment is "likely to result" in a rule that does. In most cases, the "likely result" of a regulatory experiment is the permanent adoption of either the rule that is the subject of the experiment, or a close variation to it, or reversion to the status quo. Thus, in most cases, in making the determination of whether the experiment qualifies for OIRA

---

[104] *See, e.g.*, Jacob Gersen & Adrian Vermeule, *Thin Rationality Review*, 114 MICH. L. REV. 1355 (2016) (observing that it is rational for agencies to act with highly imperfect information in certain circumstances).

[105] *See infra* notes 120-150 and accompanying text.

[106] Exec. Order No. 12,866, 3 C.F.R. 638 (1993), reprinted as amended in 5 U.S.C. § 601 (2017).

[107] *See id.*

[108] *See id.*

Review, the agency will need to consider those two likely outcomes.

3. Notice and Comment Procedures and Regulatory Experimentation

A regulatory experiment adopted through informal rulemaking, regardless of its structure, is almost certainly subject to notice and comment procedures under the APA. Under the statute, notice and comment procedures apply to all informal rulemakings, subject to a few exceptions that the D.C. Circuit has said, interpreting the APA's legislative history, are to be "narrowly construed and reluctantly countenanced."[109] One might be tempted to think that regulatory experiments should nevertheless be treated differently given their temporary nature. However, the D.C. Circuit has held that a rule's limited nature does not alter the analysis, and this likely extends to matters of temporality.[110] As for relevant exceptions, the most likely candidate is the "good cause" exception, which applies "when the agency for good cause finds (and incorporates the finding and a brief statement of reasons therefor in the rules issued) that notice and public procedure thereon are impracticable, unnecessary, or contrary to the public interest."[111]

Does the need to experiment represent a good cause under the exemption? Some commentators have suggested as much.[112] However, it seems unlikely. An agency can claim there is a "good cause" to dispense with pre-effective notice and comment procedures if such procedures are deemed by the agency to be impracticable, unnecessary or contrary to the public interest. It is not exactly clear how "the need to experiment" falls in any of these three categories. This legal conclusion is also supported by the policy concern that opening up the good cause exception to any experimental regulation whatsoever might result in too much experimentation, as agencies scramble to re-characterize rules as experiments for no reason other than to avoid costly and onerous procedures. For these reasons, it is almost certain that a regulatory experiment will trigger notice and comment procedures.

---

[109] State of New Jersey v. Environmental Protection Agency, 626 F.2d 1038, 1045–46 (D.C.Cir.1980).

[110] *See, e.g.,* Tenn. Gas Pipeline Co. v. FERC, 969 F.2d 1141, 1145 (D.C. Cir. 1992) ("'[T]he limited nature of the rule cannot in itself justify a failure to follow notice and comment procedures.'" (quoting Council of the S. Mountains, Inc. v. Donovan, 653 F.2d 573, 582 (D.C. Cir. 1981)))

[111] 5 U.S.C. § 553(b)(3)(B).

[112] *See, e.g.*, Ayres et al., *supra* note 50, at 981 ("Procedurally, an agency might argue that it should not have to go through the notice-and-comment procedure to establish an experiment, because the experiment is merely designed to produce data from which to make a subsequent policy decision.").

But what about the permanent rule adopted in the wake of an experimental period? Recall that all regulatory experiments are, by definition, temporary, even if the information they generate ends up justifying adoption of the experimental regulation on a permanent basis. In many cases, a regulatory experiment will require the agency to take some permanent regulatory action in the wake of the experiment. Specifically, the agency must take a regulatory action where the agency decides to adopt a permanent rule that differs from the rule that was the subject of the experiment (or even where the permanent rule is the same as the experimental one, but the experiment is subject to a sunset provision). The question is whether that action — the permanent regulatory decision that follows a regulatory experiment — might fall within the good cause exception.

If the permanent action is the same or similar to the rule that was the subject of the experiment, then there is a strong argument that full-blown notice and comment procedures are unnecessary since the public will have already had the chance to comment on the same or a similar rule when the agency established the experiment in the first place. This argument would be even stronger if the agency committed to accept and consider post-effective comments on the rule. This approach is sometimes referred to as an "interim final rule": a final rule that dispenses with pre-effective notice and comment, based on the good cause exception, but provides for it in the post-effective period.[113] It is fairly well established that this interim-final rule approach strengthens the case for the good cause exception.[114] And in fact, there is precedent for this in the regulatory experimentation context. Indeed, in its experiment on lending limits, the OCC adopted the permanent action that followed in the wake of its experiment as an interim final rule which lacked notice and comment in the pre-effective period but benefited from it in the post-effective one.[115] It argued that the exception applied because the modification to the experimental rule was slight and because it had already provided opportunity for comment twice before: upon adoption of the experiment and then when it extended the experiment for further review.[116]

An agency might wish to avail itself of the interim final rule methodology in the regulatory experimentation context not simply to avoid

---

[113] *See, e.g.*, Michael Asimow, *Interim Final Rules: Making Haste Slowly*, 51 ADMIN. L. REV. 703 (1999).

[114] *See id*.

[115] *See* Adoption of Interim Rule Regarding Special Lending Limits for Residential Real Estate Loans, Small Business Loans, and Small Farm Loans, 72 Fed. Reg. 31,441 (2007).

[116] *See id*.

onerous notice and comment procedures. After all, under an interim final rule, the agency commits to undergo those procedures. It just commits to do so after, instead of before, the rule becomes effective. Rather, the reason an agency might wish to adopt the interim final rule approach in this context is to minimize the disruptive effects of the experiment. This is especially true if the permanent action that the agency adopts following an experiment is the same as (or is very similar to) the rule that was the subject of the experiment and the experiment is subject to a sunset provision. In that case, allowing the experimental rule to sunset and revert to the status quo, only to then soon after follow up with a permanent rule that is the same as (or approximates) the rule that just expired, would impose unnecessary costs on regulated entities. One solution to this problem would be for the agency to extend the experiment so that it does not expire before a permanent rule is figured out. Many agencies have adopted this solution, including the SEC in the short sale pilot, discussed in greater detail below.[117]

A yet different solution might be for the agency to identify, in the rule that creates the experiment in the first place, what happens in the event that the experiment yields various results. Imagine that the Federal Motor Carrier Safety Administration ("FMCSA") of the U.S. Department of Transportation is trying to decide whether to adopt a rule that would limit the number of hours per day that a commercial trucker may drive. Perhaps the FMSCA decides to cap the number at 10 hours as part of an experiment. It could identify in the rule that creates the experiment that if the rule change does not result in some minimum percentage change in accidents involving truckers, the experiment will proceed to a second phase where the hour cap is 8 hours. This approach is essentially the scientific approach to the ex ante specification requirement, discussed above.[118] As discussed there, agencies might be reluctant to adopt this approach for the reasons discussed there.

D. Public Reaction to Regulatory Experimentation

In thinking about the public reaction to experimentation, those opposed to a given regulatory experiment are likely to have more varied positions than in the non-experimental context. This is because, not only will there be those who oppose the underlying policy and therefore the experiment, but there will also be those who favor the policy but oppose the experiment. From the perspective of this latter group, they might prefer that the policy be adopted on a permanent basis. After all, if it is a good idea, why make it temporary? Those who oppose the experiment (either because they oppose

---

[117] *See infra* notes 120-150 and accompanying text.
[118] *See supra* notes 2-41 and accompanying text.

the underlying policy or because they are in favor of the underlying policy but do not think it needs to be subject to an experiment) come in two varieties depending on the assumptions motivating their position with respect to the underlying policy. There will be those whose position is based on a belief about the underlying policy's social value. Those who oppose the policy (and oppose the experiment) believe that it is value destroying whereas those who favor it (yet oppose the experiment) believe that it is value creating, but they nevertheless have something in common: their respective position with respect to the experiment is rooted in an assumption about the policy's social welfare. Because the regulatory experiment effectively tests the social value of the regulatory action, it is likely that these groups can ultimately be persuaded to support a regulatory experiment.

But there is another variety of position with respect to regulatory experimentation: those whose position (either for or against the underlying policy) is insensitive to concerns about its social value. They oppose or favor the policy because it harms or benefits them. These groups are likely to exhibit greater resistance to an experiment because they are not actually concerned about maximizing social welfare, and therefore an experiment testing the best way to do so is of little concern to them. In fact, it might be a setback particularly if it yields information that cuts against their preferred position.

With this said, there are reasons to question whether even these groups will be successful in their opposition, either as a legal or rhetorical matter. On the legal front, as discussed above, courts may decide to defer to regulatory experiments, in which case legal challenges are likely to be unsuccessful. Of course, there is more to consider than simply legal battles. But even on the rhetorical front, the prospect of prevailing against a proposed regulatory experiment is dim. After all, the purpose of the experiment is to generate data in order to make laws that are more efficient and effective.

For these reasons, it is likely that any opposition to regulatory experimentation will take the form of opposition to the design of the experiment. Groups will argue about costs, risk or benefits. On the cost front, they will argue that the costs of the experiment are too great to justify it. On the risk front, they will argue that the probability of the experiment being successful is much less than the agency thinks, and therefore the experiment is not justifiable. Or they will argue that the probability of success is actually much greater than the agency thinks and therefore there is no need for an experiment. And finally, with respect to benefits, they will argue that the benefits of the rule in the event that is adopted on a permanent basis are not as great as the agency thinks. Or alternatively, that

they are even greater than they think and therefore they should not waste their time with an experiment. These will be issues that agencies will have to work through and address, and they will likely help the agency fine-tune the experiment. To a certain extent, we see this give and take in the example of the SEC's short sale experiment, discussed below.[119] Ultimately, however, consistent with the legal discussion above, it seems possible, if not probable, that courts will defer to the agency even on the issue of experimental design.

II. EXAMPLES

The following are specific examples of regulatory experimentation at three agencies: the SEC (which is an independent agency), the OCC (which is now an independent agency, although, at the time of the regulatory actions discussed below, was an executive agency) and the EPA (which is an executive agency). The examples are useful illustrations of many of the issues discussed up to this point: They illustrate the experimental decision itself — why regulatory experimentation might be valuable in shedding light on certain policy questions that resist easy answers through reason and logic alone. They also provide useful insights about experimental design and in particular the different structures available for conducting a regulatory experiment, as well as the different approaches that agencies might take with respect to the ex ante specification requirement and the process for ensuring learning reinforcement. They also illustrate how the public is likely to react, and how courts are likely to respond, to regulatory experimentation. The examples are organized according to the four different structures for regulatory experimentation, discussed above: temporary rules (both the randomized and non-randomized variety) and permanent rules (both the randomized and non-randomized variety). There is at least one example for each structure, except the permanent, randomized combination, which appears to be rare, if not non-existent.

1. Randomized trial + Sunset Provision

   a. The SEC's Short Sale Experiment[120]

One of the better known regulatory experiments involved the SEC's 2003 re-examination of its longstanding "Uptick Rule." This rule had been adopted in the wake of the Great Depression in an attempt to prevent a

---

[119] *See infra* notes 129-132.
[120] *See* Zachary J. Gubler, *Making Experimental Rules Work*, 67 ADMIN. L. REV. 551 (2015).

financial market phenomenon known as a "bear raid," which many policymakers believed had contributed to the devastating stock market crash of 1929.[121] A bear raid is a cascade of bets that the price of a given stock or the market as a whole will decrease.[122] These bets, formally known as short sales, might be tolerable, or even desirable, in ordinary times.[123] But during periods of financial panic, the fear was that these bets would multiply like a virus, driving down market prices to irrational levels and causing disaster in its wake.[124]

The proposed solution to this parade of horribles was the SEC's Uptick Rule, which prohibited short sales at successively lower prices.[125] Thus, if the market was on the rise, fueled by optimism, the Uptick Rule would allow for unrestricted short selling. But it would put the brakes on short selling in a declining market, thereby staving off the type of bear raids that many feared.

While the Uptick Rule appeared to work well for many decades, beginning in the 1970s, economists began to question its underlying premise. The argument was intuitively compelling: in order to be efficient, markets require input from both boosters and naysayers.[126] Stifling the naysayers, as the Uptick Rule does, only serves to artificially inflate markets, thereby undermining the goal of market efficiency more generally.[127] This argument raised important questions: Did the Uptick Rule actually do what it set out to accomplish, preventing undue downward price pressure during times of market volatility? And more generally, what effect did the Uptick Rule actually have on market efficiency? In order to answer these questions, the SEC needed data, which it could only generate through a regulatory experiment.

It took the SEC many years to realize this goal, in part due to opposition from the stock exchanges, with the New York Stock Exchange its primary detractor.[128] However, the SEC finally adopted its experiment in 2004.[129]

---

[121] *See, e.g.*, Henry T.C. Hu, *Too Complex to Depict? Innovation, "Pure Information," and the SEC Disclosure Paradigm*, 90 TEX. L. REV. 1601, 1689-90 (2012).

[122] *See, e.g.*, Jonathan R. Macey et al., *Restrictions on Short Sales: An Analysis of the Uptick Rule and Its Role in View of the October 1987 Stock Market Crash*, 74 CORNELL L. REV. 799, 802 (1989).

[123] *See* Hu, *supra* note 121, at 1612-13 (discussing early criticisms of the Uptick Rule).

[124] *See* 7 LOUIS LOSS & JOEL SELIGMAN, SECURITIES REGULATION 115 (4ᵀᴴ ED. 2012) (QUOTING S. REP. NO. 73-1455, AT 50 (1934).

[125] *See* SHO Pilot Proposal, 68 Fed. Reg. 62,972, 62,972 (Nov. 6, 2003).

[126] *See, e.g.*, Edward M. Miller, *Risk, Uncertainty, and Divergence of Opinion*, 32 J. Fin. 1151, 1166 (1977) (advancing the hypothesis that "[i]n a market with little or no short selling the demand for a particular security will come from the minority who hold the most optimistic expectations about it.")

[127] *See id.*

[128] *See* Gubler, *supra* note 35, at 564.

They designed it as a sunset rule with randomization.[130] In particular, they chose to eliminate the Uptick Rule with respect to a randomly selected group of companies consisting of roughly one-third of the Russell 3000 index, a well-known market index.[131] With respect to the remaining companies in the index, the control group, there would be no change in the rule's application.[132]

Interestingly, the SEC did not identify upfront what actions it would take given the results of the experiment. For example, it did not commit to eliminate or maintain the Uptick Rule if the experiment were to demonstrate some pre-specified effect on measures of market efficiency. In fact, the SEC did not even announce what hypothesis the experiment was trying to test. Was it trying to test whether the Uptick Rule affected market efficiency? Or was it trying to test something else, for example, the rule's effectiveness in staving off bear raids? And finally, the SEC failed to identify the criteria it would use to assess whether the experiment was a "success" or not. In other words, with respect to the ex-ante specification requirement of regulatory experimentation, the SEC adopted the exploratory approach discussed above.[133]

The experiment was set to automatically sunset after one year, at which point application of the Uptick Rule would revert to the status quo ante.[134] This time period represented a concession to the stock exchanges, who objected to the SEC's original two-year time frame.[135] However, the SEC also reserved the right "[to] extend the period of, or modify the Pilot as it determine[d] necessary or appropriate in the public interest or for the protection of investors."[136] In fact, the SEC ended up exercising this right,

---

[129] *See* Pilot Adopting Release, 69 Fed. Reg. 48,008 (Aug. 6, 2004). It is worth noting that the SEC actually proposed the rule as a temporary rule before adopting it. This allowed for comments on whether there should be an experiment in the first place, which some commenters objected to. *See* Gubler, *supra* note 35, at 564-66. This should be contrasted with the SEC's shelf registration offering, where the SEC proposed the rule but never as an experiment, and then subsequently adopted it on an experimental basis. *See infra* notes 156-157 and accompanying text.

[130] *See* Pilot Adopting Release, 69 Fed. Reg. at 48,008.

[131] *See id.* at 48,008.

[132] This experimental design reflected changes the SEC made after receiving comments about its originally proposed design. *See* Gubler, *supra* note 35, at 578-59 (discussing how the SEC's originally proposed design would have relied on a different stock index, the Russell 1000, which consists of relatively larger firms and speculating that the SEC ultimately opted to use the Russell 3000 index instead in order to avoid biases having to do with a sample that does not represent the broader population of firms).

[133] *See supra* notes 2-45 and accompanying text.

[134] *See* Pilot Adopting Release, 69 Fed. Reg. at 48,008.

[135] *See* Gubler, *supra* note 35, at 569.

[136] *Id.* at 48,033.

just days before the expiration of the experiment, by extending it one more year.[137]   The reason given had to do with minimizing the costs of the experiment, which consisted largely of systems changes that the broker-dealer community had to make in order to facilitate the experiment.[138]   The SEC explained in its order extending the experiment another year that it did not want to revert to the status quo ante if they ended up, in light of the data, deciding to subsequently eliminate the Uptick Rule on a permanent basis, since that would have required broker-dealers to undertake the same systems changes yet again.[139] Thus, although the SEC only relied on the data generated during the first year of the experiment in its analysis, it nevertheless extended the experiment pending its final regulatory decision.

In analyzing the data generated by its experiment, the SEC relied both on internal and external experts.   Internally, it relied on a group of economists housed in what is now called the Division of Economic and Risk Analysis.[140]   This group analyzed the data and published a summary report. The external experts were economists who published three different studies analyzing the data.[141]   The SEC made this external review possible by establishing a process whereby the stock exchanges released all of the data generated by the experiment.[142]   Finally, the SEC convened a public roundtable where the three external studies were each presented and then commented upon by an academic economist.[143]   The roundtable also included two former chief economists at the SEC and a former chief economist of the New York Stock Exchange.[144]

As a general matter, all of the panelists were in favor of eliminating the Uptick Rule. The external studies found no evidence that the Uptick Rule was bad for market efficiency, although surprisingly there was some evidence that it actually improved liquidity in some cases.[145]   Interestingly,

---

[137] *See* Order Extending Term of Short Sale Pilot, 71 Fed. Reg. 24,765 (Apr. 26, 2006).

[138] *See id.*

[139] *See id.* at 24,766. ("Market participants made significant changes in their systems and practices to comply with the Pilot. Absent an extension of the Pilot's end date of April 28, 2006, the pre-Pilot short sale price tests would be restored, and market participants would be required to make changes to their systems and practices to ensure that they comply with these rules. If the Commission thereafter adopts rules that remove or change the nature of price tests for some or all securities, market participants would be required to change their systems and procedures again, which could result in substantial additional costs. Extending the Pilot ending date would keep the costs of changes to a minimum and help avoid market disruption.")

[140] *See* Gubler, *supra* note 35, at 570-71.

[141] *See* Price Test Removal Proposal, 71 Fed. Reg. 75,068, 75,069 & n.12.

[142] *See id.* at 75,069 n.11.

[143] *See id.* at 75,074.

[144] *See id.*

[145] *See* Gubler, *supra* note 35, at 573-74.

the experiment did not actually allow the researchers to assess how well the Uptick Rule prevented bear raids, since the experiment did not take place during a period of extreme market volatility, which is a typical pre-condition for a "bear raid."[146] And therefore the experiment did not allow for an assessment of the very reason the Uptick Rule was adopted in the first place.[147] Nevertheless, in response to the analyses, and the unanimous opinion of the roundtable participants, the SEC subsequently proposed[148] and adopted a rule eliminating the Uptick Rule.[149] Interestingly, during the notice-and-comment process accompanying this rulemaking, the New York Stock Exchange, which had been the SEC's staunchest objector to elimination of the Uptick Rule, and which had even opposed the experiment itself, ended up supporting the Uptick Rule's elimination.[150]

2. Non-Randomized + Sunset Provision

a. The SEC's Shelf Registration Rule

While the SEC's short sale pilot program was structured as a randomized experiment subject to a sunset provision, its shelf registration rule was structured as a non-randomized experiment subject to a sunset. The shelf registration rule had to do with the way in which companies sell securities, like stocks and bonds, to the public. This process involves a company (called the issuer) who hires an investment bank (called the underwriter), which relies on its expertise at asset valuation, its investor contacts and its sales force, to bring the securities to market.[151] The process is highly regulated, requiring the issuer to file with the SEC a document called the registration statement, which discloses various information about the securities, the issuer, the nature of the business, possible risks and the company's financial data.[152] With a few exceptions, the SEC historically required securities offerings to be made immediately after its registration statement became "effective." [153] This policy had the effect of penalizing

---

[146] *See id.*

[147] *See id.*

[148] *See* Proposed Elimination of Uptick Rule, 71 Fed. Reg. 75,068 (Dec. 13, 2006).

[149] *See* Adoption of Elimination of Uptick Rule, 72 Fed. Reg. 36,348 (July 3, 2007).

[150] *See* Gubler, *supra* note 35, at 575.

[151] *See generally* 1 Louis Loss & Joel Seligman, Securities Regulation, 321-85 (3rd ed. 1989).

[152] *See id.*

[153] *See* Shelf Registration, 48 Fed. Reg. 52,889 (Nov. 23, 1983) ("These [exceptions] included securities to be issued in continuing acquisition programs or those underlying exercisable options, warrants or rights. Administrative practice, however, accommodated traditional shelf offerings beyond those specified in the Guide. Shelf registration was

companies that wanted to wait for more favorable market conditions before they actually began the sale process.[154] This preference for a delayed offering is all the more understandable in light of the fact that the entire registration process — from hiring the underwriter to having an effective registration statement — could take anywhere from several months to over a year, during which time the market price of the securities in question could have changed significantly.

In light of these concerns, the SEC engaged in a regulatory experiment. They adopted on a temporary basis Rule 415, which would allow for delayed offerings, also called "shelf offerings," in recognition that the securities could be put on the shelf in the present to be taken off and sold in the future.[155] The SEC structured the experiment as a rule that would sunset after nine months.[156] Unlike in the case of the short sale pilot, the SEC never proposed the shelf registration experiment as the subject of notice-and-comment. It simply adopted the experiment.[157] With that said, the SEC had already proposed the rule twice as a permanent rule.[158] After receiving a number of comments the first time around, the SEC revised the rule in light of the comments, and proposed a new version a second time.[159] This also produced a number of comments. It was only at this point that the SEC then adopted a further revised version on an experimental basis.[160] Shortly thereafter, the SEC announced that it would convene public hearings prior to the expiration of the experimental period. [161] The purpose of these hearings was to receive public input on a number of questions that the SEC identified in the release and that it had no doubt accumulated

---

permitted for such diverse offerings as limited partnership tax shelters, employee benefit plans, pools of mortgage backed pass through certificates offered from time to time, and customer purchase plans.").

[154] *See id.*

[155] *See* Adopting Release for Rule 415 Temporary Rule, 47 Fed. Reg. 11,380 (March 16, 1982).

[156] *See id.*

[157] *See id.*

[158] *See* Proposed Revisions of Regulation S-K, 46 Fed. Reg. 78 (Jan. 2, 1981); Integrated Disclosure Proposal, 46 Fed. Reg. 42,042 (Aug. 18, 1981).

[159] *See id.*

[160] *See id.*

[161] *See* Examination of the Registration of Securities To Be Offered and Sold on a Delayed or Continuous Basis in the Future, 47 Fed. Reg. 11,701 (March 18, 1982) ("The procedural steps include public hearings to explore the impact of a shelf registration rule and to give all interested parties further opportunity to present their views to the Commission. During the period prior to the commencement of the hearings, interested investors, issuers and members of the securities industry may wish to form groups or task forces to organize and prepare their presentations. Following the hearings, the Commission may publish further rulemaking proposals for notice and comment.").

through its experience of proposing the rule twice. [162] The SEC emphasized in this release that it preferred for the commenters to provide empirical evidence that would shed light on the various questions the SEC had identified. [163] These questions included whether the proposed shelf registration rule promoted the purposes of the securities laws, including investor protection and fairness and efficiency more generally.[164] The SEC also wanted to know how commentators thought the rule would affect markets, including the likely effect it would have on capital costs as well as how it would affect the investment banking industry.[165]

Following these hearings, the SEC decided to extend the experiment another year, until December 31, 1983.[166] It gave two reasons for the extension, both of which came from several commenters. First, it repeated the opinion of "many" commenters that the initial nine-month period was simply too short to gather the information needed to properly assess the shelf-registration rule.[167] Second, it expressed the view that the unusual market conditions weighed in favor of extending the experiment for fear of creating an unrepresentative sample.[168] Here, the SEC was no doubt referring to the recession that hit the United States in 1982, the result of sky-high oil prices, increasing interest rates and greater-than-10% unemployment.

At the same time that it announced an extension of the experiment, the SEC also disclosed data that it had compiled during the experimental period.[169] The data concerned how the market was responding to the rule and, in particular, for what types of offerings issuers were using the rule.[170] This was relevant to one of the central concerns that had been targeted at Rule 415 from the very beginning. That concern was that shelf offerings — particularly for primary offerings where the issuer, as opposed to shareholders, is the one selling the securities — would increase the dominance of large Wall Street banks and institutional investors at the expense of smaller, regional underwriters and retail investors.[171] The concern was that under the rule, an issuer's ability to sell large blocks of securities on extremely short notice would ensure that only the largest, most well-capitalized investment banks and institutional investors would be able

---

[162] *See id.*

[163] *See id.*

[164] *See id.*

[165] *See id.*

[166] *See* Extension of Temporary Rule, 47 Fed. Reg. 39,799 (Sept. 10, 1982).

[167] *See id.*

[168] *See id.*

[169] *See id.*

[170] *See id.*

[171] *See id.*

to be involved.[172]   And consequently, smaller, regional-banks and retail investors would lose out. Admittedly, this was more of a concern for equity than debt offerings, since debt offerings rarely involved smaller regional underwriters or retail investors. The data the SEC shared in the release extending the experiment showed that relatively few shelf offerings even concerned the types of primary equity offerings that commenters were concerned about.[173]   Of course, this response was not viewed as fully satisfactory by everyone, especially considering that the market response to the rule up to that point was unusual in light of the persistent recession.[174] In fact, in an unusual move that highlighted the seriousness of these concerns, one SEC commissioner, Barbara Thomas, dissented from the extension order altogether, apparently preferring that the SEC simply start over.[175]

But the SEC did not start over. And a little over one year later, and twenty months after the experiment had begun and, interestingly, one month before it was to expire, the SEC adopted a permanent version of Rule 415.[176] The word "version" is intentional here because the rule the SEC ultimately adopted was not the same one that was the subject of the experiment. Rather, in an apparent concession to those concerned about the rule's effect on the structure of the capital markets, the SEC decided to limit primary shelf offerings to large public companies that already disclose information on a periodic basis.[177] The reasoning was that these companies already sell their securities in reliance on large Wall Street banks and institutional investors, and therefore the rule is unlikely to have the feared effects on capital market structure. It is difficult to know exactly what commenters thought of this concession, however, because interestingly, the SEC did not ever propose this version of the rule. Rather, they simply adopted it as a final rule.[178] This is in contrast to what the SEC did in the short sale experiment, where it launched another notice-and-comment rulemaking following the experiment.[179] We do however know how one previously opposed voice felt about the new rule. Commissioner Thomas dissented again at least with respect to the portion of the rule permitting primary equity shelf offerings by large public companies.[180] Her position was that the rule should not apply at all to primary equity offerings.[181]

---

[172] *See id.*

[173] *See id.*

[174] *See id.*

[175] *See id.*

[176] *See* Adoption of Rule 415, 48 Fed. Reg. 52,889 (Nov. 23, 1983).

[177] *See id.*

[178] *See id.*

[179] *See supra* notes 148-149 and accompanying text.

[180] *See* Adoption of Rule 415, 48 Fed. Reg. at 52,889.

[181] *See id.*

The experimental data was clearly important to the SEC's ultimate conclusions. We've already seen how the SEC relied on data about the market's response to the rule during the first half of the experiment to address commenter's concerns about primary shelf offerings. In the release of the final rule, the SEC also relied heavily on three different event studies, which measured the effect that the rule change had on some variable of interest.[182] In this case, the researchers in all three cases were interested in testing whether Rule 415 lowered issuers' costs of raising capital.[183] The hypothesis was that shelf offerings increased the information available to potential underwriters concerning the issuer and its planned offering. This increased information, the authors reasoned, placed potential underwriters on a more equal informational footing, increasing competitive among underwriters keen on acquiring the issuer's business.[184] The result, the authors hypothesized, might be a lower cost of capital when it came time for the issuer to engage its underwriter.[185] In fact, the results of the studies were generally consistent with this hypothesis, suggesting that Rule 415 did in fact lower issuers' cost of capital.[186]

Of course, a lower cost of capital for issuers meant less revenue for investment banks that are hired to actually raise the capital. In light of these findings, one might assume that the underwriting community was not particularly excited about Rule 415. This lack of enthusiasm was undoubtedly compounded by the fact that, in addition to lower underwriter fees, the rule also increased underwriters' legal risk. The reason is because it created a market where an issuer could hire an underwriter in the morning to help it sell securities from the issuer's shelf that afternoon, with little time for the underwriter to complete its due diligence on the offering.[187] With the federal securities laws creating liability for underwriters who failed to satisfy their "due diligence defense," this meant a material increase in liability for underwriters.[188] Although the SEC did not address in the adopting release the effect that the rule would have on underwriter fees, it did address the concern that it would increase an underwriter's liability risk.[189] The SEC response was that the market had, and would continue, to respond to the changes, including the trend of an issuer hiring a single underwriter for all of its securities work and periodic due diligence sessions

---

[182] *See id.*

[183] *See id.*

[184] *See id.*

[185] *See id.*

[186] *See id.*

[187] *See id.*

[188] *See* JAMES D. COX ET AL., SECURITIES REGULATION: CASES AND MATERIALS 489-91 (7th ed. 2013).

[189] *See* Adoption of Rule 415, 48 Fed. Reg. at 52,889.

so that the issuer's underwriter would, at any time, be up to speed on its due diligence on the issuer.[190]

In the end, then, Rule 415 represented an unusual procedural approach, whereby the SEC neither subjected the experimental rule to notice and comment nor the permanent rule adopted in the wake of the experiment. However, it is not clear that such an approach would withstand judicial scrutiny in the present day.

b. The OCC's Lending Limits Experiment

Another example of a non-randomized regulatory experiment, subject to a sunset rule, comes from the Office of the Comptroller of the Currency (the "OCC"). The OCC is the primary regulator of all banks with a national charter. Its reach, however, does not extend to state chartered banks. The OCC was an executive agency until 2010, when it was made an independent agency.[191] Thus, prior to 2010, it was subject to, among other things, OIRA review pursuant to Executive Order 12,866.

Although the OCC's regulatory reach extends to all of the large financial institutions with household names, it also includes many small community banks. Not surprisingly, the difference in size between the community banks and the large Wall Street banks calls for differences in regulation. For this reason, in 1999, the OCC issued an advanced notice of proposed rulemaking in an effort to gather data on how the OCC's regulation could be better tailored to the needs of community banks.[192] This was basically a public brainstorming session to determine what, if any, regulatory action the OCC should take to help lessen the regulatory burden on community banks vis-à-vis their much larger competitors.[193] The reply that the OCC received in return was unequivocal: relax the lending limits on community banks.[194]

Lending limits are an important regulatory tool that the OCC uses to ensure prudent and sound management of institutions in its regulatory purview. In particular, federal statute, and OCC regulations, limits the amount of loans and extensions of credit a national bank can make to any one borrower to 15 percent of a bank's capital.[195] A bank may lend an additional 10 percent if the credit is secured by readily marketable

---

[190] *See id.*

[191] *See* Dodd-Frank Wall Street Reform and Consumer Protection Act, Pub. L. No. 111-203, § 315, 124 Stat. 1376, 1524 (2010).

[192] *See* Community Bank-Focused Regulation Review, 64 Fed. Reg. 25,469 (1999).

[193] *See id.*

[194] Lending Limits Pilot Program, 66 Fed. Reg. 31,114, 31,115 (2001).

[195] *See* Community Bank-Focused Regulation Review, 64 Fed. Reg. at 25,469.

collateral.[196] So, for example, if a bank has $100 million worth of common stock outstanding, under the lending limit rules, that bank can lend up to $15 million to a single borrower, an amount that increases to $25 million if the borrower can put up "readily marketable" collateral, meaning something like treasury bills or other assets that can be sold easily.[197] Because community banks are so much smaller than their larger, multinational competitors, these lending limits have a disparate effect on the two groups of institutions, potentially significantly constraining the competitiveness of the community banks.[198] While relaxed lending limits might be good for a bank's bottom line, they also likely lead to increased risk in the bank's loan portfolios. After all, the purpose of the limits in the first place is to limit default risk.

Given these concerns, the OCC decided to run a regulatory experiment.[199] The proposed rule, which was the subject of notice and comment, consisted of special, higher lending limits for the types of loans that community banks often make — in this case, loans secured by 1-4 family residential real estate and loans to small businesses.[200] The OCC restricted these special lending limits to only "eligible banks" that received approval through an application process and to banks with main offices located in states where a lending limit higher than then OCC's then-prevailing lending limits applied.[201] There were also additional safeguards imposed, including an individual borrower and aggregate borrower cap expressed as percentages of the bank's capital. The proposal stated that the OCC had the power to end the experiment early if its monitoring of the participating banks suggested that there were safety and soundness issues.[202] Otherwise, it expected to "review national banks' experience with the new exceptions over the three-year pilot period and determine whether to retain, modify, or rescind the exceptions."[203] Because the OCC was at the time an executive agency, it was potentially subject to OIRA review. However, the OCC determined that the proposed experiment was not a "significant regulatory action," the quantitative standard for determining OIRA review.[204] And OIRA did not assert review under the non-quantitative standards set forth in Executive Order 12,866.

The comments that the OCC received mainly came from banks and

---

[196] *See id.*

[197] *See id.*

[198] *See id.*

[199] *See* Proposed Lending Limits Pilot Program, 65 Fed. Reg. 57,292 (Sept. 22, 2000).

[200] *See id.*

[201] *See id.*

[202] *See id.*

[203] *See id.* at 57,294.

[204] *See id.* at 57,295.

were generally supportive of the experiment, with one exception.[205] That exception questioned the wisdom of the experiment, although not the proposed rule, on the ground that that the cost of the application process would deter participation in the experiment, thereby resulting in too little data to be meaningful.[206] Nevertheless, the OCC decided to go through with the experiment on substantially the terms proposed.[207] In other words, the OCC adopted a final rule authorizing an experiment that would expire automatically at the end of three years.[208]

A total of 169 banks headquartered in 23 states received approval to rely on the special lending limits.[209] At the end of the three-year experimental period, the OCC compared banks in the experimental group with those in the control group along a number of metrics measuring safety and soundness and found no statistically significant difference.[210] Nevertheless, deeming that the experiment had not lasted long enough, and had not generated enough data, to draw any definitive conclusions, the OCC subsequently extended the experiment for another three years.[211] They did so by submitting the proposed extension to another notice-and-comment rulemaking, which they undertook roughly five months prior to the expiration of the experiment.[212] The final rule extending the experiment was adopted roughly one month prior to the experiment's expiration.[213] The OCC also took the opportunity in that final rule to expand the special lending limits under the experiment to apply to certain farm loans.[214]

By the time the experiment had run its course, the number of banks subject to the experiment had increased to 288, approximately 15% of all

---

[205]*See* Adoption of Lending Limits Pilot Program, 66 Fed. Reg. 31,114 (June 11, 2001).

[206] *See id.* at 31,115.

[207] *See id.* The most important change was one made to ensure that the special lending limits did not give a national bank eligible for the special limits a competitive advantage in those states where the state lending limit was higher than 15% (the then applicable federal limit) but less than 25% (the limit under the proposed rule).

[208] *See id.*

[209] *See* Proposed Extension of Lending Limits Pilot Program, 69 Fed. Reg. 21,978 (2004).

[210] *See id.* at 21,980.

[211] *See id.*

[212] *See id.* To be more specific, the experiment itself was set to expire on June 11, 2007, although the original rule allowed for the banks that were subject to the experiment to continue to lend under the experimental rules until September 10, 2007, which is the date used to calculate the periods of time remaining under the experiment mentioned in the text above. *See id.*

[213] *See* Adoption of Extension of Lending Limits Pilot Program, 69 Fed Reg. 51,355, 51,356 (Aug. 19, 2004).

[214] *See id.*

national community banks,[215] suggesting that the concerns about lack of participation in the experiment were unfounded. Upon completion of the experiment, the OCC conducted a similar statistical analysis to the one before.[216] Once again, it found no statistically significant difference between the two groups with respect to the observed metrics. Consequently, the OCC adopted the proposed rule, albeit with one fairly minor modification.[217]

Importantly, however, the OCC adopted this final rule as an interim final rule, meaning that it was able to avoid pre-effective notice and comment.[218] It did so under the APA's "good cause" exception for notice and comment procedures. It reasoned that the exception applied because the modification to the experimental rule was slight and because it had already provided opportunity for comment twice before — upon adoption of the experiment and the extension.[219]

3. Non-Randomized + Permanent Rule

The examples thus far have all involved regulatory experiments that automatically expire pursuant to a sunset provision. There may be situations where a regulatory experiment is structured as a permanent rule but the agency commits from the outset to study the effects of the rule in anticipation of modifying the rule down the road. In other words, although the action is nominally permanent, the agency is nevertheless preparing from the outset for failure of a certain type. These also qualify as regulatory experiments under the definition used here.

a. The EPA's Greenhouse Gas Experiment

One particularly prominent example of this type of regulatory experimentation structure has to do with the EPA's so-called tailoring rule for greenhouse gas emissions. [220]

---

[215] *See* Adoption of Interim Rule Regarding Special Lending Limits for Residential Real Estate Loans, Small Business Loans, and Small Farm Loans, 72 Fed. Reg. 31,441 (2007) [hereinafter "Lending Limits Interim Rule"].

[216] *See id.*

[217] Lending Limits Interim Rule, 72 Fed. Reg. at 31,441. The interim rule removed the $ 10 million individual borrower cap. The OCC said that "[i]n view of the other limits and safeguards in the interim rule, and the OCC's experience with the pilot program, [it did] not believe this restriction is necessary." *Id*. at 31,443.

[218] *See id.*

[219] *See id.* .

[220] *See* Prevention of Significant Deterioration and Title V Greenhouse Gas Tailoring Rule, 41 Fed. Reg. 55,292 (2009).

These regulations arose out of the Supreme Court's holding that greenhouse gases fall within the Clean Air Act's definition of "air pollutant."[221]   While this holding paved the way for the EPA to regulate greenhouse gas emissions by both mobile sources (like cars and trucks) and stationary sources (like large industrial factories), the agency ran into problems with respect to the stationary sources. The reason is because the statute's regulatory requirements with respect to stationary sources of pollutants are triggered by specific numerical thresholds (either 100 or 250 tons per year depending on the circumstances), thresholds meant to limit the statute's applicability to significant industrial sources of pollution. Yet, if the EPA were to apply those same statutory thresholds to greenhouse gases, the result would expand the statute's reach to even the most insignificant local factories.

The EPA determined that the administrative burden of that result would be too onerous.[222]   Therefore, it decided to propose what it called a "tailoring rule."[223] In other words, it increased the statutory threshold, *with respect to greenhouse gases alone*, from 250 tons per year to 100,000 tons per year. It thereby reduced or "tailored" the statute's applicability to only large producers of greenhouse gases, which made greenhouse gas regulation of stationary sources more administratively manageable and was arguably more consistent with the legislative intent. However, the EPA recognized that these new thresholds were imprecise, with uncertain results, and tended to view them "as interim measures that will need to be reassessed in terms of their administrative necessity."[224]   To that end, the EPA committed to conduct and complete a study within five years from the effective date of the final rule that would be used to make necessary modifications in light of administrative resources.[225]

The EPA adopted the final rule in 2010.[226] The final rule included the EPA's planned study, complete with an "enforceable commitment" to undertake that study within a specified period of time.[227]   The EPA

---

[221] *See* Massachusetts v. EPA, 549 U.S. 497 (2007). In fact, the Court held that this was the case only if the EPA were to find that such gases satisfied the statutory standard—in other words, that they "contribute to . . . air pollution which may reasonable anticipated to endanger public health or welfare." *See id*. The EPA made this finding, and therefore greenhouse gases came to be included in the category of "any pollutant" under the Clean Air Act. *See* 75 Fed. Reg. 17,004 (April 2, 2010).

[222] *See* Proposed Tailoring Rule, 74 Fed. Reg. 55,292, 55,294 (Oct. 27, 2009).

[223] *See id*.

[224] *See id*. 55,336.

[225] *See id*.

[226] *See* Adoption of Tailoring Rule, 75 Fed. Reg. 31,514 (June 3, 2010).

[227] The Supreme Court later invalided the tailoring rule on the ground that it was not permitted by the statute. *See* Utility Air Regulatory Group v. Environmental Protection Agency, 134 S. Ct. 2427 (2014).

explained that they did not know exactly what form final action would take, which depended on the conclusions of the study.[228]

It is not clear why the EPA adopted the tailoring rule as a permanent as opposed to a temporary rule. It is, after all, somewhat puzzling in light of the agency's recognition of the imprecision of the whole exercise. One can only speculate as to the reason. It is possible that it wanted to send a message to industry that even though the rule might change on the margins, there would be no change regarding the EPA's decision to bring greenhouse gases within the scope of the statute. It is also possible that it simply did not see any need for an actual experiment but instead anticipated to fine-tune the rules over time. But while we do not know for sure, what does seem more certain is that a regulatory experiment would not have avoided the Supreme Court's ultimate decision to invalidate the rule on the ground that it was not supported by the statute. In this sense, the case of the EPA's tailoring rule is different from the SEC's experience with the proxy access rule. Whereas the SEC's proxy access rule might have survived judicial review if it had been structured as a regulatory experiment,[229] this is not the case with the EPA's tailoring rule. The difference is attributable to the fact that the D.C. Circuit's objection to the SEC's rule in *Business Roundtable v. SEC* had nothing to do with the Commission's statutory authority to adopt the proxy access rule but rather centered on its analysis of the evidence underlying its decision. In *Utility Air Regulatory Group v. EPA*, by contrast, the Court simply did not think the EPA had statutory theory to act, and no cost-benefit analysis, no matter how thorough and precise, would save the day.

III. RECOMMENDATIONS

Now that we have considered the theoretical, legal and practical considerations relevant to regulatory experimentation, and having seen examples of regulatory experimentation in its various guises, we are in a position to distill and synthesize this information into the following

---

[228] *See* Adoption of Tailoring Rule, 75 Fed. Reg. at 31,525 ("We cannot predict at this time what form that final action will take. It could function as a Step 4, bringing in additional sources based on, for example, streamlining actions, increased permitting authority resources, and experienced and more efficient permitting staff; and it could further indicate that we intend to follow-up with a Step 5 to bring in more sources. Alternatively, it could also function as a final step excluding certain sources permanently based on our application of the Chevron framework, taking account of the "absurd results" doctrine, and subjecting the remaining sources to permitting. However, whatever final action we take would explain any necessary changes to the Step 3 thresholds and would supersede the 6-year exclusion for sources and modifications below 50,000 tpy CO_2 e.").

[229] *See* Gubler, *supra* note 8, at 170-73.

recommendations.

For the purposes of these recommendations, it is important to underscore the broad definition of regulatory experimentation that has been adopted in this report. Under that definition, a "regulatory experiment" is any regulatory action designed with the express purpose, from the outset, of generating information that would reasonably be expected to inform a more permanent decision down the road.

1. The Decision Whether to Run a Regulatory Experiment

   i.   Agencies should consider adopting a rule on an experimental basis even if there is a low probability that the rule in question will survive beyond the experimental period. The value of a regulatory experiment lies in its temporary nature and the information it generates about the long-term desirability of the rule in question. If the experimental results end up suggesting that the rule is undesirable for some reason, the experiment can be terminated. In other words, an experimental approach to rulemaking is one where failure is anticipated and planned for in advance. Additionally, even if the rule does not survive the experiment, the process results in valuable information the agency might not have been able to gather otherwise.

   ii.   In deciding whether to engage in a regulatory experiment, agencies should focus resources on rules with significant potential benefits relative to the status quo. It is these types of high-risk, high-reward rules where regulatory experimentation is likely to generate the greatest bang for the regulatory buck. To re-purpose a famous investing adage, it probably makes sense to run a regulatory experiment if one can say of the rule in question, "Heads we win big; tails, we might lose, even significantly, but not catastrophically."

   iii. In deciding whether to engage in a regulatory experiment, agencies would do well to consider a break-even analysis, particularly if the agency finds that it is more difficult to estimate the net benefits of the rule in question than the costs of experimentation. A break-even analysis allows the agency to estimate, given the costs of the experiment, the magnitude of the net benefits the rule must generate in order to justify the experiment. In conducting a break-even analysis for regulatory experimentation, agencies must estimate the costs of the experiment; the probability that the rule fails under the experiment (that is to say, that the net benefits of the rule do not end up justifying adoption of the rule on a permanent basis); and the net benefits of the rule in the event the experiment fails. Armed with that information, the

agency can then calculate the net benefits the rule would need to generate in the best case scenario in order to justify the experiment.

2. Structuring Regulatory Experiments

i.   In structuring regulatory experiments, agencies should consider two key design choices. First, they should consider the choice between, on the one hand, a randomized trial — where a rule is applied to a random set of regulated entities and then compared along various metrics with a control group — and, on the other hand, a non-randomized trial where the rule in question applies to all relevant regulated entities. Second, agencies should consider the choice between a temporary rule that automatically expires pursuant to a sunset provision and a permanent rule that does not. While agencies should strongly consider randomized trials, this approach will not always be feasible. In particular, a randomized regulatory experiment will not always be justified by the costs, in particular the costs resulting from the disruptive effects of randomization. Additionally, randomized trials can raise legitimate fairness concerns among regulated actors.

ii.   In choosing whether to structure the experiment as a sunset or permanent rule, agencies should choose the structure that is consistent with the probability that the experiment will be a success. In other words, if it is sufficiently likely that the experimental results will justify adopting, on a permanent basis, the rule that is the subject of the experiment, the agency should structure it as a permanent rule when it adopts the experiment. Otherwise, it should be structured as a sunset rule, which expires automatically once the experiment is completed.

iii. Agencies should give considerable thought to how to ensure that correct lessons are drawn from the experiment. Options include commissioned studies, expert panels, open forums and public calls for papers. This decision turns on a number of different factors, including the scope of the issues in question and the need to bring to bear expert analysis on those issues; the agency's internal expertise; the availability of, and the agency's awareness of, outside experts; the scope of the issues to be analyzed; and the extent to which the agency is aware of relevant experts in the field.

3. Legal and Other Considerations

i.   Agencies should give serious thought to how courts are likely to review a given regulatory experiment. Although the issue is far from settled, there is reason to believe that courts will defer to regulatory experiments. This conclusion is based on the general deference courts apply to an agency's evaluations and assessments of data within its expertise and the preference of courts for procedures that cause agencies to produce more information about a given regulatory action. Where courts have shown a reluctance to defer to an agency's evaluations and assessment of data within its expertise, courts seem to be motivated by a desire to cause the agency to produce more information. Regulatory experimentation is inherently information-forcing, which serves this judicial concern.

ii.   Executive agencies should give special consideration to whether a regulatory experiment requires review by OIRA. One case where OIRA review is warranted is where a regulatory action is "likely to result" in a rule that has an annual effect on the economy of at least $100 million. Special consideration should be given to those regulatory experiments where the experiment itself doesn't involve a rule that meets this economic significant test, but the experiment is likely to result in a rule that does.

iii. Agencies should give consideration to how stakeholders are likely to react to a regulatory experiment. The preferences of stakeholders are likely to be more complicated when it comes to a regulatory experiment than in the non-experimental context. This is because opposition to a regulatory experiment can come not simply from stakeholders who oppose the rule that is the subject of the experiment, but also from those who favor the rule but simply do not think it should be adopted on an experimental basis. In both cases, agencies can likely encourage stakeholder support by marshalling theoretical work from the relevant community of experts suggesting that, even though the theory alone would not necessarily justify adopting the rule on a permanent basis, it would justify making it the subject of an experiment.

One useful approach for thinking about whether to adopt a regulatory experiment is a break-even analysis.[1] In other words, instead of quantifying presumably unquantifiable or unmonetizable benefits, agencies could specify how high the benefits would need to be, relative to the next best alternative, in order to justify the costs of experimentation. They could then evaluate the likelihood that the benefits of the rule would meet or exceed that threshold.

Breakeven analyses are used all of the time. In fact, they are so ubiquitous, we probably sometimes do not realize that we're in the middle of one. For example, let us say that a family must decide whether to stay in a cheap or expensive hotel for a vacation. Which one should they choose? It is difficult to quantify the benefits of the hotel. How does one put a number on the utility derived from being met with fresh cut flowers in the lobby or fine crafted molding on the ceiling? Although one might not know that number, one can calculate the cost of the experience fairly easily. Once armed with that information, one can then ask whether that cost is likely to be recouped in utility from the experience. The answer to that question might depend on a number of other considerations, like whether the vacation involves the whole family, kids and all, or whether it is a more quiet getaway just involving the parents. The point is that this type of cost-driven breakeven analysis can be a useful tool not only in one's domestic live but in agency decision-making too.

When it comes to experimentation, the reasoning is very similar to our hypothetical family making decisions about their hypothetical vacation. The reasoning focuses on the costs of experimentation, which it is assumed we know, and then extrapolates to what the net benefits would need to be in order to justify the experiment. However, the calculation is slightly more complicated because it also requires one to know the risk that the rule will be a success and, if not the precise magnitude of downside risk, at least an idea of how the net benefits in the event the rule is a failure compare to those of the status quo or the next best alternative.

To illustrate, let us consider the following hypothetical. Let us imagine that the agency is once again choosing between a rule with near certain net benefits and a riskier rule. Let us assume that the agency calculates that the non-risky rule generates $50 in net benefits on an annual basis and $500 in

---

[1] *See, e.g.*, Cass R. Sunstein, *The Real World of Cost-Benefit Analysis: Thirty-Six Questions (And Almost as Many Answers)*, 114 COLUM. L. REV. 167, 195 (2014) (citing Cass R. Sunstein, *The Limits of Quantification*, 102 CALIF. L. REV. 1369 (2014).

net benefits over its useful life, assuming a discount rate of 10%.[2] Let us further assume that the agency knows that the risky rule has a 20% chance of generating an unknown yet greater net benefit than the non-risky rule (in other words, more than the $500 in net benefits over its life) but an 80% chance of generating $0 in net benefits. Finally, let us assume that the cost of experimenting is $25. The question then is how much value does the risky rule have to create, in the event that it is a success, in order to justify an experiment of a duration of one year?

The answer is given by the following "break-even formula for regulatory experimentation":

$$\text{Net benefit if risky rule is a "success"}^3 = \frac{B_{SQ}*(1-M+M*p+d*p)-c}{p*(1+d)}$$

Where,

1. B_{SQ} is the annual net benefit of the status quo or the "non-risky" rule;

2. M is the net benefit of the risky rule in the event it is a failure, expressed as a percentage of B_{SQ};[4]

3. p is the probability that the risky rule will be a success;

4. d is the multiple to which the annual net benefits of a given rule are applied to calculate the net present value of those net benefits, beginning at the end of the one-year experiment; and

5. c is the costs of experimentation.

Thus, in our example, B_{SQ} equals $50; M equals 0; p equals .20; d equals 9.09[5] and c equals $25. Plugging those values into our formula above yields the following result: $102. In other words, given these assumptions, the risky rule would need to generate an annual net benefit of at least $102

---

[2] In that case, the lifetime net value of the rule is given by $50/.10, which equals $500.

[3] For a derivation of this formula, see *infra*.

[4] In other words, if we thought that, in the event it was a failure, the rule would result in a net benefit that is one-half of that of the status quo, M would be equal to ½.

[5] We assumed that the discount rate was 10% and that the experiment would last for one year. This means that d is given by the following expression: $\sum_{t=2}^{\infty} \frac{1}{(1+.01)^{\wedge}t}$.

(or more than twice as much as the status quo of $50) in order to justify a regulatory experiment. Of course, after running this analysis, the agency may still not know exactly what the benefits of the experimental rule are. But it now knows how much those benefits would have to be in order to justify the experiment.

What follows is the derivation of this break-even formula for regulatory experimentation:

Let us assume that B represents the annual net benefits of a non-risky rule (what we have also called in this report the status quo or the next best alternative). Assuming a discount rate of 10%, the net present value of these annual net benefits can be represented as follows:

(1) $B + \sum_{t=2}^{\infty} \frac{B}{(1+.01)^t} = 10B$

To simplify this expression, let's redefine as d the expression, $\sum_{t=2}^{\infty} \frac{1}{(1+.01)^t}$, such that the expression in (1) can be rewritten simply as, $B + B \cdot d$

Let us assume that a risky rule produces a high net benefit if it is successful and a low net benefit if it is a failure. More specifically, let us assume the risky rule will with probability p yield an annual net benefit of x and will, with probability $(1-p)$, yield an annual net benefit that is a fraction a of the non-risky rule's annual net benefit of B. In other words, with probability $(1-p)$, the risky rule will yield an annual net benefit of B x a.

Assume a one-year experimental period whereby the risky rule is adopted and then can be reversed and the status quo restored after one year, in the event that it turns out to be a failure. The expected value of that experimental rule is given by the following expression:

(2) $-c + p \cdot (x + x \cdot d) + (1 - p) \cdot (B \cdot a + B \cdot d)$,

where the expression $p \cdot (x + x \cdot d)$ represents the expected net benefits of the risky rule in the event that it is successful, discounted by probability p, and where $(1 - p) \cdot (B \cdot a + B \cdot d)$ represents the expected value of the experiment in the event that the risky rule is a failure. In the latter case, the risky rule will, solely for the duration of the one-year period of the experiment, yield a low net benefit of $(B \cdot a)$, with a probability of $(1 - p)$, and it will produce a net benefit of $(B \cdot d)$, which reflects the net present

value of restoring the status quo after the one-year experiment, once again discounted by the probability of that occurring, which is $(1 - p)$.

We now need to set the expected value of the regulatory experiment with the risky rule, expression (2), equal to the expected value of the non-risky rule, expression (2).

(3) $B + B \cdot d = -c + p \cdot (x + x \cdot d) + (1 - p) \cdot (B \cdot a + B \cdot d)$

Re-arranging and solving for x, expression (3) simplifies to the following:

(4) $x = \dfrac{B \cdot (1 - a + a \cdot p + d \cdot p) + c}{p(1 + d)}$

Expression (4) is the break-even formula for regulatory experimentation. In other words, it tells us how high the net benefits of a risky rule have to be in order to justify the regulatory experiment, given the costs and some assumptions about the downside risk, relative to the status quo or the next best alternative.