Chapter 7

Redundancy

Our final concern with safeguards—trigger mechanisms that may induce compliance with the federal distribution of authority—is their imperfection, leading to inappropriate punishment frequency.\(^1\) With imperfect safeguards, the trigger may react too quickly, or not at all. If they fail to react, then the system’s ability to maintain compliance is reduced. On the other hand, punishment can be too frequent. Although the equilibria derived in Chapters 3 and 4 acknowledged that punishment will occur even when no one deviated (beyond the tolerated level of non-compliance), if the safeguard is flawed, punishing too frequently (more than efficient), frustration with the system mounts, and the union cannot be sustained as member governments exit.

Referring to the baseline model of safeguards in Chapter 4, a safeguard’s imperfection has two sources: a signal—observation (\(\omega\))—which is the safeguard’s impression of reality, distorted by noise, and a threshold (\(T\)). The latter imperfection arises because safeguards are not necessarily designed efficiently. Instead, safeguards are often political creatures, or designed for other purposes, or have interests of their own that are tangential to the compliance problem. To induce the optimal degree of compliance with the distribution of

authority, participants count on trigger mechanisms, either their own threats of retaliation or institutional safeguards, to react when opportunism is suspected.

If we were only concerned that the safeguard would fail to react then the solution would be so straightforward that it would hardly merit a chapter: the more the merrier, two heads are better than one, pack a second parachute. Add safeguards. But we have the former problem as well: safeguards sting; if they trigger too frequently, or inaccurately, then they flip from being incentives to dis-incentives, reducing both utility and compliance.

In addition to extraneous punishment, multiple safeguards further create obstacles for the federation to remain productive in a changing environment by limiting its ability to adapt. This problem is known generally as a status quo bias: when decision-makers have conflicting opinions about the direction of change, policy stagnates. We know of the well-established argument that as the number of veto players increases, policy is more stable (e.g. Tsebelis 2002). The down side of stability is that the system cannot adjust when circumstances demand it. Fernandez and Rodrik (1991) argue that utility-enhancing institutional reform is often resisted because of the uncertainty associated with how the new system will distribute benefits. This uncertainty can only be resolved through experience. In theory, an advantage of federalism is experimentation, but often innovation requires structural change (perhaps, by decentralizing policy jurisdiction). Rather than enabling experimentation, safeguards exacerbate the tendency identified by Fernandez and Rodrik. A federal system of safeguards that works too well may stifle the very system it is designed to protect.

Alas it would seem that we have a trio of contradictory concerns: failing to fire, firing too frequently, and creating a status quo bias. In this chapter we take up this design riddle.

---

2See the discussion in Persson and Tabellini 2003:27–9.
7.1 The Issue: Imperfect Safeguards

To this point in the book we have considered the safeguards to be impersonal, disinterested trigger mechanisms, even employing the language of “construction” to make them appear to be distinct from strategic players. However, these institutions are themselves composed of strategic actors, and any sanctioning reaction must be self-enforcing. In order for any safeguard—structural, political, or judicial—to deter free riders and therefore encourage productivity, it must establish a threshold that when crossed, triggers a punishment. On paper, we can derive the efficient threshold and punishment combination which balances the degree of compliance against the probability of a punishment. But real safeguards are not simple if-then rules. Rather than being the passionless, detached referees of theory, the safeguards are composed of people with ambitions, idiosyncrasies, and weaknesses. The real people behind the triggers may set thresholds and punishments for private reasons, or may have perspectives that causes them to interpret observations differently. Real people may pull the mechanism away from optimality.

Threshold Inefficiencies

The sole source of inefficiency that we examined in the last chapter was a safeguard’s punishment capacity. Chapter 6 showed that an insufficient punishment could be supplemented by a complementary mechanism; together the safeguards would more closely approximate efficiency. In some cases, the complementary safeguard failed, as in Figure 6.2 where the mild institution was ineffective, and ignored. While Chapter 6 focused on the punishment, $Q$, as the source of these deficiencies, the safeguard’s trigger could also be at fault. The threat capacity of a trigger mechanism is a combination of the punishment and its frequency: a weak punishment can be invigorated through a lower threshold, with more frequent punishment. In this figure, it is easy to see that a small shift in the mild institution may cause a lumpy shift in behavior, as equilibrium levels of shirking shift from approaching full compliance, to
that sustainable by the severe trigger alone. The disutility of a punishment cannot be calculated without both the threshold and the severity of the punishment, since the threshold indicates the likelihood of the punishment.

For example, perhaps the mild safeguard is a court. The court sets its threshold for reasons external to this model: strategic, behavioral, or legal reasons. The forces that act upon the court may change, causing its threshold value to shift. Or consider political safeguards derived from parties. The party’s motivation is clear: get its candidates elected. Individual politicians within the party want to be reelected, or elected to higher office, or to be offered an economically secure position after they leave office. None of these goals necessarily puts them in a position of wanting to patrol the federation when punishing may clash with their own self-interest. Also, these goals may shift. Therefore, with any of the safeguards, the operation of the trigger may shift: the definition of what actions will trigger a punitive response may change over time, and institutions may drift away from efficiency. If the safeguards have inconsistent threshold settings, or generally set them for reasons only indirectly related to maximization of the social welfare function, then the safeguard may be inefficient, and even ineffective. In what follows, we continue to assume that the threshold and punishment are exogenously set, but now let us consider that the threshold may be politically determined, and therefore imperfect.

Diverse Observations

The trigger is composed of two variables: the information received by the mechanism (the observation, \( \omega \)) and the threshold, \( T \), that when crossed causes the trigger to fire. Knowing the threshold is not sufficient to know the likelihood of punishment. Signals, too, are a source of imperfection. Modeling convention flows from analytical convenience: typically we assume that the signal—the observation—is common to all. That is, we assume that all safeguards process the observation in the same way, categorizing it identically. But just as witnesses to a crime differ in what they remember of events, safeguards too have different
7.1. **THE ISSUE: IMPERFECT SAFEGUARDS**

perspectives which cause them to focus on different facts, or interpret facts diversely. We may gain something important when allow for diverse perspectives, so we pursue the alternative in this chapter.

Signals are not singular bits of information: instead they have multiple attributes exhibited across multiple dimensions. Any particular signal may not convey the complete set of attributes. When this distortion is random, it creates inference problems as described above, in Chapter 3, when we were motivating stochastic error. In interviewing eyewitnesses, one expects certain features of a crime scene to stick in the memories of all witnesses—the amount of blood, the speed and direction of the escaping car—but other details to be remembered differently or not at all—how many shots were fired, even the sequence of events. We wouldn’t necessarily expect any pattern or predictability to the distribution in reported observations, but they would differ from another.

With imperfect mechanisms, we are particularly interested in systematic flaws connected to monitoring capacities of the safeguarding agent. It is possible that an agent’s perception may be influenced by characteristics particular to the agent; one witness noticed the quality of the perpetrator’s coat, having tried a similar one on at the store; another recognized the shout to the accomplice, unintelligible to others, because her grandmother spoke the same dialect at home. Valuable information connecting a suspect to the crime may come from only one witness, and, importantly, could only be provided by that particular witness.3

Likewise with federalism, a safeguard may only “notice” certain attributes. The subset that it considers is its interpretation of the event. Interpretations are closely tied to a safeguard (or the agent directing that safeguard)’s set of past experiences. This bias creates imperfection in the safeguard’s response. Politicians, the soul of the political safeguards, tend to notice what their constituents bring to their attention, or focus on the elements of policies that are most likely to influence votes. Voters, en masse, may be surprisingly

---

3The theory of diversity of perspectives and interpretations is developed in Page 2006 and a model of signal subjectivity is from Hong and Page 2005.
good at making informed choices when issues are salient to them (Lupia and McCubbins 1998, Hutchings 2005), but much policy never hits their radar. For other safeguards, signal perception can be independent of interests. The court is an excellent example. Any case has many attributes. A judge’s legal training can cause her to focus her attention on a particular subset, weighing, say 5 of the case’s 10 attributes more heavily. A different judge, trained at a different law school, may look at 3 of the same attributes and 2 distinct ones. Notice that this diversity in judgment is different from error correction where appeals improve the judiciary’s accuracy because it provides a second look (Shavell 1995). Instead, consider the potential for interpretations of the same facts, as Daughety and Reinganum 2000 justify the appellate process.\footnote{In their model, the appeals court and the appellant from the trial stage each have a private signal about the superior court’s interpretation of the law; the appeals court uses the appeal as a signal of the superior court’s position.}

Finally, there is diversity in the way that different agents will map their observation into a decision about what happened. Kornhauser 1992 and Cameron and Kornhauser 2005, in theory applied exclusively to the court, describe a court’s decision-making procedure as the application of a legal rule that partitions a multidimensional case space; the method of partitioning the space will ultimately reduce the case’s vector of attributes into a simple binary set of guilt or innocence. It is not the same as setting a threshold: instead it is a method of locating the observation on one side or the other of the threshold. Page 2006 describes the same process generally as a predictive model: a mapping of facts, as encoded by an agent, into a judgment or an inference about the truth.

Signals then can be distorted—can deviate from the assumption of common observation—in three ways: agents may observe different aspects of the signal, sometimes in a manner that is not random but instead a function of the agent (an encoding or interpretation); agents may have different perceptions of what the dimensions mean; and agents may have different decision rules or mental models that cause them to process the signal differently, arriving at
7.2. OVERCOMING—AND HARNESSING—IMPERFECTION

a different judgment about reality, and therefore perhaps even disagreeing about which side of the threshold the signal falls.

These flaws—in the way the threshold is set and in the consistency of the perceived signal—are observationally difficult to disentangle but analytically distinct, and in fact have contradictory influences on the federation’s robustness. If the threshold is set too low, the punishment is triggered too frequently; if the signal is systematically perceived as too low, the punishment isn’t triggered frequently enough.

7.2 Overcoming—and Harnessing—Imperfection

Our trio of problems that arise from institutional imperfection—the two-horned calibration of punishment frequency as well as the orthogonal hold-up problem—resonate with problems throughout the social sciences. In the next few pages we briefly examine these parallel fields for advice on the resolution of our problem.

System Reliability: Von Neumann to Ostrom, by way of Bendor

The literature on system reliability captures half of our puzzle very well. Bucking the conventional wisdom favoring governmental streamlining and equating redundancy with waste, Bendor argued convincingly that it is the system without redundancy that is poorly designed. A system without redundancy depends upon every component to function perfectly. If any component fails, the system fails. The old adage “a chain is only as strong as its weakest link” may have inspired a cult game show but it is defeatist advice for organizational design. Instead, Von Neumann 1956, Landau 1969, Bendor 1985, and Ostrom 1999 all embrace inherent human fallibility: it is impossible to eliminate error, but we can design an organization to overcome internal flaws, to minimize the consequences of the failure of any one component. Von Neumann and Bendor’s concern is a failure in functionality, a failure to react. When the problem is information failure, parallel information processing is bene-
ficial (Simon YEAR; Cohen 1981). Ostrom’s (e.g. 1990, 1999, 2005) research on successful management of common pool resources engages both information and reaction concerns.

Asks Landau, “Can we . . . build an organization that is more reliable than any of its parts?” (1969:350) The recommendation is straightforward: introduce redundancy when failure is possible. Successful parallel systems have two characteristics: (1) redundant components should have fully overlapping functionality and (2) as much as possible, they should have uncorrelated vulnerabilities. That is, while their capacity should overlap, they should fail for different reasons. Note that the components need not be identically designed; actually, it is their diversity that makes them useful as insurance, because the redundant component can only be useful if it does not fail when the first does. The advantages of redundancy are easy to compute: if one component has an estimated failure rate of 10%, then two, established in tandem, would have only a 1% chance of failure.

But what of reliability’s endurance? Momentary reliability is an oxymoron. Reliability implies that a system functions now and will continue to function in the future when circumstances are different, perhaps unpredictably so. A reliable system adapts to changing circumstances. It innovates.

Part of what makes a system healthy and adaptive is the element that cannot be planned: the adjustments that will emerge through experience. Landau 1969:347–50 praises redundancy for increasing the potential for innovation. Landau cites Von Neumann to suggest that redundancy not only improves reliability, but enhances adaptability.

Self-organizing systems exhibit a degree of reliability that is so far superior to anything we can build as to prompt theorists to suggest ‘that the richly redundant networks of biological organisms must have capabilities beyond anything our theories can yet explain.’ In Von Neumann’s phrasing, they ‘contain the necessary adjustments to diagnose errors as they occur, to readjust the organism so as to minimize the effects of errors, and finally to correct or to block permanently the faulty component.’ Error refers here to malfunction, and Von Neumann states that there is now little doubt that they are ‘able to operate even when malfunctions set in . . . [while] their subsequent tendency is to remove them.’ (Landau 1969:350)
7.2. OVERCOMING—AND HARNESSING—IMPERFECTION

Therefore, what we are looking for, to quote Ostrom 1999, is a parallel adaptive system. Ostrom expanded Bendor’s phrase “parallel systems” to emphasize the organic nature of her social systems, as opposed to the physical architecture of Bendor’s transportation systems; she also wanted to emphasize the evolution and flexibility of the system. While in a social system like federalism we are not dealing with Von Neumann’s biological organisms, much of how a social system evolves is unplanned by the designers, even as the mutations occur within the framework established at the founding. While one hopes that the initial design is intelligent, the aim is the antithesis of “intelligent design”: a complex environment introduces challenges and needs unforeseeable to the social system’s planners. As with Von Neumann, Landau, Bendor, and Ostrom, we are interested in a system that maintains its functional capacity despite a changing environment. To be robust, it must be adaptive, a point we will return to below.

We have received excellent advice on solving the problem of safeguards that fail to react: we need to construct redundancies. We have been given hints about solving our third problem of adaptation, escaping the tendency to stick with the status quo, but we will want to probe the literature further for increased understanding. And we have not addressed the concern of the safeguard that triggers too easily, punishing too often.

**Type I and Type II Errors**

Punishing the innocent is a significant concern of any system of justice, and must be a priority for designing a system of safeguards knowing that they are inevitably flawed. Redundancy here can only hurt if the safeguards are flawed. If one resists firing (correctly) but the other punishes, the damage is done. In the movies, if two women misunderstand a guy, thinking that he has said something rude, the fact that one resists the temptation to slap him across the cheek doesn’t make up for the drink the other one threw in his face. He’s still wet, even after the misunderstanding is cleared up.

Our problem has a familiar cognate in statistics: the tradeoff between the risks of making
CHAPTER 7. REDUNDANCY

a Type I and Type II error. With the Type I error, or false positive, the null hypothesis is rejected incorrectly, while the Type II error, or false negative, fails to reject the null hypothesis when it should. As opposed to the reliability problem posed above, where the concern is that a component will fail to act—the Type II error—it is the opposite issue that statistics views as more problematic: rejecting a hypothesis incorrectly.

For any fixed sample size, Type I error and Type II error risks are off-setting: to reduce the risk of one, you increase the risk of the other. The prescription is again straightforward. Consider the Type I problem first. Set the critical value (the threshold) based upon an acceptable failure rate. To augment the power of the test (its ability to avoid Type II errors, false negatives), increase the sample size.

Translating the false positive to the federal context, the consequence of a safeguard failure that causes it to trigger a punishment too often could be huge if the safeguard triggers intergovernmental retaliation, including retaliatory opportunism and even withdrawal from the union. So from statistics we learn that these problems compete with one another, and if we fear convicting the innocent more than undue leniency, then we should solve that problem first. So much for the reliability lessons above; it seems that adding safeguards will only augment the problem by making punishment more certain. We can’t have it both ways. Or can we?

Insurance and Confirmation

We have conceived of redundancy too narrowly. There is a second form of redundancy that is helpful for solving the problem of overly frequent punishment. If a safeguard punishes too frequently, one problem may be that its perception of its signal is flawed in one of the ways described above. If this is the case, then a second safeguard, with an independently-drawn signal, could improve the efficiency of the system’s reaction. Here the second signal (or safeguard) is not insurance but confirmation. Recall the Green & Porter world of oil cartels, a model described in Chapter 3. For them, the unique signal is the price of oil: if
the price drops, oil producers take it as a signal of a possible increase in supply, caused by
the defection of one of the cartel members. Green and Porter did not consider the possibility
of two signals, but they could have. For example, a second signal could come from input
pricing. If independent truckers and tankers suddenly demand a higher price for transporting
oil, it could be due to a higher demand for their services, also a signal that could indicate
a deviation. Or, conversely, if the output price of oil drops, but the input price remains
steady, then cartel members may be more confident that the price hit a demand shock and
no mutual punishment is needed.

The theory of the second signal is related to the results in Sah & Stiglitz (1985, 1986),
who compare two organizational forms, the hierarchy and the polyarchy. In a polyarchy, a
project is accepted if one (of two) agents accepts it. Both must reject it for it to be rejected.
In a hierarchy, two agents are necessary to approve a project, and therefore only one is
sufficient to reject it.\(^5\)

Translation of their model to federalism goes as follows. Consider their “project” to be an
action taken by a member government, and the polyarchy and hierarchy to be characteristics
of a system of safeguards. The hierarchy echoes the reliability literature. It is a redundant
system of safeguards, with one backing up the other (insurance); punishment is inflicted if
either one of the safeguards is triggered. Likewise, equivalent to the polyarchy is a second
form of redundancy; this time, two safeguards must be triggered before the punishment is
levied. In this sense, the redundancy is confirmatory.\(^6\)

\(^5\)It is temptation to draw parallels between the names for these organizational forms and federalism's
federation or unitary system, but within this section we are concerned only with the relationship between
the safeguards, not between the governments.

\(^6\)Adapting the model from Chapter 4, incorporating the Sah and Stiglitz approach, might look as follows.
Consider an additional institution that does nothing more than provide a second look—an additional measure
on the signal. Now, if \(x_t^2 + \epsilon > T\), rather than immediately triggering the a punishing sanction, let a second
institution get its own independently-drawn signal. Let the two signals be \(\omega_1\) and \(\omega_2\). Note that since
the second signal is drawn from the same distribution (that is, is a result of the same action), its mean
and variance are identical, so the signals are correlated. Remember that in equilibrium no one deviates
from the common behavior (usually some degree of partial compliance), therefore all signals less than the
threshold are due to stochastic error. In equilibrium, it was unlikely before that \(\omega > T\); now, we need both
CHAPTER 7. REDUNDANCY

The hierarchy/insurance commits more Type I errors: it builds in a level of conservatism that rejects more good projects than the polyarchy, but the danger of the polyarchy/confirmation is that it accepts more bad projects, thereby committing more Type II errors through its relative leniency. Preference for one type of organizational structure over another would depend upon the relative utility of making a Type I versus a Type II error. In Sah & Stiglitz, this calculation is based upon the relative prevalence of good projects to bad projects, as well as the relative loss from bad projects compared to the gain from good projects. We are urged by both the statistical theory and Sah and Stiglitz’s theory of polyarchy and hierarchies to weigh the cost of punishing the innocent against the cost of undue leniency.

While this theory gave us more aid in defining how to avoid overpunishment, it didn’t directly provide a method for eliminating both problems. However, the theory was limited by the simplicity of the problem space: review of undifferentiated projects by undifferentiated agents. With federalism we have varying levels of non-compliance as well as safeguards that are quite distinct from one another. There is no reason why we couldn’t design a system that employed different relationships between the safeguards, where some confirm while others act as insurance.

In medicine, the confirmation is known as the second opinion. If a doctor says your cholesterol would be lower if you lost weight and exercised more, the recommendation is

\[ \omega_1 \text{ AND } \omega_2 \text{ to exceed the threshold to trigger a punishment. Adding an institution in this way, by requiring a “second look,” means that fewer punishments will be triggered. The punishment regime is best thought of as a compliance-maintenance regime; that is, as long as the incentives are adequate, in equilibrium no one deviates and the punishment is triggered by random environmental noise. When we reduce the frequency of the punishment, the equilibrium behavior induced may be higher, since each player gains more from being in a cooperative state. Therefore a second, identical, complementary institution works on the first two goals, maximizing cooperation and minimizing the cost of enforcement. This is the most common application of redundancy: reduction of the Type I error.}

On the other hand, if the concern instead is about the likelihood of a Type II error, that the institution might not trigger frequently enough, one might introduce a second signal and trigger a punishment if \( \omega_1 \) OR \( \omega_2 \) exceed the threshold. The frequency of punishment increases. This might be a good solution in the cases where a more severe punishment is not available or the threshold cannot be moved to a more efficient position.
only beneficial, and you’d be wasting your time to seek another doctor’s advice. But when treatment for a diagnosis is itself potentially injurious, such as embarking upon a long-term drug intervention or surgery, patients regularly seek a second, and even third, opinion, wisely. When the downside of incorrect punishment is high, we’d want to have a redundant safeguard as confirmation, to double-check the observation.

**Exploration via Imperfection**

We are still left with the problem of how to ensure adaptation. Just as Type I and Type II errors involve a trade off, so it would seem that improving the reliability (reducing Type II errors) of the system of safeguards naturally reduces the adaptability of the federation: strict compliance means no experimentation. There is no straightforward way to eliminate this problem without softening up on perceived opportunism.

March 1991 describes the tradeoff between exploration and exploitation in organizational learning about a multi-dimensional reality space. In his model, learning is bi-directional. In any given time period, with some probability the organization adapts its beliefs along a dimension of reality if its belief differs from a preponderance of its members. It uses these beliefs to develop the organizational code, a set of rules and procedures. Likewise, with some probability individuals are in turn socialized into the organization, adopting its beliefs and conforming to its prescribed practices. Those individuals that conform with high probability are called “fast learners” while those that have a low probability of conforming are “slow learners”. Which type of learner is better for the organization?

The organizational code is a set of best practices meant to maximize the organization’s productivity given the aggregated knowledge of its individuals. The organization can choose to *exploit* current knowledge by rewarding compliance with its code, essentially by retaining fast learners (conformers). This method has intuitive appeal until one remembers that the environment changes, and if the individuals have conformed to the organizational code, they stop sampling the environment. The organization has no other way to learn than through
CHAPTER 7. REDUNDANCY

its members, and so it fails to adapt to new circumstances. Alternatively, it could employ slow learners, who continue to explore the environment. By bucking the organizational code they do not take advantage of all that the organization has learned, but there is some chance that in resistance they teach the organization something valuable that helps it to adapt. “The fraction of slow learners [those who socialize slowly, most resistant to adopting the organization’s code] is a significant factor in organizational learning” (March 1991:77). Non-conformists improve the organization’s performance.

The most immediate application of March’s theory to federalism is to treat the distribution of authority as the organizational code. A system of safeguards is faced with the choice between adhering to the distribution of authority as currently understood versus tolerating deviance. Deviations are a form of experimentation and can help the federation’s members to learn more about the current state of the environment and therefore how to (and when to) adapt the distribution of authority to improve the federation’s performance.

The conflicting belief, as evidenced through the opportunism of one or two member governments, is unlikely to cause the whole federation to change course. But with some probability the union will revise its beliefs about what is best for the union when more and more governments share beliefs dissimilar from the status quo. This updating, this learning about the environment, and therefore the best design, is only possible if governments are allowed to explore. Opportunism must be possible.

At this point I’ll pause to deliver a sympathetic note to the reader: We’ve certainly changed our tune about opportunism, the story’s villain until now. But we see from March that some non-conformity is the source of learning about a changing environment. Without this experimentation, the federation cannot adapt and will not remain robust.

As an illustration, consider the case of Iowa’s regulation of the length of truck trailers. Iowa set a maximum trailer length that was lower than that commonly used by interstate
truckling agencies. In 1981, the Supreme Court struck down Iowa’s regulatory legislation, affirming lower court decisions that it was unconstitutional because it burdened interstate commerce. Iowa defended its regulation with an argument that the reduced trailer length was safer. What if it had been able to show that the regulation reduced the number of fatal accidents? Then its shirking would have taught the rest of the union something about the state of the world, about the connection between actions and outcomes. (And if all states simultaneously adopted the same regulation, or (ideally) if haulers did so voluntarily, then the regulation would not be held unconstitutional. It was its variance from standard practice, not the regulatory content itself, that made the regulation unacceptable to the court.)

Although opportunism can be useful, we needn’t embrace it unconditionally. We can qualify our enthusiasm for it in three ways:

1. Only mild opportunism is tried.

2. Mild opportunism is still subjected to the same standards, including the potential of punishment.

3. Accepted changes are in the best interest of the union as a whole.

Experimentation should not put the union at risk. Flagrant acts of opportunism are likely to trigger a severe reaction like intergovernmental retaliation; their potential benefit as learning exercises are swamped by the destructive cost to the union. Second, to prevent any member government from taking advantage of this need for experimentation, the policy of punishing mild opportunism should not change. Finally, we need a theory of how opportunism/innovation leads to changes in the intergovernmental relationship.

Downs and Rocke’s terrifically titled book *Optimal Imperfection* (1995) presents that case that a system of safeguards that seems too mild may actually be beneficial. Governments often are uncertain about future domestic demands, and therefore hesitate to commit to

---

treaties that they think they might have to (temporarily) break in the future. With low punishments, governments can deviate when necessary to appease domestic audiences, and return to compliance when interests suit. We see the value of this repeated in the federation: I’ve called mild safeguards that because their punishments are not particularly severe. Governments can deviate in small ways at small cost when they have short-term supplementary benefits from opportunism. Returning to the truck trailer regulation, Iowa knew that its legislation had a good chance of being struck down, but the negative consequence of the judgment was minimal. Between enactment and the Court’s judgment it was able to reduce the traffic on the interstate highways, and although it was unable to show a broader benefit (in terms of improved safety), without this experimentation we would not have evidence to the contrary.

There is a second sense in which federations are optimally imperfect. Under the conditions we’ve specified in this book, full compliance cannot be sustained (Bednar 2004a). Opportunism is built into the system. Some slippage—some mild deviations from the distribution of authority—are a normal part of federalism. It is as if federalism couldn’t help but heed March’s advice; non-conformists are automatically incorporated into the union. By being incapable of enforcing perfect compliance, governments can experiment around the edges of the distribution of authority at no cost, and the system as a whole may learn from what they find.

For example, consider California’s anti-pollution measures, including requiring catalytic converters on new cars sold within the state, and more recently, a quota of low- and zero-emission vehicle sales. Automakers cannot afford to avoid the regulations by pulling out of the state’s market, and production technologies limit their capacity to build a separate line of cars for California alone. Therefore California is in effect regulating national auto sales. The pecuniary externalities imposed by California have both diffuse and concentrated incidence: the changes raise the cost to purchase a new vehicle, and when prices rise, demand lessens,
7.2. OVERCOMING—AND HARNESSING—IMPERFECTION

so the effect is felt doubly in states with an economic dependence on auto production, like Michigan. Legal challenges are mounting, but unlike the truck regulation case, it is not clear that they will prevail. And political and electoral safeguards are showing no signs of intervening: to the contrary, several northeastern states have suggested that they will soon follow California’s lead. As some automakers begin to make adjustments, California’s experimentation will help us to better understand two variables: first, what is the effect of reduced emissions, even locally, on a global environmental problem? And second, what is the real cost of the new technology? The externalities generated by California’s regulation are indisputable. But a system that tolerates mild opportunism may end up learning more about the environment than one that shuts down all non-compliance.

Two problems: first, our plan seems to contradict what we thought we understood about design above. That is, we believed that for mild opportunism, redundant safeguards would be useful, as insurance. But to minimize the likelihood of punishing the innocent, we proposed a confirming redundancy for severe punishments. Wouldn’t this make major acts of opportunism even more likely? Second, these insights are useful, but now we need a better sense of adaptation. How do we have confidence that the alterations are in the common interest?

A Common Culture?

A second application of March’s theory of the value of non-conforming explorers can help us to understand a force that haunts the federalism literature: culture.

THE EXCERPT ENDS HERE WITH MY APOLOGIES. REMAINDER AVAILABLE WITHIN DAYS FOR THOSE WHO WANT TO KNOW HOW THE STORY ENDS.