

Technology Law

Chapter 10: The Precautionary Approach

BJ Ard and Rebecca Crootof

Last updated May 2026. Comments welcome!

Cite as: BJ Ard & Rebecca Crootof, [The Precautionary Approach](#), in [Technology Law](#) (2026).

10. The Precautionary Approach	2
A. Introducing the Precautionary Approach	2
The Risks of Risk Management	4
B. Bans	5
Rhetorical Strategy: Types of Arguments	6
Discussion Questions	7
Rebecca Crootof, The Killer Robots Are Here: Legal and Policy Implications, 36 Cardozo L. Rev. 1837, 1883-94 (2015)	8
C. Other Implementations	21
Cass R. Sunstein, The Paralyzing Principle, Regulation, Winter 2002-2003	22
Comprehension Check	34
Discussion Question	35
Noah M. Sachs, Rescuing the Strong Precautionary Principle from Its Critics, 2011 U. Ill. L. Rev. 1285, 1304-26 (2011)	35
Comprehension Check	52
D. Case Study: Generative AI	52
Matthew Tokson, Uncertainty, Catastrophic Risk, and AI Regulation, Lawfare (Sep. 16, 2024)	53
Putting it All Together: Chapter 10	57

10. The Precautionary Approach

A. Introducing the Precautionary Approach

Notwithstanding Zittrain's celebration of procrastination,¹ the risks and unpredictability of generative and other technologies prompt some to favor a more proactive regulatory stance. However, this approach is frequently mischaracterized and dismissed. In the interest of sidestepping the ongoing definitional debates surrounding the appropriate characterization of the precautionary *principle*, we focus instead on discussing the benefits and drawbacks of the precautionary *approach*.

While the permissive approach entails a presumption against regulation unless and until specific harms have been identified as sufficiently likely and significant, the **precautionary approach** is a presumption towards having or enacting regulation to minimize likely or significant harms. Because advocates of the precautionary approach are arguing against the more typical laissez faire approach to regulation, their arguments are most persuasive when they argue both (1) that a precautionary approach is needed and (2) what form it would ideally take. For example, a precautionary approach could range from incorporating additional regulatory burdens (like a licensing regime or proof of compliance with safety standards) to a moratorium on certain uses to a complete ban on a particular technology.

There is an obvious and intuitive appeal to avoiding likely, significant, or irreversible catastrophes. Even a technology that seems safe may ultimately prove harmful, and in worst case scenarios—like overfishing, unchecked pollution, and nuclear contamination—harms may prove impractical or impossible to reverse. By slowing the development or proliferation of a technology, we can minimize its negative impacts.

The precautionary approach can also be used to facilitate information gathering while extending the time frame during which it is possible to craft effective regulation, undermining the common assumption that the permissive approach is the only one that permits regulators to understand how a technology operates before regulating it. New regulation can only do so much when design choices and social understandings of a particular technology are already settled.² In contrast,

¹ See [Chapter 9](#).

² See [Chapter 11](#).

banning a technology postpones the stabilization of both a technology's design and its social uses; less dramatic restrictions can also be designed to facilitate a period of controlled study or to compel developers to provide information. If regulators ultimately lift restrictions on a technology, they will be both more informed and better-positioned to use law to impact the design and usage norms of the technology as it enters widespread use.

This approach also creates a limited mechanism for correcting overly cautious regulation. If a "safe" technology is presumptively and incorrectly classified as unsafe, its proponents have a direct interest in establishing its safety. Public-choice theory predicts that a concentrated interest group like this will mobilize to oppose the regulations; given the likelihood that those promoting a technology are likely to be more concentrated and better able to mobilize than those opposing it, erroneously classifying a technology as harmful is more likely to be corrected than erroneously classifying it as safe.

Of course, the precautionary approach comes with its own risks. It may well chill innovation. The benefits of a technology—and the necessary momentum to overcome restrictive rules—may never manifest if industries experience the precautionary approach as chilling and decide to invest their resources elsewhere.³ Early regulations may also fail to target a technology's harms due to a misunderstanding of the actual risks, and the sense of having already addressed the issues may make it difficult to establish more appropriate regulations later.

The precautionary approach may fail to take sufficient account of the risks and opportunity costs of not using a technology, especially if those harmed by its restriction are a diffuse group or have little political power. For example, the FDA has adopted a precautionary approach towards food irradiation, possibly at the cost of significantly reducing deaths due to bacterial infection. And while further testing may reduce the risk of lethal side effects of a new cancer treatment, the delay may also lead to deaths among those who would have benefitted from it. Again, public choice theory and political economy dynamics suggest that the ability of groups harmed by inaction to engage in regulatory correction will depend on

³ See, e.g., Steven Sinofsky, [211. Regulating AI by Executive order is the Real AI Risk](#), Hardcore Software (Nov. 1, 2023) (arguing that the Biden Executive Order on AI will "stifle innovation before it can happen" and noting also the democratic concerns with regulating by Executive Order, rather than the traditional federal lawmaking process). For a contrary perspective, see Yonathan A. Arbel, Matthew Tokson & Albert Lin, [Systemic Regulation of Artificial Intelligence](#), 56 Ariz. St. L.J. 545 (2023) (arguing for a precautionary approach to regulating AI, both domestically and internationally).

their concentration and relative power. But even if the rule is changed following a successful mobilization, that change may be of little comfort to those harmed in the interim.

Ultimately, it would be impractical to adopt a precautionary approach towards all technologies. Legal actors can, however, adopt broad, tech-neutral prophylactic rules for particularly risky or ethically ambiguous classes of technology, such as new medical treatments or human cloning; for industries that have a history of causing widespread harm, such as reckless bankers; or for dangerous kinds of conduct, such as indiscriminate attacks in warfare.

The Risks of Risk Management

Because the precautionary approach focuses on preventing harm, it may seem intuitive to connect this approach to the language of “risk management.” But this is not quite right. The permissive approach, the precautionary approach, and their conceivable alternatives all seek to mitigate risks. They simply prioritize different risks and different ways of dealing with it.

In fact, many critics have argued that risk management, as a way of framing regulation, actually serves to justify or excuse activities that pose a likely risk of harm.⁴ This puts risk management at odds with the precautionary approach. STS pioneer Langdon Winner was at the forefront of this debate. Writing in the 1980s, he argued that relying on risk assessment and related forms of cost-benefit analysis when discussing technology was treacherous because they displaced the more robust idea of harm prevention.⁵ To portray negative impacts on health and the environment as risks, he argued, was to presuppose the idea that these were acceptable wagers and that the most we could hope for was mitigation.⁶

Contemporary legal scholarship has developed the criticism further. For example, Margot Kaminski critiques risk regulation as an approach to AI.⁷ Drawing on lessons from other fields of law, she identifies the “policy baggage” that risk-based approaches carry—such as the assumption that the technology will be

⁴ See, e.g., Samuel J. Rascoff & Richard L. Revesz, [The Biases of Risk Tradeoff Analysis: Towards Parity in Environmental and Health-and-Safety Regulation](#), 69 U. Chi. L. Rev. 1763, 1793 (2002).

⁵ Langdon Winner, [The Whale and the Reactor: A Search for Limits in an Age of High Technology](#) 144–45 (2d ed. 2020).

⁶ See *id.* at 150–51.

⁷ Margot E. Kaminski, [Regulating the Risks of AI](#), 103 BU L. Rev. 1347 (2023).

used and so harms must be mitigated (rather than limiting the use) and how its focus on quantifiable harms doesn't adequately protect less quantifiable ones, such as dignity, fairness, privacy, and autonomy.⁸

Kaminski's work is also significant in showing that risk management is not a single unified approach—its implementation has taken many forms that impose greater or lesser constraint and accountability.⁹ That core lesson does align with our discussion of the precautionary approach: how a regulatory strategy is implemented can sometimes matter as much as the strategy itself.

B. Bans

Implementing a permissive approach is easy: do nothing. Implementing a precautionary approach requires also addressing the question of what should be done. The simplest—and most extreme—means of implementing a precautionary approach is a ban: if a given technology isn't permitted, there's no need to worry about its problematic side effects!

Since 2012, there has been an ongoing international discussion around banning autonomous weapon systems—weapon systems that, based on gathered information and preprogrammed constraints, are able to independently select and engage targets. In 2014, Human Rights Watch and Harvard International Human Rights Clinic published a report in which they identified and rebutted various arguments against a ban.¹⁰ In response to the claim that “[a] ban on fully autonomous weapon systems is premature given the possibility of a technological fix,”¹¹ they argued, “These highly problematic weapons should be preemptively banned to prevent serious humanitarian harm before it is too late and to accord with the precautionary principle.”¹² Advocates for a ban voiced skepticism that technological developments would address the legal issues associated with autonomous weapon systems, such as the possibility that these systems would be able to comply with the requirement that combatants must distinguish between lawful targets (like other combatants) and unlawful targets (like civilians or

⁸ *Id.* at 1400–03.

⁹ *Id.* at 1403–06.

¹⁰ Human Rights Watch & Int'l Human Rights Clinic, Harvard Law Scho., [Advancing the Debate on Killer Robots: 12 Key Arguments for a Preemptive Ban on Fully Autonomous Weapons](#) 24 (2014).

¹¹ *Id.* at 18.

¹² *Id.*

wounded or surrendering combatants), as well as concerns about “the potential for an arms race, an accountability gap, and moral objections.”¹³ But their main response was solidly grounded in the precautionary approach:

The prevalence of humanitarian concerns and the uncertainty regarding technology make it appropriate to invoke the precautionary principle. . . . Fully autonomous weapons implicate the three essential elements of the precautionary principle—threat of serious or irreversible damage, scientific uncertainty, and the availability of cost-effective measures to prevent harm. The development, production, and use of fully autonomous weapons present a threat to civilians that would be both serious and irreversible, as the technology would revolutionize armed conflict and would be difficult to eliminate once developed and employed. Scientific uncertainty characterizes the debate over these weapons. Defenders argue there is no proof that a technological fix could not solve the problem, but there is an equal lack of proof that a technological fix would work. Finally, while treaty negotiations and implementation would carry costs, these expenses are small compared to the significant harm they might prevent.¹⁴

The advocates for a ban also noted the importance of acting preemptively, before the technology proliferated:

Given the host of concerns about fully autonomous weapons, they should be preemptively banned before it becomes too late to change course. It is difficult to stop technology once large-scale investments have been made. The temptation to use technology already developed and incorporated into military arsenals would be great, and many countries would be reluctant to give it up, especially if their competitors were deploying it.¹⁵

Rhetorical Strategy: Types of Arguments

If you have ever answered a policy question on a law school exam or attempted to convince your family of something over dinner, you are probably already

¹³ *Id.*

¹⁴ *Id.* at 19.

¹⁵ *Id.* at 18.

familiar with the fact that one of the more effective rhetorical strategies is to make multiple types of arguments in favor of your overarching point. A legal argument. A moral argument. A fairness argument, or an efficiency argument, or an administrability argument. A public health, national security, or environmental impact argument. The more types of arguments you make, the more likely you are to hit on a point that resonates with your audience.

The same is true when arguing for or against regulatory action. The more types of arguments you can muster, the more persuasive your overall presentation will be. Of course, each argument must be credible—"reaching" with unfounded arguments may undermine the other points!

Discussion Questions

1. What types of arguments might be made in favor of a more permissive approach to autonomous weapon systems?
2. Conversely, what types of arguments might be made in favor of banning autonomous weapon systems?
3. And, to take the Goldilocks route, what types of arguments might be made in favor of a precautionary approach that was less extreme than a ban? Perhaps a moratorium? A licensing regime? Or regulations that limit the use by certain actors or for certain activities?

Excerpt From

Rebecca Crootof, [*The Killer Robots Are Here: Legal and Policy Implications*](#), 36 Cardozo L. Rev. 1837, 1883-94 (2015)

III. Can Autonomous Weapon Systems Be Successfully Banned?

Human beings are remarkably ingenious at developing new means and methods of killing each other. Throughout history, new weapons have been critiqued for being too ignoble, gruesome, or haphazard—but only some of these weapons have been successfully banned.

Scholars on both sides of the ban debate have mined historical examples for evidence that states will or will not be able to successfully ban autonomous weapon systems. Proponents of a ban point to widely-adhered-to prohibitions on chemical weapons and permanently blinding lasers.¹⁶ Skeptics, noting failed attempts to ban crossbows and aerial bombardment, respond that weapon bans are generally ineffective.¹⁷ But there is little analysis of which of these historical examples is most aptly analogized to autonomous weapon systems. To the extent scholars do discuss similarities and distinctions between earlier banned weaponry and autonomous weapon systems, they tend to ignore the crucial fact that these weapons are already integrated into states' armed forces.

Based on commonly cited attempted and enacted weapon bans,¹⁸ this Part teases out eight traits which seem to increase the likelihood that a given ban will be successful—which is to say, both enacted and effective at limiting the usage of the banned weapon. Based on these characteristics, it concludes that states are unlikely to negotiate a successful ban on autonomous weapon systems.

¹⁶ See, e.g., Human Rights Watch & Int'l Human Rights Clinic, Harvard Law Scho., [Advancing the Debate on Killer Robots: 12 Key Arguments for a Preemptive Ban on Fully Autonomous Weapons](#) 24 (2014) [hereinafter *Advancing the Debate*]; [Memorandum from the Mines Action Can. to Convention on Conventional Weapons Delegates](#) (May 2014) [hereinafter *Mines Action Can. Memorandum*] (originally note 187).

¹⁷ Shane R. Reeves & William J. Johnson, *Autonomous Weapons: Are You Sure These Are Killer Robots? Can We Talk About It?*, 2014 Army Law. 25, 31 (2014) (originally note 188).

¹⁸ The Appendix reviews various attempted and enacted bans on different classes of weaponry and types of warfare, including crossbows, aerial bombardment, submarines, nuclear weapons, cluster munitions, anti-personnel landmines, biological weapons, chemical weapons, and permanently blinding lasers (originally note 189).

A. Qualities of a Successful Ban

Every weapon ban success story is the product of a unique combination of factors, including the weapon's inherent traits, its recent usage, prevailing moral and ethical concerns, and the status and interests of concerned states. It is tempting—and not inaccurate—to conclude that each ban is *sui generis*, the product of hard work and happy coincidence.

That being said, successful bans do seem to have certain qualities in common, and failed attempts to ban a class of weapon share a number of contrary characteristics. At the risk of oversimplifying the complex interactions that have resulted in successful bans, this Section attempts to tease out some common qualities.¹⁹

1. Weapons Causing Superfluous Injury or Unnecessary Suffering

The Saint Petersburg Declaration of 1868, which banned explosive bullets and provided incentive for the eventual banning of expanding bullets, proclaimed that the object of warfare “would be exceeded by the employment of arms which uselessly aggravate the sufferings of disabled men, or render their death inevitable” and that the use of such weapons would therefore “be contrary to the laws of humanity.”²⁰ This prohibition has since evolved into a customary rule of armed conflict, reiterated in numerous treaties.²¹ Technically, because their use is already forbidden, there should be no need to explicitly ban weapons causing superfluous injuries—they are already *per se* unlawful.

In practice, states naturally disagree about which weapons qualify as *per se* unlawful on these grounds, and advocates of nearly all modern attempted and enacted bans have argued that the weapon in question causes superfluous injury or unnecessary suffering. Accordingly, while claims that specific weapons cause superfluous injuries have undergirded weapon bans since the 1868 Declaration, an argument that a given weapon is *per se* unlawful on this ground does not

¹⁹ This list is not intended to be exhaustive, nor does it discuss the likely relative weight of the different factors (originally note 190).

²⁰ Declaration Renouncing the Use, in Time of War, of Certain Explosive Projectiles, Nov. 29–Dec. 11, 1868, 18 Martens Nouveau Recueil (ser. 1) 474, 138 Consol. T.S. 297 (originally note 191).

²¹ [Rule 70. Weapons of a Nature to Cause Superfluous Injury or Unnecessary Suffering](#), Int'l Committee Red Cross (last visited Apr. 27, 2015) (originally note 192).

significantly increase the likelihood of a ban being concluded or of its ultimate effectiveness.

That being said, if injuries caused by a weapon inspire widespread public concern and civil society engagement, they may indirectly contribute to the likelihood of a ban's success. Additionally, a ban might clarify whether or not there is consensus as to whether the use of a given weapon generally causes superfluous injury or unnecessary suffering.

2. Inherently Indiscriminate Weapons

Because indiscriminate attacks are prohibited, inherently indiscriminate weapons—which includes weapons that cannot be directed at a military objective and weapons whose effects cannot be controlled—are per se unlawful. This is both a practical and humanitarian prohibition: to the extent a weapon's effects cannot be directed or controlled, it threatens both the users' troops and nearby civilians.

Ban advocates regularly argue that the weapon in question is inherently indiscriminate. But, like weapons causing superfluous injuries, states can disagree in good faith on this question. For example, many advocates of the Mine Ban Convention argue that anti-personnel landmines are per se unlawfully indiscriminate;²² the United States has responded that "smart" mines, which could destruct or deactivate automatically after a certain period of time or at the end of active hostilities, would not be.²³ Nor will near-universal recognition that a weapon is inherently indiscriminate ensure that states will be able to conclude a treaty ban, as evidenced by nuclear weapons.²⁴

²² See, e.g., [Why the Ban](#), Int'l Campaign to Ban Landmines (last visited Apr. 27, 2015) (originally note 194).

²³ See Emily Alpert, [Why Hasn't the U.S. Signed an International Ban on Landmines?](#), L.A. Times Blog (Apr. 5, 2012, 4:45 AM) (originally note 195).

²⁴ [Ed. note: In 2021, the Treaty on the Prohibition of Nuclear Weapons entered into force. Treaty regimes are consent-based. For those states that are party to it, the treaty prohibits the development, testing, production, stockpiling, stationing, transfer, use and threat of use of nuclear weapons, as well as assistance and encouragement to the prohibited activities. As of 2026, none of the states with nuclear capabilities have joined; a few—including nuclear powers like the United States and non-nuclear powers that benefit from the nuclear security umbrella like South Korea—have expressed outright opposition on the grounds that it would undermine non-proliferation agreements and established deterrence regimes.]

It is worth noting that, just by existing, a ban may contribute to its own effectiveness. In situations where a weapon is not universally recognized as indiscriminate at the time a ban is concluded, the passage of the ban itself may lead to the recognition of the weapon's indiscriminate nature. The Mine Ban Convention is largely credited with the increasing stigmatization of anti-personnel landmines,²⁵ and advocates of the Convention on Cluster Munitions hope that it will have a similar effect.²⁶ Similarly, states' military manuals now tend to describe biological weapons as prohibited because they are indiscriminate—not because their use is forbidden by treaty law.²⁷

3. Ineffective Weapons

Discussing the Hague Conference's bans of various forms of weaponry, one scholar reflected:

Such destructive weapons, for instance, as the high explosive shell, the shrapnel, mines or torpedoes, were retained as legitimate means of warfare, whereas the inefficient expanding and explosive bullets were condemned along with the perfectly useless free balloons. The proceedings of the Hague Conference demonstrate rather that a weapon will be restricted in inverse proportion, more or less, to its effectiveness; that the more efficient a weapon or method of warfare the less likelihood there is of its being restricted in action by the rules of war.²⁸

This realist analysis has been borne out to some degree in recent practice: many credit the relative success of the biological and chemical weapon bans with the

²⁵ See, e.g., Rachel Good, Note & Comment, *Yes We Should: Why the U.S. Should Change Its Policy Toward the 1997 Mine Ban Treaty*, 9 NW. J. Int'l Hum. Rts. 209, 210 (2011) (originally note 196).

²⁶ See, e.g., Human Rights Watch, [Twelve Facts and Fallacies About the Convention on Cluster Munitions](#) 6 (2009) (originally note 197).

²⁷ See [Practice Relating to Rule 71. Weapons That Are by Nature Indiscriminate](#), Int'l Committee Red Cross (last visited Apr. 27, 2015) (citing sources) (originally note 198).

²⁸ M.W. Royse, *Aerial Bombardment and the International Regulation of Warfare* 131–32 (1928); see also Bonnie Docherty, *The Time Is Now: A Historical Argument for a Cluster Munitions Convention*, 20 Harv. Hum. Rts. J. 53, 59–61 (2007) (originally note 199).

fact that they are weapons with little military utility for powerful states, especially relative to newer weaponry.²⁹

Equally, once a weapon has demonstrated its utility, it will be difficult to convince states to cease using it. Most failed bans share two crucial characteristics in common, one of which is that they attempted to prohibit the use of extremely effective weapons.

However, the conclusion that the law will always bow to effective weapons³⁰—which is to say, that effective weapons can never be banned—is not entirely accurate. Chemical weapons have proven their utility in many situations, yet the chemical weapons ban is one of the more effective ones in existence today.³¹ The Mine Ban Convention is another possible exception to the general rule: despite its many arguments that anti-personnel mines have military value, the United States has now committed to ceasing to produce or acquire banned mines in the future with an eye toward compliance with the Convention.³²

4. Other Means Exist for Accomplishing the Same Military Objective

Certain weapons may efficiently accomplish a military objective: permanently blinding lasers, for example, can be used to disable enemy troops. However, when a state may use another weapon—say, temporarily blinding lasers—to

²⁹ See, e.g., Mark J. Osiel, *Obeying Orders: Atrocity, Military Discipline, and the Law of War*, 86 Calif. L. Rev. 939, 992 n.199 (1998). The United States unilaterally renounced its biological weapons research after concluding they were of limited military effectiveness. See Docherty, *supra*, at 60 (originally note 200).

³⁰ See Hays Parks, *Submarine-Launched Cruise Missiles and International Law: A Response*, 103 U.S. Naval Inst. Proc. 120, 120 (1977) (“In the consideration of any new weapon, technology must yield to the law, or the law to technological change. Where a weapon has proved effective, the latter usually has occurred.”) (originally note 201).

³¹ See Richard M. Price, *The Chemical Weapons Taboo* 5 (1997); Jozef Goldblat, *The Biological Weapons Convention—An Overview*, 318 Int’l Rev. Red Cross 251, 264 (1997) (describing chemical weapons as “predictable, capable of producing immediate effects and, consequently, useful in combat”) (originally note 202).

³² Press Release, Office of the Press Sec’y, The White House, [Fact Sheet: Changes to U.S. AntiPersonnel Landmine Policy](#) (June 27, 2014). *But see* Docherty, *supra*, at 60 (arguing that the success of the Mine Ban Convention is due in part to the fact that anti-personnel landmines have declining military utility) (originally note 203).

accomplish the same military objective by other means, it will be more willing to relinquish the other one.³³

Again, the converse is also true: states will be less willing to voluntarily surrender a weapon which provides the only means of accomplishing certain goals. The second crucial characteristic most failed weapon bans have in common is that they attempted to ban weapons which, at least at the time, were unique in their ability to wreak a certain type of devastation or accomplish certain goals.

5. Clear and Narrowly Tailored Prohibitions

Bans are more likely to be effective if they clearly describe what weapons are and are not permitted. The Chemical Weapons Convention, for example, has an extensive Annex on Chemicals delineating why and what specific agents are prohibited.³⁴ Clear descriptions increase the effectiveness of a ban at three critical points in time. First, states want to understand precisely what capabilities they are foregoing at the time of treaty ratification. Second, because states know precisely what is forbidden, they will be less likely to use or invest research monies in developing related weapons. Finally, in most cases, other states will be able to easily identify violations of the ban.³⁵ For treaty regimes with enforcement mechanisms, this will make formal enforcement easier; for those without, states and civil society can still engage in informal enforcement actions, such as naming and shaming.

Furthermore, many bans owe their existence and success to the fact that they are narrowly tailored. The prohibition on blinding laser weapons does not prohibit

³³ See Nicholas Marse, Peace Research Inst. Oslo, *Defining the Scope of Autonomy: Issues for the Campaign to Stop Killer Robots 2* (2014); cf. Allyson Hauptman, *Autonomous Weapons and the Law of Armed Conflict*, 218 Mil. L. Rev. 170, 192 (2013). ("Laws that focus more on methodology, rather than outcomes, are more likely to gain adherence because state will not feel as if their legitimate options for attaining military victory have been prohibited.") (originally note 204).

³⁴ Convention on the Prohibition of the Development, Production, Stockpiling and Use of Chemical Weapons and on Their Destruction, Annex on Chemicals, Jan. 13, 1993, S. Treaty Doc. No. 103-21, 1974 U.N.T.S. 45 (originally note 205).

³⁵ There is reason to believe, however, that this generalization will not hold true for autonomous weapon systems, no matter how precisely they are defined (originally note 206).

the development or use of *temporarily* blinding weapons, such as dazzlers.³⁶ Additionally, it only forbids the use of weapons where the primary purpose is to cause blinding—it explicitly does not cover blinding which occurs “as an incidental or collateral effect of the legitimate military employment of laser systems, including laser systems used against optical equipment.”³⁷ The Biological Weapons Convention excepts agents that have “justification for prophylactic, protective or other peaceful purposes,”³⁸ and the Mine Ban Convention excludes anti-tank mines equipped with anti-handling devices from its definition of “anti-personnel mines.”³⁹

Broader, vague bans have had far less success. The 1907 Hague Declaration’s prohibition of “the launching of projectiles and explosives from balloons” may have enjoyed greater ratification rates had it not also banned similar launchings “by other new methods of a similar nature.”⁴⁰ Additionally, the United States—originally one of the key initiators of a ban on anti-personnel landmines—may have joined the Mine Ban Convention had the other state parties been willing to make certain exceptions that restricted its application.⁴¹

6. Prior Regulation

The existence of a prior treaty or customary regulation, either on the same form of weaponry or on a similar type or class of weapons, may encourage states to conclude and abide by a later, stronger treaty ban. The success of the chemical weapons ban has been traced to the prohibitions on the use of poison or

³⁶ See Protocol on Blinding Laser Weapons (Protocol IV) art. 1, Oct. 13, 1995, S. Treaty Doc. No. 105-1, 2024 U.N.T.S. 167. In 2006, the United States employed laser dazzlers in Iraq to temporarily incapacitate drivers who ignored warnings at checkpoints. James Rainey, [A Safer Weapon, With Risks](#), L.A. Times, May 18, 2006 (originally note 207).

³⁷ Protocol IV, *supra*, art. 3 (originally note 208).

³⁸ Convention on the Prohibition of the Development, Production and Stockpiling of Bacteriological (Biological) and Toxin Weapons and on Their Destruction art. I, Apr. 10, 1972, 26 U.S.T. 583, 1015 U.N.T.S. 163 (originally note 209).

³⁹ Convention on the Prohibition of the Use, Stockpiling, Production and Transfer of AntiPersonnel Mines and on Their Destruction art. 2, Sept. 18, 1997, 2056 U.N.T.S. 211 (defining an “anti-handling device” as a device “which activates when an attempt is made to tamper with or otherwise intentionally disturb the mine”) (originally note 210).

⁴⁰ Declaration (XIV) Prohibiting the Discharge of Projectiles and Explosives from Balloons, Oct. 18, 1907, 36 Stat. 2439, 1 Bevans 739 (originally note 211).

⁴¹ See David E. Sanger, [U.S., in Shift, Says It May Sign Treaty to Ban Land Mines](#), N.Y. Times, Sept. 15, 1997 (originally note 212).

poisoned weapons,⁴² and the relatively effective anti-personnel landmine ban was preceded by a regulatory treaty.

If, however, new technological developments have increased the military utility of a weapon, the existence of a prior, related treaty will not be decisive. States were relatively willing to ban aerial bombardment by balloon; they were far less willing to ban aerial bombardment after the invention of the airplane.

7. Public Concern and Civil Society Engagement

The existence and growing success of the Mine Ban Convention and the Convention on Cluster Munitions has been widely attributed to the active participation of nongovernmental organizations and other civil society representatives, both prior to, during, and after the treaty negotiations.⁴³ Harold Koh argues that these “transnational norm entrepreneurs” can play an important role in altering states’ interests.⁴⁴ Some have suggested that such actors will likely play an increasingly influential role in the future.⁴⁵

That being said, this factor is far from decisive. Despite longstanding crusades for a ban on nuclear weapons by a plethora of prominent individuals and international and domestic organizations—including direct action organizations, environmental groups, consumer protection groups, professional associations, and political and religious organizations—a ban on nuclear weapons still seems out of reach.⁴⁶

8. Sufficient State Commitment

All of the aforementioned characteristics will help determine whether a state will be willing to voluntarily renounce its right to use an otherwise lawful weapon.

⁴² See Price, *supra*, at 15 (originally note 213).

⁴³ See, e.g., Richard Falk & Andrew Strauss, *On the Creation of a Global Peoples Assembly: Legitimacy and the Power of Popular Sovereignty*, 36 Stan. J. Int’l L. 191, 199–201 (2000) (originally note 214).

⁴⁴ Harold Hongju Koh, Address, *The 1998 Frankel Lecture: Bringing International Law Home*, 35 Hous. L. Rev. 623, 656–63 (1998) (discussing the role of these norm entrepreneurs with regard to the Mine Ban Treaty) (originally note 215).

⁴⁵ See Eric Talbot Jensen, *The Future of the Law of Armed Conflict: Ostriches, Butterflies, and Nanobots*, 35 Mich. J. Int’l L. 253, 282 (2014) (originally note 216).

⁴⁶ [See Editors note above.]

And, ultimately, states' commitment is the best indicator of a weapon bans' success: without it, no ban can be concluded, let alone be effective.

But state commitment to a ban is not purely a product of, say, whether the weapon is uniquely effective or abhorrent to the general public. It may also depend on the number of other states willing to renounce the weapon. Thus, there may be a snowball effect: as more states [publicly] declare their support for a ban, more states may be willing to support it. The status of supporting states is also relevant to a ban's success. If a treaty ban is ratified by the vast majority of states in the international community, but not by states that produce or use the weapon in question, it would be difficult to argue that the ban is successful.⁴⁷ Finally, a ban with a small number of initial state parties could potentially become effective over time, if its language evolves into customary international law and thereby becomes binding on all states. However, without widespread subscription or buy-in from the most relevant states, a treaty is unlikely to contribute significantly to the development of relevant customary norms.

B. Implications for Autonomous Weapon Systems

The prior section identified eight traits often associated with successful bans. Of these, only one characteristic—civil society engagement—suggests that a ban on autonomous weapon systems would be enacted or effective; the others are inconclusive or currently weigh against the likelihood of a successful ban.

Acknowledging that autonomous weapon systems have already been integrated into states' armed forces affects the analysis of many of these factors. First, there is evidence that autonomous weapon systems can be highly effective in certain circumstances, such as where faster-than-human reaction times are necessary. Second, many of the objectives accomplished by autonomous weapon systems

⁴⁷ See Reeves & Johnson, *supra*, at 29 n.55 (originally note 217).

could not be similarly achieved by other means.⁴⁸ Because autonomous weapon systems now in use are uniquely effective, it will be difficult to convince states to voluntarily relinquish them.⁴⁹ Finally, there has not been any great public outcry over the current deployment of autonomous weapon systems.

Two aspects of the clarity factor undermine the likelihood of a successful ban on autonomous weapon systems. First, because the possibilities for their use is far from understood, let alone exhausted, states are unlikely to agree to prohibit them. Second, even were a ban enacted, autonomous weapon systems are unusual in that it may not be possible to determine if and when the ban is violated. As evidenced by conflicting reports on the Korean SGR-A1, it is difficult for outside observers to determine the autonomous capabilities of any given system. As a result, only those deploying a weapon system will know if it is semi-autonomous or autonomous, supervised or unsupervised.

⁴⁸ Because the ban on permanently blinding lasers does not prohibit the use of lasers generally or even the use of temporarily blinding lasers, it does not prevent states from accomplishing similar military objectives. Therefore, contrary to many assertions, this ban provides little precedential value for a ban of autonomous weapon systems. See Michael N. Schmitt & Jeffrey S. Thurnher, “*Out of the Loop*”: *Autonomous Weapon Systems and the Law of Armed Conflict*, 4 Harv. Nat’l Security J. 231, 234, 281 n.158 (2013); see also Hauptman, *supra*, at 195 n.89. *Contra* Mines Action Can. Memorandum, *supra* (discussing the relevance of the ban on permanently blinding lasers); see also Mary Ellen O’Connell, *Banning Autonomous Killing: The Legal and Ethical Requirement that Humans Make Near-Time Lethal Decisions*, in *The American Way of Bombing: Changing Ethical and Legal Norms, From Flying Fortresses to Drones* 224 (Matthew Evangelista & Henry Shue eds., 2014); [Q&A on Fully Autonomous Weapons](#), Hum. Rts. Watch (Oct. 21, 2013). Additionally, while they had not yet been used in combat, the effects of blinding lasers were fairly predictable, see Tyler D. Evans, *Note, At War with the Robots: Autonomous Weapon Systems and the Martens Clause*, 41 Hofstra L. Rev. 697, 721–22, 731 (2013); and unnecessarily harmful, see *Advancing the Debate*, *supra*, at 17—unlike the potential effects of as-yet-undeveloped autonomous weapon systems (originally note 218).

⁴⁹ Cf. Marsh, *supra*, at 4 (“At the very least, if the Brimstone and other ‘fire and forget’ missiles are assumed to fall under the definition used by the Campaign, then consider effort may need to be expended on explaining why they are not to be considered militarily essential.”); see also *Advancing the Debate*, *supra*, at 18 (noting that, once autonomous weapon systems technology is developed, “many countries would be reluctant to give it up, especially if their competitors were deploying it”) (originally note 219).

While states do seem interested in how autonomous weapon systems might be regulated,⁵⁰ there is hardly sufficient state commitment at present to suggest a ban would be successful. At the May 2014 Experts Meeting, only a few states called for an outright ban.⁵¹ Nor is there any evidence that the military powerhouses currently investing in autonomous weapon research and development would join such a treaty. In fact, the United Kingdom opposes any such ban.⁵²

Other factors do not weigh heavily one way or the other. There is nothing intrinsic to autonomous weapon systems that would cause superfluous injury or unnecessary suffering; a bullet fired by an autonomous sentry robot causes the same amount of injury as one fired by a human sentry.⁵³ Certainly some autonomous weapon systems could be designed to cause such harm—just as certain bullets could be designed to cause excessive damage—but that does not mean the class of weapons as a whole should be banned. And, as discussed above, even assuming that autonomous weapons cannot discriminate between lawful and unlawful targets does not make them per se unlawfully indiscriminate. As evidenced by autonomous weapon systems in use today, they can be employed in a discriminating manner and with controlled effects. Nor is there any directly relevant prior regulation.

The only characteristic which seems to increase the likelihood of a successful ban is the strong civil society campaign, comprised of a growing coalition of legal scholars, philosophers, human rights activists, scientists, Nobel laureates, and

⁵⁰ See Sarah Knuckey, [Scientists from 37 Countries Call for Ban on Autonomous Lethal Targeting](#), Just Security (Oct. 16, 2013) (noting that, at the May 2014 Experts Meeting, all thirty states making opening statements welcomed the conversation, many states took an active role in discussions, and no state argued that it was the wrong time or forum to be addressing these issues—instead, some states proposed that the issue be considered additionally in other forums) (originally note 220).

⁵¹ See Bonnie Docherty, [Taking on “Killer Robots.”](#) Just Security (May 23, 2014) (reporting that five states have called for a ban); Knuckey, *supra* (stating that Ecuador, Egypt, and Pakistan called for a ban on fully autonomous weapon systems) (originally note 221).

⁵² Owen Bowcott, [UK Opposes International Ban on Developing ‘Killer Robots.’](#) The Guardian, Apr. 13, 2015 (originally note 222).

⁵³ See George R. Lucas, Jr., *Automated Warfare*, 25 Stan. L. & Pol’y Rev. 317, 330 n.24 (2014). *Contra* Mark Gubrud, [The Principle of Humanity in Conflict](#), ICRC (Nov. 19, 2012) (suggesting that there “is a human right not to be killed on the decision of machines”); Wendell Wallach, [Terminating the Terminator: What to Do About Autonomous Weapons](#), Sci. Progress (Jan. 29, 2013) (arguing that permitting machines to make life-or-death decisions is mala in se—evil in itself—like the use of rape or biological weaponry) (originally note 223).

diverse organizations.⁵⁴ Due to their efforts, “killer robots” are no longer perceived as the stuff of science fiction; rather, they are the subject of Special Rapporteur reports,⁵⁵ international meetings,⁵⁶ and U.S. congressional briefings.⁵⁷ These norm entrepreneurs are working to frame the debate on autonomous weapon systems to increase the likelihood of a successful ban.⁵⁸ Whether this will be sufficient, especially when balanced against the other factors, has yet to be seen.

Perhaps the most critical determinant of an autonomous weapon system ban’s success will be how narrowly tailored the chosen definition is. If the official threshold for autonomy in a weapon system is set excessively high, it will exclude weapon systems with autonomous capabilities currently in use today or likely to be developed in the near future, which would increase the likelihood of state subscription. However, this would grant states wide latitude in using weapon systems capable of independently selecting and engaging targets, which undermines the original impetus for a ban. Additionally, as Nicolas Marsh points out, there is a paradoxical effect to a definition of autonomy that would require machines to engage in human-level reasoning: “A key motivation for the Campaign [to Stop Killer Robots] is that robots cannot make the ethical and contextual assessments that humans can. However, a robot with a decisionmaking capability as advanced as human cognition could presumably make such judgments.”⁵⁹ Alternatively, if the threshold for autonomy in weapon systems is set too low, states will be extremely unlikely to ratify the treaty, as they will be unwilling to give up what they perceive as necessary technology.

Finally, regardless of how autonomous weapon systems are defined, a ban might exclude certain types or uses no one seems to find objectionable, like

⁵⁴ See Advancing the Debate, *supra* (originally note 225).

⁵⁵ Special Rapporteur on Extrajudicial, Summary or Arbitrary Executions, *Report of the Special Rapporteur on Extrajudicial, Summary or Arbitrary Executions*, 20–21, Human Rights Council, U.N. Doc. A/HRC/23/47 (Apr. 9, 2013) (originally note 226).

⁵⁶ Jean-Hugues Simon-Michel, [Report of the 2014 Informal Meeting of Experts on Lethal Autonomous Weapons Systems \(LAWS\)](#), UN Off. Geneva (May 16, 2014) (originally note 227).

⁵⁷ Tim Starks, [Killer Robots, Outer Space and Defense Spending Bill in Week Ahead](#), Roll Call (July 14, 2014), (originally note 228).

⁵⁸ See, e.g., Article 36, Key Areas for Debate on Autonomous Weapon Systems (2014); Advancing the Debate, *supra*; Human Rights Watch & Int’l Human Rights Clinic, Harvard Law Sch., [Losing Humanity: The Case Against Killer Robots](#) 46 (2012); Marsh, *supra*; Mines Action Can. Memorandum, *supra* (originally note 229).

⁵⁹ Marsh, *supra*, at 3 (originally note 231).

autonomous weapon systems now employed for purely defensive purposes or “smart” weapons which better distinguish between lawful and unlawful targets.⁶⁰ While such exclusions will likely increase ratification rates, they would also undermine many ban advocates’ goals, such as limiting military investment in weapon autonomy and preventing the proliferation of autonomous weapon systems.

Developments in weapons technology often challenge basic precepts of the existing law governing armed conflicts. The crossbow changed the common understanding of who could be a valuable soldier; submarines, airplanes, and cyberattacks expanded the scope of potential battlefields; nuclear weapons raised the possibility of total war and even human-caused human extinction. Autonomous weapon systems similarly raise important questions as to the necessary level of human involvement in decisions to use lethal force, leading many to call for their complete prohibition.

But autonomous weapon systems share few qualities with weapons successfully banned in the past. They are not inherently indiscriminate, and they can be designed such that they do not cause superfluous injury or unnecessary suffering. Moreover, as evidenced by autonomous weapon systems in use today, they can be used lawfully, effectively, and carry out tasks that could not be otherwise achieved. As a result, states are unlikely to conclude—let alone comply with—a treaty banning their use, unless the ban is so narrowly tailored that it effectively defines autonomous weapon systems out of existence.

All in all, this analysis suggests that states and other parties interested in the governance of autonomous weapon systems should focus not on banning them, but rather on determining how best to regulate their use.

As Crootof’s piece showcases, when arguing for a ban, advocates will need to consider two questions.

⁶⁰ See Michael C. Horowitz & Paul Scharre, [Do Killer Robots Save Lives?](#), Politico (Nov. 19, 2014); Paul Scharre, [Reflections on the Chatham House Autonomy Conference](#), Lawfare (Mar. 3, 2014) (originally note 232).

First, *should* we ban a technology? Do the various considerations weigh in favor or against this most extreme version of the precautionary approach?

Second, *can* we ban that technology? In general, it is easier to regulate—and thus ban—technologies when they (1) are narrowly defined; (2) well-understood; and (3) have a regulatory bottleneck. That bottleneck may be a controllable interaction, like a point of sale. Or it may be a trait of the technology itself, like a dependence on access to particular materials (like enriched uranium) or particularized expertise (like the ability to enrich uranium).

Conversely, the more widely available a technology is, the more difficult it may be to enforce a ban.⁶¹ Additionally, the social context matters: the more useful or widely used a technology is, the more difficult it will be to enforce a ban. In short, among other considerations, a technology's architecture and affordances will affect a ban's likely effectiveness.⁶²

Accordingly, advocates for banning a technology must be prepared to argue not only that a ban is warranted, but also that it is feasible.

C. Other Implementations

Again, a precautionary approach requires not only arguing against a permissive approach, but also addressing the question of what regulation is needed. If a ban is undesirable or unfeasible, there are many other ways of implementing a precautionary approach.

Those attempting to argue against a permissive, wait-and-see approach have coalesced around the concept of the “precautionary principle.” Unfortunately, this conversation has been muddled by a lack of consensus around what the precautionary principle actually requires (which is why we instead focus on a

⁶¹ Thus, one of the traits that contributes to a technology's generativity is inversely related to that technology's regulability. See [Chapter 9](#).

⁶² It's tempting to adopt a tech-determinist view and take the design features of a technology as a given, then ask what challenges that architecture poses for regulation. But we also want to be aware of the constructivist point that human agents create technological architectures in response to social, economic, and legal constraints. The development process for the kinds of mobile autonomous weapon systems that scare many people is long, complex, and expensive—and thus could be nipped in the bud if funding is cut. Of course, co-constructivists would acknowledge the powerful incentives to keep pouring military dollars into such research and development. See [Chapter 4](#).

precautionary approach). In its weakest form, the principle merely suggests that a lack of scientific certainty regarding harms cannot be a justification for postponing regulation aimed at mitigating those harms. The strongest version would ban technologies entirely until their safety is firmly established. Alternatively, some describe the precautionary principle as a procedural tool—namely, as a burden-shifting presumption—that favors regulatory restrictions on technologies until proponents justify lifting those restrictions.

Using the “precautionary principle” language, the next two readings explore reasons to be skeptical of a precautionary regime, reasons to employ it, and ways in which such regimes might be structured. As you read, attend to the definitions being employed and who would bear the burden of arguing for the creation, satisfaction, or dissolution of regulations under the different definitions.

Excerpt From

Cass R. Sunstein, [*The Paralyzing Principle*](#), Regulation, Winter
2002-2003

All over the world, there is increasing interest in a simple idea for the regulation of risk: the Precautionary Principle. Simply put, the principle counsels that we should avoid steps that will create a risk of harm; until safety is established through clear evidence, we should be cautious. In a catchphrase: Better safe than sorry.

In ordinary life, pleas of this kind seem quite sensible. People buy smoke alarms and insurance. They wear seatbelts and motorcycle helmets, even if they are unlikely to be involved in an accident. Should rational regulators not follow the same approach as well? Many people believe so.

In many ways, the Precautionary Principle seems quite sensible, even appealing. To justify regulation, a certainty of harm should not be required; a risk, even a low one, may well be enough. It makes sense to expend resources to prevent a small chance of complete disaster; consider the high costs, pecuniary and otherwise, that are spent to reduce the risk of terrorist attack. On reasonable assumptions, the costs are worth incurring even if the probability of harm—in individual cases or even in the aggregate—is relatively low.

The Precautionary Principle might well be seen as a plea for a kind of regulatory insurance. Certainly the principle might do some real-world good, spurring us to attend to neglected problems. Nonetheless, the principle cannot be fully defended in those ways, simply because risks are on all sides of social situations. Any effort to be universally precautionary will be paralyzing, forbidding every imaginable step, including no step at all.

Definition and Appeal

The Precautionary Principle enjoys widespread international support. But what does the principle mean or require? There are numerous definitions, and they are not all compatible with one another. We can imagine a continuum of understandings. At one extreme are weak versions to which no reasonable person could object; at the other extreme are strong versions that would appear to call for a fundamental rethinking of regulatory policy.

The most cautious and weak versions suggest, quite sensibly, that a lack of decisive evidence of harm should not be a ground for refusing to regulate. Regulation might be justified even if we cannot establish an incontrovertible connection between, say, low-level exposures to certain carcinogens and adverse effects on human health. Thus, the 1992 Rio Declaration states, "Where there are threats of serious or irreversible damage, lack of full scientific certainty shall not be used as a reason for postponing cost-effective measures to prevent environmental degradation."

Strong Version

The weak versions of the Precautionary Principle are unobjectionable and important. Every day, people take steps (and incur costs) to avoid hazards that are far from certain. We do not walk in moderately dangerous areas at night; we exercise; we buy smoke detectors; we buckle our seatbelts; we might even avoid fatty foods. Because the weak versions are sensible, I will not discuss them here. Instead, I will understand the principle in a strong way, to suggest that regulation is required whenever there is a possible risk to health, safety, or the environment, even if the supporting evidence is speculative and even if the economic costs of regulation are high. To avoid palpable absurdity, the idea of "possible risk" will be understood to require a certain threshold of scientific plausibility. To support regulation, no one thinks that it is enough if someone, somewhere, urges that a

risk is worth taking seriously. But under the Precautionary Principle as I shall understand it, the threshold burden is minimal, and once it is met, there is something like a presumption in favor of stringent regulatory controls.

In 1982, the United Nations World Charter for Nature apparently gave the first international recognition to the strong version of the principle, suggesting that when “potential adverse effects are not fully understood, the activities should not proceed.” The widely publicized Wingspread Declaration, from a meeting of environmentalists in 1998, is another example of the strong version:

When an activity raises threats of harm to human health or the environment, precautionary measures should be taken even if some cause and effect relationships are not established scientifically. In this context the proponent of the activity, rather than the public, should bear the burden of proof.

Unlike the weak version of the Precautionary Principle, the strong version is not limited to threats of serious or irreversible damage and reverses the burden of proof.

Belief in the strong version of the Precautionary Principle is not limited to any particular group. All over the world, the idea has been a staple of regulatory policy for several decades. In the United States, both Congress and the federal courts, without using the term explicitly, have built in a notion of precaution in some important cases, allowing or requiring regulation on the basis of conservative assumptions. The Precautionary Principle has played a significant role in international documents, to the point where it has become ubiquitous.

Paralysis

The most serious problem with the Precautionary Principle is that it offers no guidance—not that it is wrong, but that it forbids all courses of action, including inaction. To understand that point, it will be useful to anchor the discussion in some concrete problems:

- One of the first controversies faced by the current Bush administration involved the regulation of arsenic in drinking water. There is a serious dispute over the precise level of risks posed by low levels of arsenic, but in

the “worst case” scenario, over 100 lives might be lost each year as a result of the 50 part-per-billion standard that the Clinton administration sought to revise. At the same time, the proposed 10 ppb standard would cost over \$200 million each year, and it is possible that it would save as few as six lives annually. (See “The Arsenic Controversy” special report, Fall 2001.)

- Genetic modification of food has become a widespread practice. But the risks involved are not known with precision. Some people fear that genetic modification will result in serious ecological harm and large risks to human health. Other people claim that genetic modification will have significant health benefits.
- Scientists are hardly in full accord about the dangers associated with global warming, but there is general agreement that global warming is occurring. It is possible that global warming will produce, by 2100, a mean temperature increase of 4.5 degrees C; that it will result in well over \$5 trillion in annual monetized costs; and that it will also produce a significant number of deaths from malaria. The Kyoto Protocol would require most industrialized nations to reduce greenhouse gas emissions to between 92 and 94 percent of 1990 levels in an effort to reduce the degree of warming. Such reductions would impose substantial costs.
- Many people fear nuclear power on the ground that nuclear power plants raise various health and safety issues, including some possibility of catastrophe. But if a nation does not rely on nuclear power, it is likely to rely on fossil fuels, and in particular on coal-fired power plants. Such plants create risks of their own, including risks associated with global warming. China, for example, has relied on nuclear energy as a way of reducing greenhouse gases and other air pollution problems.
- There is a possible conflict between the protection of marine mammals and military exercises. The U.S. Navy, for example, engages in many such exercises, and it is possible that marine mammals are threatened as a result. Military activities in the oceans might well cause significant harm, but a decision to suspend those activities might also endanger military preparedness.

In those cases, what guidance does the Precautionary Principle provide? It is

tempting to say that the principle calls for strong controls on arsenic, genetic engineering of food, greenhouse gases, threats to marine mammals, and nuclear power.

In all of those cases, there is a possibility of serious harms, and no authoritative scientific evidence suggests that the possibility is close to zero.

If the burden of proof is on the proponent of the activity or processes in question, the Precautionary Principle would seem to impose a burden of proof that cannot be met. Put to one side the question of whether the principle, so understood, is sensible; let us ask a more fundamental question: Is more stringent regulation really compelled by the Precautionary Principle?

Opportunity Benefits

The answer is that it is not. In most of the cases above, it should be easy to see that in its own way, stringent regulation would actually run afoul of the Precautionary Principle. The simplest reason is that such regulation might well deprive society of significant benefits, and for that reason produce risks and even deaths that would otherwise not occur. In some cases, regulation eliminates the “opportunity benefits” of a process or activity, and thus causes preventable deaths. If that is so, regulation is hardly precautionary.

The most familiar cases involve the “drug lag” produced by a highly precautionary approach to the introduction of new medicines and drugs into the market. If a government takes such an approach, it might protect people, in a precautionary way, against harms from inadequately tested drugs. But it will also prevent people from receiving potential benefits from those drugs. Is it “precautionary” to require extensive premarketing testing, or to do the opposite?

Or consider the case of genetic modification of food. Many people believe that a failure to allow genetic modification might well result in numerous deaths, and a small probability of many more. The reason is that genetic modification holds out the promise of producing food that is both cheaper and healthier, which would have large benefits in developing countries. Now the point is not that genetic modification will definitely have those benefits, or that the benefits of genetic modification outweigh the risks. The point is only that if the Precautionary Principle is taken literally, it is offended by regulation as well as by nonregulation.

Substitute Risks

Sometimes regulation would violate the Precautionary Principle because it would give rise to substitute risks in the form of hazards that materialize, or are increased, as a result of regulation. Consider nuclear power. It is reasonable to think that in light of current options, a ban on nuclear power will increase dependence on fossil fuels that contribute to global warming. If so, such a ban would seem to run afoul of the Precautionary Principle. Or consider the Environmental Protection Agency's effort to ban asbestos, a ban that might well seem justified or even compelled by the principle. The difficulty, from the standpoint of that very principle, is that substitutes for asbestos also carry risks. Or return to possible risks to marine mammals from the U.S. Navy. Some people are concerned that efforts to eliminate those risks will endanger military preparedness, if only because of the rise of new administrative barriers to training exercises. In those circumstances, what is the appropriate approach, according to the Precautionary Principle?

The problem is pervasive. In the case of arsenic, EPA has expressed concern that regulation, by virtue of its cost, will lead people to cease using local water systems and to rely on private wells, which have high levels of contamination. If that is so, or even if it might possibly be so, stringent arsenic regulation violates the Precautionary Principle just as less stringent regulation does. That is a common situation, for opportunity benefits and substitute risks are the rule, not the exception.

Regulatory Costs

A great deal of evidence suggests the possibility that an expensive regulation can have adverse effects on life and health simply by reducing income. Richer societies are healthier societies; richer individuals tend to be healthier too. If regulatory policies are expensive and lead to higher costs, less employment, and more poverty, the net effect may be to harm individual health. The empirical question is: How much money, in terms of regulatory costs, will lead to the loss of a statistical life because of the wealth-health relationship?

To be sure, both the phenomenon and the underlying mechanisms are disputed. Low-end estimates suggest that a statistical life is lost for every expenditure of \$7

million resulting from regulation; it has also been estimated that the requisite expenditure is \$50 million per statistical life; and one of the most careful studies suggests a cutoff point of \$15 million per statistical life. (See “Safety at Any Price?” Fall 2002.) A striking study suggests that poor people are especially vulnerable to this effect—that a regulation that reduces wealth for the poorest 20 percent of the population will have twice as large a mortality effect as a regulation that reduces wealth for the wealthiest 20 percent.

I do not mean to accept any particular amount here, or even to suggest that there has been an unambiguous demonstration of an association between mortality and regulatory expenditures. The only point is that reasonable people believe in that association. It follows that a multimillion dollar expenditure for “precaution” has—as a worst case scenario—significant adverse health effects, with an expenditure of \$200 million leading to perhaps as many as 30 lives lost.

If the Precautionary Principle argues against any action that carries a small risk of significant harm, then we should be reluctant to spend a lot of money to reduce risks, simply because those expenditures themselves impose risks. Here is the sense in which the Precautionary Principle, taken for all that it is worth, is paralyzing: It stands as an obstacle to regulation and nonregulation, and to everything in between.

Precautions and Rationality

But if the Precautionary Principle, taken in a strong form, is unhelpful, how can we account for its extraordinary influence, and indeed for the widespread belief that it can and should guide regulatory judgments? Undoubtedly, self-interested political actors invoke the principle strategically. For example, European farmers invoke the idea of precaution to stifle American competitors who are far more likely to rely on genetically modified crops. But apart from that point, I suggest that an understanding of human cognition provides some useful clues:

Loss Aversion. People dislike losses far more than they like corresponding gains. The result is that out-of-pocket costs, or deteriorations from the status quo, seem much worse than opportunity costs, or benefits lost as a result of continuing the status quo. In the context of risks, people often tend to focus on the losses that are associated with some activity or hazard, and to disregard the gains that might be associated with that

activity or hazard. A closely related point is that unfamiliar risks produce far more concern than familiar ones, even if the latter are statistically larger; the Precautionary Principle, in practice, is much affected by that fact.

The Myth of a Benevolent Nature. Loss aversion is often accompanied by a mistaken belief that nature is essentially benign, leading people to think that safety and health are generally at risk only or mostly as a result of human intervention. A belief in the relative safety of nature and the relative risk of new technologies often informs the Precautionary Principle. Because natural processes are often dangerous and human interventions often promote safety, a commitment to nature can be life threatening.

The Availability Heuristic. It is well known that people focus on some risks simply because they are cognitively “available,” whereas other risks are not. When the Precautionary Principle seems to require stringent controls on one risk, even though other risks are in the vicinity, the availability heuristic is a common reason. And when the availability heuristic is at work, certain hazards will stand out whether or not they are not statistically large. The hazards associated with heat waves, for example, receive little public attention, while the hazards associated with air travel are a significant source of public concern; one reason is that the latter hazards come readily to mind. That is a serious problem because the less salient risks, including those from heat waves and poor diet, can be the serious ones.

Probability Neglect. People are sometimes prone to neglect the probability that a bad outcome will occur; they focus instead on the outcome itself. The Precautionary Principle often embodies a form of probability neglect. At least, that is the case when people invoke the principle to favor stringent controls on a low-probability risk and when the consequence of those very controls is to give rise to new risks of equal or greater probability. In the context of the sniper attacks in the Washington, D.C. area in October 2002, people were far more concerned, and took many more precautions, than the statistical realities warranted, in part because the high salience of the attacks led to a form of probability neglect. It is highly likely that some of those precautions, including those that involved extra driving, actually increased people’s risks.

System Neglect. The Precautionary Principle often reflects a general neglect

of the systemic effects of regulation. When a single problem is placed in view, it can be difficult to see the full consequences of legal interventions. Sometimes, the principle has the appearance of being workable only because a subset of the relevant effects is “on screen” — and hence there seems to be no need to take precautions against other possible adverse effects that do not register. I suggest that the Precautionary Principle seems appealing to many people in large part for the same reason.

Rejoinders, Adjustments, and Salutary Goals

Is there anything that might be said, at this stage, by proponents of the Precautionary Principle? There are several possibilities:

The Weak Version and Context

It might be tempting to revert to the weak version of the principle—a version that is entirely unobjectionable. Alternatively, it might be urged that in many cases in which the principle is invoked, the risk at issue is the one that deserves the most sustained attention. In the context of global warming, for example, the Precautionary Principle might be triggered on the (controversial) ground that the potential risks of warming are far greater than the risks associated with the reduction of greenhouse gases. But that step points toward a sensible and substantial refashioning of the principle, one that ensures that low-probability catastrophes are given careful attention, and that the various risks at issue will be weighed and balanced in accordance with the facts.

Biases

Advocates of the Precautionary Principle might urge that environmental values are systematically disregarded in the regulatory process or not given their due, and hence that the principle helps counteract systematic biases. A particular problem here is myopia: Perhaps government officials, uninformed by the principle, would fail to attend to risks that will not occur, or be seen to occur, in the short-run. Another problem is that some people tend to be unrealistically optimistic. As a result, many low-level risks do not register at all. When people think that they are “safe” even though they face a statistical risk, they might well be responding to emotions, seeking to avoid the anxiety that comes from an understanding of the inevitability of (some) risk.

On that view, the principle can be defended pragmatically, if not theoretically, as a way of emphasizing the importance of attending to issues, especially environmental issues, that might otherwise be neglected. In some settings, the pragmatic defense is undoubtedly plausible, and the Precautionary Principle, applied narrowly, undoubtedly leads to some good results.

But two problems remain. The first is that environmental values are sometimes on both sides; consider the nuclear power controversy. The same is certainly true of health and safety, as shown by the case of extensive premarket testing of pharmaceuticals. The second is that, even when environmental values are on only one side, the interests and values on the other side might well be at a comparative disadvantage too; consider the potential beneficiaries of genetic modification of food.

Distribution

It might be tempting to defend the Precautionary Principle on distributional grounds. The Clean Air Act takes a precautionary approach, requiring an “adequate margin of safety” and hence regulation in the face of scientific uncertainty. At the same time, the Clean Air Act appears to be giving disproportionate benefits to poor people and members of minority groups.

Aggressive action to combat climate change might be more beneficial to poor countries than to wealthy ones. That is partly because wealthy countries are better able to adapt; it is partly because agriculture (potentially vulnerable to climate change) is responsible for a small percentage of the economy of wealthy nations but a large percentage of the economy of poor nations. It is also because one of the most serious health risks posed by climate change consists of an increased incidence of malaria, a nonproblem for wealthy countries. In the context of global warming, at least, the Precautionary Principle might be invoked to prevent especially severe burdens on those in the worst position to bear them.

Of course, it makes sense to be concerned with the distribution of domestic or international risks. The problem of global warming owes its origin to the actions of wealthy nations, and hence it makes special sense to ask those nations to bear a disproportionate cost of correction if poor nations are likely to be hit hardest. The distributional effects of global warming are among the strongest points in

favor of aggressive regulation of greenhouse gases – at least if (and it is a big if) that regulation does not hit poor countries economically. But in too many cases, the Precautionary Principle, as applied, would have unfortunate distributional effects. The case of genetic modification of food is an example; here the benefits are likely to be enjoyed by poor people, not the wealthy.

The case of DDT is similar. While a ban on DDT, supported by reference to the Precautionary Principle, is probably justified in wealthy nations, such a ban is likely to have bad effects in at least some poor countries where DDT is the cheapest and most effective way of combating serious diseases, most notably malaria. Distributional issues should indeed be a part of a system of risk regulation, but the Precautionary Principle is a crude, indirect, and sometimes perverse way of incorporating distributional concerns.

Risk v. Uncertainty

A more subtle point involves situations of risk (where probabilities can be assigned to various outcomes) rather than uncertainty (where no such probabilities can be assigned). In a situation of uncertainty, when existing knowledge does not permit regulators to assign probabilities to outcomes, some argue that people should follow the “Maximin” Principle: Choose the policy with the best worst-case outcome.

Perhaps the Precautionary Principle, as applied, is a form of the Maximin Principle. In this form, the principle asks officials to identify the worst case among the various options, and to select the option that has the least-bad worst case. Perhaps the Maximin Principle would support some proposed applications of the Precautionary Principle by, for example, urging aggressive steps to combat terrorism or global warming.

The problem is that the Precautionary Principle is not the Maximin Principle. If the latter principle is what is meant, then we should be discussing that principle directly, and evaluating it against the alternatives. The Precautionary Principle obscures those issues. It does so because it is invoked across-the-board, in situations of risk as well as uncertainty.

Salutary Goals

We are now in a position to appreciate some of the goals of those who invoke the Precautionary Principle. In the context of tobacco, for example, a serious public health movement was muted simply by virtue of some scientific doubt – even though reasonable people take steps to reduce likelihoods, not only certainties, of adverse effects. The Precautionary Principle can be taken as a reminder not to require proof. To the extent that the principle emphasizes obligations to the future, it is entirely salutary. Those who invoke the principle undoubtedly are motivated, much of the time, by the goal of protecting the most vulnerable people against risks to their safety and health.

Nothing I have said is meant to draw those goals into doubt. My claim is that the Precautionary Principle is not a helpful way of promoting the relevant goals – and that if it is taken seriously, it is paralyzing and not helpful at all.

Conclusion

In this article, I have argued not that the Precautionary Principle leads in the wrong directions, but that if it is taken for all that it is worth, it leads in no direction at all. The reason is that risks of one kind or another are on all sides of regulatory choices, and it is therefore impossible, in most real-world cases, to avoid running afoul of the principle. Frequently, risk regulation creates a (speculative) risk from substitute risks or from foregone risk-reduction opportunities. And because of the (speculative) mortality and morbidity effects of costly regulation, any regulation—if it is costly—threatens to run afoul of the Precautionary Principle. We have seen that both regulation and nonregulation seem to be forbidden in cases involving nuclear power, arsenic, global warming, and genetic modification of food. The Precautionary Principle seems to offer guidance only because people blind themselves to certain aspects of the risk situation, focusing on a mere subset of the hazards that are at stake.

To some extent, those who endorse the principle are responding to salutary political or moral motivations that it might be thought to embody. Well-organized private groups sometimes demand conclusive proof of harm as a precondition for regulation; the demand should be resisted because a probability of harm is, under many circumstances, a sufficient reason to act. Both individuals and societies sometimes have a tendency to neglect the future; the Precautionary Principle might be understood as a warning against that form of neglect. There are good reasons to incorporate distributional considerations into risk regulation;

the Precautionary Principle seems, some of the time, to be a way to protect the most disadvantaged against risks of illness, accident, and death.

The problem is that the Precautionary Principle, as applied, is a crude and sometimes perverse method of promoting those various goals, not least because it might be, and has been, urged in situations in which the principle threatens to injure future generations and harm rather than help those who are most disadvantaged. A rational system of risk regulation certainly takes precautions. But it does not adopt the Precautionary Principle.⁶³

Comprehension Check

1. What is Sunstein's version of the "weak precautionary principle"? How does Sunstein's "weak precautionary principle" differ from the permissive approach?
2. What is Sunstein's version of the "strong precautionary principle"? How does Sunstein's "strong precautionary principle" differ from the permissive approach?
3. Who bears the burden of proof in rebutting the regulatory default under Sunstein's version of the strong precautionary principle? How great is that burden?
4. Sunstein observes that, "if the Precautionary Principle is taken literally, it is offended by regulation as well as by nonregulation." What does he mean by this?
5. What alternative(s) does Sunstein propose?

⁶³ [Editors' note: For further development of Sunstein's argument, see Cass R. Sunstein, [*Beyond the Precautionary Principle*](#), 151 U. Pa. L. Rev. 1003 (2003).]

Discussion Question

1. Is a cost-benefit analysis aligned with the precautionary approach? Or are there ways in which it could be described as permissive? Or is it neutral between them?

Excerpt From

Noah M. Sachs, [*Rescuing the Strong Precautionary Principle from Its Critics*](#), 2011 U. Ill. L. Rev. 1285, 1304-26 (2011)

II. Rescuing the Strong Precautionary Principle From Its Critics

The Strong Precautionary Principle has become a punching bag for many scholars of risk regulation. Cass Sunstein and other prominent scholars have unsparingly attacked it, and lacking many defenders, the Principle's reputation has been battered.

In this Part, I respond to these critiques. My rebuttal here is qualified. I fully concede that the Strong Precautionary Principle cannot be a universally applicable approach to all health and environmental dangers that society faces. But it can serve as the platform for protective risk decision making in discrete areas of law, including chemical regulation. By caricaturing the Principle as extreme and dismissing it out of hand as unworkable, the Principle's detractors have overlooked both its utility and the myriad examples where it already operates successfully in U.S. law.

A. The Critics Complaints

The critics share some common objections. They often maintain that the Principle is extreme, inflexible, anti-science, anti-growth, or anti-technology. Because Strong Precaution shifts the burden of proof on safety, critics allege that its implementation would prevent promising new technologies from getting off the ground, especially in emerging fields with a high degree of uncertainty about risks. Sunstein claims that the Strong Precautionary Principle would "eliminate technologies and strategies that make human lives easier, more convenient,

healthier, and longer.”⁶⁴ John Graham names the internal combustion engine, electricity, plastics, and the Internet as examples of technologies that would have been prohibited.⁶⁵ Harvey Miller and Gregory Conko further charge that “[i]f the precautionary principle had been applied decades ago to innovations such as polio vaccines and antibiotics, . . . that precaution would have come at the expense of millions of lives lost to infectious diseases.”⁶⁶

The claim that Strong Precaution is antithetical to technological progress is closely related to another view in the critical literature: Strong Precaution requires that manufacturers prove “zero risk” or “absolute safety” for an activity to proceed.⁶⁷ These, of course, are impossible standards to meet (and the Strong Precautionary Principle, as I and others have defined it, does not impose them). Other critics suggest that the Principle will be applied on the mere “conjecture”⁶⁸ or “speculation”⁶⁹ of a threat to health or the environment, suppressing important technologies without any scientific basis. In addition, because of the strong government role and the shift in the burden of proof, critics have argued that Strong Precaution is overly statist; its “guilty-until-proven-innocent” approach to addressing risk is allegedly contrary to U.S. values.⁷⁰

Cass Sunstein, the leading U.S. critic of the Strong Precautionary Principle, claims that the Principle is, ultimately, paralyzing as an approach to risk decision making.⁷¹ If the Strong Precautionary Principle presumptively requires regulation of activities that present serious threats to health, safety, or the environment, then, according to Sunstein, the Principle will paralyze decision makers because threats may be created both by regulatory inaction and by regulatory action.⁷²

⁶⁴ Cass R. Sunstein, *Laws of Fear: Beyond the Precautionary Principle* 25 (2005) (originally note 90).

⁶⁵ John D. Graham, *The Perils of the Precautionary Principle: Lessons from the American and European Experience* (Oct. 20, 2003), in *Heritage Lectures*, Jan. 15, 2004, at 3 (originally note 91).

⁶⁶ Henry I. Miller & Gregory Conko, *The Science of Biotechnology Meets the Politics of Global Regulation*, 79 *Issues in Sci. & Tech.* Online (2000) (originally note 92).

⁶⁷ Lawrence A. Kogan, *The Extra-WTO Precautionary Principle: One European “Fashion” Export the United States Can Do Without*, 17 *Temp. Pol. & Civ. Rts. L. Rev.* 491, 506–07 (2008) (originally note 93).

⁶⁸ See Miller & Conko, *supra* (originally note 94).

⁶⁹ See Cass R. Sunstein, *The Precautionary Principle As a Basis for Decision Making*, 2 *Economists’ Voice* 1, 5 (2005) (originally note 95).

⁷⁰ See Ronald Bailey, *Precautionary Tale*, *Reason*, Apr. 1999, at 37; Kogan, *supra*, at 601–03 (contrasting the Precautionary Principle with the “founding principles of our society, chief among them economic and political freedom and the rule of law”); Miller & Conko, *supra* (originally note 96).

⁷¹ See Sunstein, *Laws of Fear*, *supra*, at 14–15 (originally note 97).

⁷² *Id.* (originally note 98).

This “paralysis” critique begins from the proposition that regulation itself has costs and potential risks. For instance, introduction of genetically modified crops may harm ecosystems, but harm may also occur (from increased hunger or more widespread pesticide use) if genetically modified crops are prohibited.⁷³ With respect to drug approvals, Sunstein notes that stringent Food and Drug Administration (FDA) review requirements for new drugs may block introduction of dangerous medications, but the delay involved in FDA review (so-called “drug lag”) means that life-saving drugs may not reach patients who need them.⁷⁴ Sunstein questions: “Is it ‘precautionary’ to require extensive premarketing testing, or to do the opposite?”⁷⁵ The Strong Precautionary Principle allegedly provides no guidance. According to Sunstein, the Principle “forbids all courses of action, including regulation. It bans the very steps that it requires.”⁷⁶

Given the presence of these risk-risk tradeoffs, Sunstein maintains, a principle that calls generally for caution in the face of risk, or that calls for shifting the burden of proof in response to serious threats to human health or the environment, is logically incoherent. Strong Precaution only *seems* appealing as an approach to risk because we naturally focus our attention on certain novel “target” risks (such as genetically modified foods or a chemical recently in the news, such as Bisphenol-A) and ignore both the opportunity costs of regulation and the risks presented by the status quo.⁷⁷ The way out of these decision-making

⁷³ *Id.* at 31 (originally note 99).

⁷⁴ *Id.* at 29 (originally note 100).

⁷⁵ *Id.* In another example, Sunstein notes that many chemical agents may be carcinogenic at very low levels, while others may be beneficial at very low levels even if harmful at high levels (the so-called “hormesis” effect). *Id.* at 30–31. Therefore, the “simultaneous possibility of benefits at low levels and of harms at low levels makes the Precautionary Principle paralyzing.” *Id.* (originally note 101).

⁷⁶ *Id.* at 26 (originally note 102).

⁷⁷ Sunstein argues that five different cognitive biases operate to make the Strong Precautionary Principle appear functional: (1) individuals focus on salient risks that come to mind from media reports or recent memory, and they ignore less visible risks; (2) probability neglect leads people to concentrate on worst-case outcomes, even if they are highly improbable; (3) loss aversion causes people to avoid changes from the status quo, even if the status quo poses its own risks; (4) a widespread belief that nature is benign makes human-created risks seem particularly suspect; and (5) system neglect prevents people from appreciating opportunity costs and the substitute risks that may arise when a particular technology or product is restricted. *See id.* at 36–49; Cass R. Sunstein, *Beyond the Precautionary Principle*, 151 U. Pa. L. Rev. 1003, 1011 (2003) (arguing that the Principle’s “puzzling appeal” reflects its dependence on cognitive biases) (originally note 103).

dilemmas, Sunstein maintains, is comprehensive cost-benefit analysis,⁷⁸ or, as John Graham and Jonathan Wiener advocate, “risk tradeoff analysis.”⁷⁹ Through careful quantitative analysis, critics suggest, regulators will be able to assess and compare the risks that would be created by different decision-making paths.⁸⁰ Once these risks are identified, regulators can then choose the most cost-effective risk-reducing interventions.

B. Countering the Critics

To defend use of the Strong Precautionary Principle in regulatory law, I make three principal claims. First, I demonstrate that Strong Precaution already operates successfully in U.S. law, undermining critics’ arguments that the Principle is somehow inherently unworkable or paralyzing. Second, I show that the critics’ charge of extremism is overheated. The Principle does not require prohibiting all risky activities, and in fact it provides flexibility for policy makers to determine how much risk will be tolerable or acceptable in a given area of law, as well as flexibility to determine proportionate regulatory responses. Finally, I show that the critique that Strong Precaution is “paralyzing” because of risk-risk trade-offs is overstated, especially against the backdrop of its practical implementation in U.S. law. Properly implemented, the Principle allows regulators to consider tradeoffs and alternatives. It commands neither extreme regulation nor abdication of judgment.

1. Strong Precaution in Existing Law

Critics often paint Strong Precaution as a new kid on the block, a yet-to-be-tried alternative to cost-benefit analysis, or an exotic import from Europe that has not

⁷⁸ Sunstein, *Laws of Fear*, *supra*, at 129 (cost-benefit analysis gives people “a more accurate sense of the actual harms against which protection is sought” and provides “a clearer sense of the stakes”) (originally note 104).

⁷⁹ Jonathan Baert Wiener & John D. Graham, *Resolving Risk Tradeoffs*, in *Risk Versus Risk: Tradeoffs in Protecting Health and the Environment* 226, 19–22 (John D. Graham & Jonathan Baert Wiener eds., 1995) Graham and Wiener describe risk tradeoff analysis as a three-part process that identifies tradeoffs that might result from an intervention, weighs the comparative importance of target risks and countervailing risks, and analyzes the possibility of “risk-superior” moves that might result in overall risk reduction. *Id.* (originally note 105).

⁸⁰ Sunstein, *Basis*, *supra*, at 6 (“We do not believe there is any principled way of making policy decisions without making the best possible effort to balance all the relevant costs of a policy against the benefits.”) (originally note 106).

been embraced in the United States.⁸¹ These attempts at delegitimization fail to recognize that the Strong Precautionary Principle already operates successfully in U.S. law. From Capitol Hill to state houses, legislators have frequently (and sensibly) turned to *ex ante* gatekeeping mechanisms to protect public health and the environment against serious risk.

The FDA's review process for new drugs is a prime example. Under the Federal Food, Drug, and Cosmetic Act,⁸² all substances meeting the definition of a drug⁸³ are presumptively banned from sale in the United States, unless the manufacturer produces relevant data on risks, side effects, and efficacy; conducts clinical trials; and receives affirmative FDA approval for sale.⁸⁴ In this system, the FDA stands in the exact gatekeeping role called for by the Strong Precautionary Principle. Indeed, the FDA's new drug review system can be viewed as a particularly potent form of Strong Precaution, because the precautionary measure implemented as a response to serious threats to human health from untested drugs is a complete prohibition. That prohibition, backed by criminal penalties,⁸⁵ remains in place (without any cost-benefit analysis) until the drug manufacturer can overcome the default and carry its burden of proof on safety and efficacy.⁸⁶

Belying the argument that a gatekeeping role for government is inherently anti-science or anti-technology, the United States has maintained this FDA review process for decades while also developing the most innovative and profitable pharmaceutical industry in the world. As noted above, some critics point to antibiotics as the leading example of a useful product that would have been

⁸¹ See, e.g., Lawrence A. Kogan, *The Extra-WTO Precautionary Principle: One European "Fashion" Export the United States Can Do Without*, 17 Temp. Pol. & Civ. Rts. L. Rev. 491, 493 (2008) (originally note 107).

⁸² Federal Food, Drug, and Cosmetic Act (FFDCA), 21 U.S.C. §§ 301–399 (2006). The FFDCA was enacted in 1938 and was substantially amended in 1962 to require the affirmative approval of the FDA before a new drug could be marketed. Prior to 1962, new drug applications were deemed approved unless a contrary order was issued. See *USV Pharm. Corp. v. Weinberger*, 412 U.S. 655 (1962) (originally note 108).

⁸³ See 21 U.S.C. § 321(g) (definition of "drug"); *id.* § 321(p) (definition of "new drug") (originally note 109).

⁸⁴ See *id.* § 355(a) (prohibiting the sale of new drugs without FDA approval); *id.* § 355(b)–(d) (describing the application and approval process) (originally note 110).

⁸⁵ *Id.* § 333(b) ("[A]ny person who violates [the FFDCA] by . . . knowingly selling, purchasing, or trading a drug . . . shall be imprisoned for not more than 10 years or fined not more than \$250,000, or both.") (originally note 111).

⁸⁶ See *id.* § 355 (originally note 112).

squelched if the United States had embraced Strong Precaution decades ago. But pharmaceutical manufacturers have developed life-saving antibiotics since the 1940s and will continue to do so, within a regulatory system that reflects Strong Precaution.⁸⁷ Allocation of the burden of proof to drug manufacturers has neither extinguished the market for, nor the supply of, life-saving antibiotics.

[Sachs then discusses the EPA's registration system for pesticides and the Clean Air Act as other examples of successful burden-shifting precautionary regulations.]

[...]

Critics are overlooking that the Principle can provide a workable accommodation between the needs of industry and the need to ensure harm prevention and adherence to ecological limits. Specifically, putting government in a risk gatekeeping role serves several important purposes, including:

- Ensuring that the applicant is competent to engage in the activity and has the required expertise and resources;
- Regulating the location of potentially risky activities and ensuring that they occur in places where risks to the public are minimized;
- Ensuring that activities presenting serious threats to public health or the environment can be prohibited (or have safety precautions placed on them) before harm occurs;
- Ensuring, through establishing a uniform review process for every applicant, that the cumulative amount of a risky activity does not exceed limits that would be damaging to the environment or human health; and
- Minimizing risks while further research is conducted and making that research the responsibility of firms that will benefit the most from the activity.

I am not trying to defend every permitting and licensing scheme, of course. Government permitting programs can be burdensome and prone to political

⁸⁷ It was the *unregulated* introduction of an antibiotic, causing widespread illness in the 1930s, which led to public outcry in favor of a premarket government review system. The purported antibiotic was called "Elixir Sulfanilamide," and more than one hundred people were poisoned by it in Tennessee before the FFDCA was enacted. See Richard A. Merrill, *The Architecture of Government Regulation of Medical Products*, 82 Va. L. Rev. 1753, 1761–64 (1996) (describing the origins of the FFDCA) (originally note 114).

favoritism and rent-seeking behavior. They are often complex. If inadequately funded and staffed, a governmental review may be no more than a fig leaf of risk management (witness the Deepwater Horizon Oil Spill and the lax oversight of the Minerals Management Service). But the long-standing practice in U.S. law of establishing government agencies as *ex ante* gatekeepers for risk does suggest that the Strong Precautionary Principle cannot be so easily dismissed. It is not as alien to U.S. law and values as the critics would have us believe, and it hardly seems “paralyzing” in the many contexts in which it has been applied.

2. Countering the Critique of Extremism

Rather than seriously grapple with examples of Strong Precaution in existing law, many critics instead challenge a straw man—a jumble of extremist positions and hard-to-defend results that allegedly flow from Strong Precaution. Critics frequently contend, for example, that the Strong Precautionary Principle amounts to an intolerable command for proponents of potentially harmful technologies to prove zero risk.⁸⁸ [...]

Zero risk is an undesirable and unattainable goal.⁸⁹ Regulators rarely face a choice between risk and no risk. They must choose instead among various options with different degrees of risk. There is nothing in the Strong Precautionary Principle that precludes priority setting To claim that Strong Precaution is a zero-risk standard is merely to set the straw man on his post. Zero risk is an easy target to knock down, but it misses the crux of the debate.

[...]

Most formulations of the Strong Precautionary Principle, including my own definition and that of the Wingspread Statement, offer far more flexibility than

⁸⁸ See, e.g., Giandomenico Majone, *What Price Safety? The Precautionary Principle and Its Policy Implications*, 40 J. Common Market Stud. 89, 101 (2002); Christopher D. Stone, *Is There a Precautionary Principle?*, 31 Envtl. L. Rep. 10790, 10791 (2001) (originally note 126).

⁸⁹ See *Indus. Union Dep’t, AFL-CIO v. Am. Petroleum Inst.*, 448 U.S. 607, 642 (1980) (“[S]afe’ is not the equivalent of ‘risk-free.’ There are many activities that we engage in every day—such as driving a car or even breathing city air—that entail some risk of accident or material health impairment, nevertheless, few people would consider these activities ‘unsafe.’”) (originally note 129).

this extreme portrait painted by the critics.⁹⁰ The Strong Precautionary Principle does not impose a burden on any party to prove zero risk, nor does it state that all activities that pose a possible risk must be prohibited. Rather, the Principle, while shifting the *burden of proof* away from government, deliberately leaves the relevant *standard of proof* open-ended and subject to democratic deliberation. The Principle requires that a proponent demonstrate that its product or activity meets a specified standard (predetermined through legislation) for how much risk is tolerable or acceptable in a given issue area.

This standard of proof can be set in a variety of ways, just as governing risk standards vary considerably in current public health and environmental statutes, which may prescribe specific safety technologies, set performance benchmarks, impose cost-benefit tests, or set minimal health-based standards.⁹¹ In some cases, policy makers might choose to adopt a cost-benefit test within a Strong Precautionary framework, requiring the applicant to show that the benefits of its activity outweigh potential harms. The Strong Precautionary Principle is not necessarily antithetical to cost-benefit analysis, though it has often been framed this way in the literature. Under a Strong Precautionary framework, cost-benefit analysis could still be conducted, but it would be the applicant, rather than the government, who would carry both the analytical burden of conducting the analysis and the burden of persuasion. Utilitarian welfare maximization would remain the ultimate goal of the process, but in contrast to the way cost-benefit analysis is traditionally performed, some *ex ante* precautionary restrictions would be imposed on the activity until the applicant has met its burden to prove that the activity is worth the risks.

There are many options other than cost-benefit balancing that could be used to implement the Strong Precautionary Principle in concrete legislation. In situations where both the probability of harm and the nature of the outcomes are uncertain, cost-benefit balancing is accomplished only through heroic modeling

⁹⁰ The Wingspread Statement, for example, provides that the burden of proof shall be borne by the proponent of the activity, but it does not itself supply the standard for determining when the burden of proof is met. *Id.* at 19 (originally note 134).

⁹¹ See Toxic Substances Control Act § 6(a), 15 U.S.C. § 2605(a) (2006) (stating that the EPA must prove “unreasonable risk” to restrict a chemical); Clean Air Act § 109, 42 U.S.C. § 7409(b)(1) (2006) (stating that the EPA is required to set ambient air quality standards at a level “requisite to protect the public health” with an “adequate margin of safety”) (originally note 135).

assumptions that obscure the underlying uncertainties and value choices.⁹² Cost-benefit analysis also deals awkwardly with equity concerns, including inter-generational equity. It sends a “go” signal to activities that may externalize billions of dollars in health costs on a population, as long as the related benefits in corporate profits and employment are higher than the expected costs. This Article is not meant as a comprehensive critique of cost-benefit analysis, but it does highlight why it is appropriate to establish precautionary risk management regimes that do not rest on such analysis—and not just for a narrow class of “catastrophic” risks.⁹³

Utilitarian welfare maximization need not be the driving goal of every regulatory regime, and the Strong Precautionary Principle can incorporate a wider variety of goals and values. For example, a statute might require the applicant to prove that the proposed activity will not cause substantial harm to human health, that it will incorporate best available safety or pollution control technology, or that it will not substantially harm significant ecosystems.⁹⁴ In the hazardous waste context, the Principle would not forbid a firm from engaging in long-term storage of hazardous waste nor would it require complex cost-benefit balancing for each individual storage facility. Rather, the Principle, quite sensibly, would impose a default prohibition on long-term hazardous-waste storage until the firm can demonstrate, *ex ante*, that the storage will be conducted safely—meaning in accordance with prescribed standards for construction, monitoring, financial

⁹² See, e.g., Frank Ackerman & Lisa Heinzerling, *Priceless: On Knowing the Price of Everything and the Value of Nothing* 10 (2004) (“Cost-benefit analysis of health and environmental protection rests on simplistic, implausible hypotheses”); Douglas A. Kysar, *Regulating from Nowhere: Environmental Law and the Search for Objectivity* 71–75 (2010) (contrasting situations of “risk,” which can be modeled reasonably well because outcomes and probabilities are well defined, with situations of “uncertainty” and “ignorance,” typical in environmental policy making, in which probabilities are poorly defined) (originally note 136).

⁹³ Sunstein “reconstructs” the Precautionary Principle as an Anti-Catastrophe Principle, and he argues that so reconstructed, his Anti-Catastrophe Principle “deserves to play a role in regulatory choices.” Sunstein, *Laws of Fear*, *supra*, at 114 [...] (originally note 137).

⁹⁴ That Strong Precaution may introduce values other than strict economic efficiency into the regulatory process is one of the reasons it has drawn such heated criticism. But many existing health, safety, and environmental statutes also impose requirements other than strict cost-benefit tests, such as the use of minimum safety equipment, the adoption of best-available technology, or the avoidance of significant risk. As Sunstein himself has noted, a democratic society may appropriately set regulatory standards with reference to values other than economic efficiency. See Sunstein, *Laws of Fear*, *supra*, at 129 (“Efficiency is relevant, but it is hardly the only goal of regulation. Citizens in a democratic society might well choose to protect endangered species, or wildlife, or pristine areas, even if it is not efficient for them to do so.”) (originally note 138).

security, labeling, and emergency response. The Principle is contextual and pragmatic. It allows policy makers to assess the seriousness of risks and to develop appropriate responses in a much more qualitative way than staunch advocates of a cost-benefit test would tolerate.

The governing risk standard—what, exactly, the proponent must show or prove—is not set forth in the Principle itself, but rather in implementing legislation, enacted through democratic processes. Sunstein and other critics are asking far too much from a brief principle when they criticize it for failing to answer complex questions about the valuation of human life or the distributional consequences of risk decision making.⁹⁵ The Principle was never intended for these purposes; the proper venue for resolution of these questions is implementing legislation.⁹⁶

In addition to providing flexibility to design risk standards, the Strong Precautionary Principle also affords flexibility to design default regulatory responses in the time period before the proponent has met the burden of proof. The extremist critique suggests that the only response of regulators, acting in accordance with the Strong Precautionary Principle, is to ban activities that pose possible or potential risks. But the Strong Precautionary Principle can be implemented through a variety of defaults beyond simplistic, binary (ban or

⁹⁵ See Sunstein, *Laws of Fear*, *supra*, at 51 (“[T]he Precautionary Principle is a crude, indirect, and sometimes perverse way of incorporating distributional concerns.”). *But see* John S. Applegate, *Embracing a Precautionary Approach to Climate Change*, in *Economic Thought and U.S. Climate Change Policy* 176 (David M. Driesen ed., 2010) (arguing that it is “patently not [the] purpose” of the Precautionary Principle to resolve distributional concerns) (originally note 140).

⁹⁶ The necessity of interpretation through implementation does not, however, carve out so much from the Principle itself that the Principle dissolves into vacuity. *See also* David A. Dana, *A Behavioral Economic Defense of the Precautionary Principle*, 97 NW U. L. Rev. 1315, 1317 (“Principles can express and reinforce value commitments and procedurally structure decisionmaking without dictating a single set of specific, substantive outcomes . . .”). Moreover, in existing legislation that incorporates a Strong Precautionary approach, the burden that the proponent or applicant must carry is typically substantial. There is little point in shifting the burden of proof, of course, if the ultimate burden the proponent must carry is a de minimis one. In the ex ante regulatory regimes described *supra*, the proponent bears a substantial burden to prove that risks or impacts will be within acceptable bounds—a burden that normally takes significant investment of time and resources to meet. The burden-shifting mechanism therefore does real “work” in allocating responsibilities (originally note 141).

no-ban) choices.⁹⁷ Appropriate defaults might include partial prohibitions, worker-safety precautions, locational restrictions, or warnings. Even where a regulatory regime provides for a default policy of complete prohibition (as is the case with FDA drug reviews, EPA pesticide registration, and most governmental licensing and permitting programs), the Strong Precautionary Principle clearly provides that the default can be overcome by proof of acceptable risks submitted by the applicant.

3. Countering the Critique of “Paralysis”

No one seriously disputes that regulation has opportunity costs, and no serious scholar of risk regulation believes that opportunity costs should be ignored. [...]

The presence of tradeoffs, however, does not mean that the Strong Precautionary Principle is incoherent or paralyzing. [...]

Several scholars have already examined how policy makers can address tradeoffs in a way that is consistent with the Precautionary Principle. David Dana has suggested, for example, that the Precautionary Principle helps us choose among alternative regulatory actions (including no action) by countering the cognitive tendency to avoid decisions that involve upfront costs and more distant benefits.⁹⁸ Steffen Foss Hansen and Joel Tickner have argued that true risk-risk tradeoffs are less prevalent than the critics have alleged,⁹⁹ and John Applegate has asserted that even where risks inhere on many different sides of a decision, the Precautionary Principle still plays a useful role by ruling out arguments that *all* decisions should be postponed due to scientific uncertainty.¹⁰⁰ In a thoughtful

⁹⁷ Indeed, it would be a gross waste of resources to implement the same precautionary measures for chemicals whose effects on human health cannot be determined with accuracy as for chemicals that are known carcinogens. “Precautionary measures should be taken in both cases, but with different degrees of stringency.” Sven Ove Hansson, *Can We Reverse the Burden of Proof?*, 90 Toxicology Letters 224 (1997) (originally note 143).

⁹⁸ Dana, *supra*, at 1333 (“[T]he precautionary principle may have a role to play as a corrective of cognitive biases” in situations where we face “choices between the avoidance of sure, immediate losses and the avoidances of unsure, future losses.”) (originally note 144).

⁹⁹ See Steffen Foss Hansen & Joel A. Tickner, *Putting Risk-Risk Tradeoffs in Perspective: A Response to Graham and Wiener*, 11 J. Risk Res. 475, 476 (2008) (arguing that many purported risk-risk tradeoffs are hypothetical, never emerged in practice, or were adequately addressed through precautionary regulation) (originally note 145).

¹⁰⁰ Applegate, *Climate Change*, *supra*, at 173, 183–84 (explaining that the function of the Precautionary Principle is primarily procedural and evidentiary) (originally note 146).

examination of risk-risk tradeoffs, Doug Kysar has noted that while the opportunity costs of regulation must be considered, the Precautionary Principle offers a more contextual, dynamic approach for making these tradeoffs than cost-benefit analysis.¹⁰¹ The Precautionary Principle, according to Kysar, approaches complex tradeoffs through deliberation, moral reflection, and appreciation of the limits of human cognition, while cost-benefit analysis “proceeds awkwardly in the absence of fully characterized risks” and involves questionable monetization of environmental and health harms.¹⁰²

I do not intend to revisit these arguments in detail here. Instead, I offer two responses, underdeveloped in prior literature, to the critique that Strong Precaution is paralyzing for risk decision making—responses that are particularly pertinent to the field of chemical regulation.

a. The Problem of Tradeoffs in Legislation

My first response is that the detailed, quantitative risk tradeoff analysis that Sunstein, Wiener, Graham, and other critics advocate is often difficult to implement, as a practical matter, within the legislative process. Enactment of legislation is both the primary focus of this Article and the proper focus of analysis of the Strong Precautionary Principle. The Principle should be viewed primarily as a framework to guide legislators, rather than a framework to guide regulators, because, at least in U.S. law, the shift in the burden of proof usually requires some legislative authorization.¹⁰³

In deciding whether to enact a new statute, legislators should of course attempt

¹⁰¹ Douglas A. Kysar, *It Might Have Been: Risk, Precaution and Opportunity Costs*, 22 J. Land Use & Envtl. L. 1, 9 (2006) (“Undoubtedly, CBA proponents are correct to note that no society should flatly ignore the opportunity costs of precautionary regulation. But this is a trivial observation, for no serious proponent of the [Precautionary Principle] disagrees with it.”) (originally note 147).

¹⁰² *Id.* at 10 (originally note 148).

¹⁰³ Scholarship on weak versions of the Precautionary Principle has focused overwhelmingly on the decisions of regulatory agencies. See, e.g., Elizabeth Fisher & Ronnie Harding, *The Precautionary Principle and Administrative Constitutionalism: The Development of Frameworks for Applying the Precautionary Principle*, in *Implementing the Precautionary Principle: Perspectives and Prospects* 113, 116 (Elizabeth Fisher et al. eds., 2006) (The Principle applies to “administrative activities” that are “delegated by a primary law maker to a non-elected secondary law maker.”). The Strong Precautionary Principle, however, more directly addresses legislative design, given that the key choices, such as the degree of risk that will trigger burden shifting and the showing that must be made to overcome default prohibitions, are usually legislative choices (originally note 149).

to identify the tradeoffs involved, including opportunity costs and any countervailing risks. They should also seek to identify ancillary *benefits* from legislation, other than the intended target effects.¹⁰⁴ This is simply good decision making.

[...]

Consider ... the licensing of doctors and nurses. There is clearly a risk-risk tradeoff in a decision on whether to enact a licensing statute, with human lives at stake “on all sides” of the choice, to use a favorite phrase of the critics. Not requiring a license could result in untrained people performing surgeries and prescribing medications. Yet, requiring a license clearly limits the supply of doctors and nurses, which could put lives at risk in medically underserved communities.

But why should applying the Strong Precautionary Principle be “paralyzing” here? Would we really say that either decision would be equally precautionary, or that a precautionary legislator would find it impossible to decide? Consistent with the Strong Precautionary Principle, legislators have made a judgment (in all fifty states) that it is better to have a default rule prohibiting unlicensed medical practice, with the applicant bearing the burden of proof on their qualifications for the license, rather than to allow any person to practice medicine and then respond to any resulting harm on an *ex post* basis. We do not consider costs and benefits in deciding whether a particular nurse or doctor should be allowed to enter the market for medical services (the kind of individualized determination that Sunstein has advocated for chemical substances). We instead implement a firm rule that applies to all practitioners in every state: if you want to practice medicine, you must bear the burden to prove, *ex ante*, that you have the required skills, training, and knowledge. That there are some inevitable tradeoffs in grounding medical licensing rules in the Strong Precautionary Principle does not

¹⁰⁴ Interest in risk-risk tradeoff analysis arose in the 1980s and 1990s in tandem with the rise of cost-benefit analysis in regulatory decision making, and, according to some scholars, risk-risk tradeoff analysis consequently developed a distinct antiregulatory bias. Samuel J. Rascoff & Richard L. Revesz, *The Biases of Risk Tradeoff Analysis: Towards Parity in Environmental and Health-and-Safety Regulation*, 69 U. Chi. L. Rev. 1763, 1793 (2002). In particular, Rascoff and Revesz find bias in the identification of the countervailing risks from regulation, without attention to the ancillary benefits of regulation. As an illustrative example of ancillary benefits, they explain that statutes aimed at reducing carbon dioxide emissions, through fuel switching or energy efficiency measures, would also lead to reductions in conventional pollutants such as sulfur dioxide and ozone. *Id.* at 1808. They conclude that “[r]isk tradeoffs and ancillary benefits are simply mirror images of each other. There is no justification for privileging the former and ignoring the latter.” *Id.* at 1793 (originally note 150).

eviscerate the coherence of Strong Precaution as an approach to risk.

[...]

b. Addressing Tradeoffs Within Precautionary Legislation

My second response to the criticism that Strong Precaution necessarily ignores risk-risk tradeoffs is that such tradeoffs can often be addressed *within* regulatory regimes grounded in Strong Precaution. Indeed, the shift in the burden of proof to proponents may highlight policy alternatives that may not come to light under traditional decision-making frameworks.

There are numerous examples of permitting, licensing, and preapproval statutes that implement a default prohibition on a certain targeted activity, yet also contain procedures to address known countervailing risks of the prohibition. In drug approvals, for example, the FDA has developed three separate programs (fast track, accelerated approval, and priority review) to reduce “drug-lag,” shortening the time in which breakthrough drugs can move to market.

Pharmaceutical manufacturers under those programs still cannot market a new drug until they receive FDA approval, but they proceed down a separate and faster track within the agency. Environmental statutes that rely on licensing and preapproval mechanisms are replete with “minor source,” “minor use,” and “experimental use” exemptions, to avoid applying strict research, preapproval, and pollution control requirements to activities that, as a class, pose less severe risks.

In arguing that the presence of risk-risk tradeoffs makes precaution incoherent, Sunstein, Graham, Wiener, Cross, and other critics often point to the same example: DDT. The critics explain that banning DDT could lead to increased deaths from malaria in the developing world. Given this countervailing risk of a DDT ban, critics suggest, a goal of caution is paralyzing. Is it cautious to ban DDT or to do the opposite? When international negotiators actually confronted this tradeoff in the design of the Stockholm Convention on Persistent Organic Pollutants, they formalized a global ban on the production and use of DDT and then provided an exception for parties that register their intent to continue to use

DDT for mosquito control.¹⁰⁵ Through permitting opt-outs from the global ban, negotiators achieved a compromise that addresses both the target risk (environmental and health harms from DDT) and one countervailing risk of a ban (malaria).

Strong Precaution does not inexorably lead to abdication of judgment. It does not doom legislators to extremist positions in which known, countervailing risks and the costs of legislation must be ignored. Rather, as the above examples illustrate, legislators can structure a Strong Precautionary regulatory regime to address both the “target risk” and countervailing risks.

For a particularly elegant example of this in chemical regulation, consider REACH, the EU chemical legislation discussed previously. The EU explicitly grounded the legislation in the Precautionary Principle,¹⁰⁶ and the legislation relies heavily on burden shifting. In particular, REACH imposes a default sunset date after which a chemical identified as a “very high concern” substance cannot be marketed in the EU.¹⁰⁷ “Very high concern” is a term of art, defined in REACH in reference to the intrinsic properties and hazards of the chemical (such as carcinogenicity or bioaccumulative properties).¹⁰⁸ REACH further provides that chemical manufacturers can overcome the default and receive authorization to market a “very high concern” chemical through demonstrating, among other things, that the socioeconomic benefits of the chemical exceed potential costs.¹⁰⁹

REACH provides an important reminder that a Strong Precautionary regulatory regime need not preclude consideration of the costs and benefits of regulation. In a twist on prevailing approaches in the United States, however, REACH shifts the burden of conducting the cost-benefit analysis to the manufacturer that seeks to

¹⁰⁵ Stockholm Convention on Persistent Organic Pollutants, Annex B.II, May 22–23, 2001, *opened for signature* May 23, 2001, 40 I.L.M. 532 (2001). Under Annex B.II of the Convention, continued use of DDT is subject to developing an implementation plan that includes a search for suitable alternatives to DDT. *Id.* (originally note 178).

¹⁰⁶ See Commission Regulation 1907/2006, 2006 O.J. (L 396) 1 (EC), at 396 (originally note 180).

¹⁰⁷ See Noah M. Sachs, *Jumping the Pond: Transnational Law and the Future of Chemical Regulation*, 62 Vand. L. Rev. 1817, 1837–38 (2009) (providing an overview of REACH) (originally note 181).

¹⁰⁸ *Id.* at 1837 (originally note 182).

¹⁰⁹ *Id.* at 1838 (originally note 183).

continue to market a chemical substance with known hazardous properties.¹¹⁰ REACH also requires that manufacturers, in the authorization process, analyze whether less hazardous substitute chemicals exist.¹¹¹ In this way, the “stop and think” mechanism of Strong Precaution, far from ignoring tradeoffs, becomes the spur for a formal alternatives analysis that has never been a prominent part of U.S. chemical regulation.

C. The Limits of Strong Precaution

The Strong Precautionary Principle provides a framework to protect public health and safety, and some advocates of the Principle naturally want to universalize it as a guiding polestar for all areas of risk. As the authors of the Wingspread Statement put it, “[c]orporations, government entities, organizations, communities, scientists, and other individuals must adopt a precautionary approach *to all human endeavors*.”¹¹²

The Principle has some significant limitations, however, and it needs to be applied judiciously. For one thing, many forms of human health and environmental risks have no “proponent” to which the burden of proof can be shifted. Consider natural disasters, the spread of contagious disease, or the remediation of historic contamination at a hazardous waste site. We need government to take appropriate preparations for these risks, and in the case of historic waste sites, to set appropriate cleanup standards, but there is no private “proponent of the activity” who can bear the burden of proof. Climate change is another example of a problem not easily addressed through burden-shifting mechanisms. A climate change strategy must include promoting energy efficiency and pricing greenhouse gas emissions,¹¹³ and arguments over burden of proof would merely

¹¹⁰ *Id.*; see also Wendy E. Wagner, *The Precautionary Principle and Chemical Regulation in the U.S.*, 6 Hum. & Ecological Risk Assessment 459, 473 (2000) (noting that a system in which cost-benefit analysis is used by manufacturers to rebut precautionary health-based regulation would be consistent with the Precautionary Principle) (originally note 184).

¹¹¹ See Sachs, *Jumping*, *supra*, at 1841–42 (discussing REACH’s incentives for development of substitutes for hazardous chemicals) (originally note 185).

¹¹² See, e.g., Protecting Public Health & the Environment: Implementing the Precautionary Principle app. A, at 353 (Carolyn Raffensperger & Joel A. Tickner eds., 1999) (emphasis added). The context of this quotation makes it clear that the authors were describing the Strong Precautionary Principle, as the Wingspread Statement goes on to provide a definition of precaution that includes burden shifting. See *id.* at 353–54 (originally note 186).

¹¹³ See Noah M. Sachs, *Greening Demand: Energy Consumption and U.S. Climate Policy*, Duke Envtl. L. & Pol’y 295, 298 (2009) (originally note 187).

distract from those urgent national tasks.¹¹⁴

A second reason that Strong Precaution should not be viewed as universally applicable is that putting government in a risk gatekeeping role can be complex, potentially anticompetitive, and expensive. PhD-level government personnel would often be needed to review risk data compiled by private parties and to determine (subject to layers of judicial review) whether the proponent has met its burden of proof. The expansion of agency bureaucracy that would result from indiscriminate application of Strong Precaution would be a counterintuitive outcome from a principle designed to shift the burden of proof *away* from government.

For these reasons, the gatekeeping mechanism of the Strong Precautionary Principle should be reserved for serious threats that cannot be addressed through less intrusive mechanisms. For many forms of risk, there is no need to resort to the aggressive *ex ante* approach to risk implicit in the Strong Precautionary Principle. Consider, for example, performing surgery, marketing a dangerous recreational activity (e.g., sky diving), or marketing a dangerous power tool (e.g., chainsaw, jackhammer). These all raise serious safety risks, but the liability system, insurance, and tort-based deterrence have performed reasonably well for addressing these risks in the United States. There is no need to institute complex legal mechanisms for risks that are already being handled acceptably through less aggressive means.

The most pressing issue for legislators, therefore, is how to make the fundamental judgment alluded to in the prior section: determining when it makes sense, despite attendant complexity, to implement the Strong Precautionary Principle as the basis for a risk regulatory regime. Relevant criteria should include the administrative costs of establishing an *ex ante* review mechanism; the suitability of tort remedies or criminal penalties to address the risk; whether risk creators are likely to be judgment proof in the event that harm occurs; whether the risk is one to which participants consent (as in surgery or in dangerous recreational activities), or is in the nature of a nonconsensual externality; and whether the risk could be addressed adequately through postmarket regulatory

¹¹⁴ I do not mean to minimize the role of precautionary steps in addressing climate change. Policy making should still be guided by weaker versions of the Precautionary Principle, allowing early action in the face of scientific uncertainty. For more on the role of precaution in the climate debate, see Applegate, *Climate Change*, *supra*, at 171 (originally note 188).

mechanisms.¹¹⁵

Comprehension Check¹¹⁶

1. What is Sachs's version of the "strong precautionary principle"? How is it similar to and different from Sunstein's?
2. Who bears the burden of proof in rebutting the regulatory default under Sachs's version of the strong precautionary principle? Who determines how high the burden is?
3. What does it mean to bear the "burden of production" or the "burden of persuasion" in this context?
4. What is "paralysis by analysis"?
5. When does Sachs think that a precautionary approach should be avoided?

D. Case Study: Generative AI

There is an ongoing debate as to whether we should adopt a more permissive or precautionary approach to regulating AI. On one hand, some predict that the risks are immense and existential ("x-risk"), up to and including human extinction. On the other hand, the probability of this risk seems quite small. Arvind Narayanan and Sayash Kapoor, for example, consider the probability of human extinction due to AI

¹¹⁵ Steve Shavell explored some of these same criteria nearly thirty years ago in his classic article on the merits of liability rules versus public regulation as tools for risk management. Steven Shavell, *Liability for Harm Versus Regulation of Safety*, 13 J. Legal Stud. 357 (1984) (originally note 189).

¹¹⁶ For further critique of Sunstein's argument, see Gregory N. Mandel & James Thus Gathii, *Cost-Benefit Analysis Versus the Precautionary Principle: Beyond Cass Sunstein's Laws of Fear*, 51 U. Ill. L. Rev. 1037 (2006). For a complementary articulation of a "non-naive" precautionary principle, see Nassim Nicholas Taleb, Rupert Read, Raphael Douady, Joseph Norman & Yaneer Bar-Yam, *The Precautionary Principle (with Application to the Genetic Modification of Organisms)*, Extreme Risk Initiative—NYU School of Engineering Working Paper Series.

and conclude that “AI x-risk forecasts are far too unreliable to be useful for policy.”¹¹⁷

The following reading responds to their argument, noting reasons to favor a precautionary approach in this context, albeit without clarifying what form it should take.

Matthew Tokson, [*Uncertainty, Catastrophic Risk, and AI Regulation*](#),
Lawfare (Sep. 16, 2024)

Imagine this scenario: America’s top scientists are researching a new technology on behalf of a U.S. regulatory agency. They report to the agency head that there is a substantial risk that the new technology, if deployed, could cause enormous harm to the U.S. civilian population. However, they cannot quantify the risks involved with any precision. There is a great deal of uncertainty around the potential effects of the novel technology. But the [majority of researchers](#) in the field believe that it is likely to be remarkably dangerous. Should the relevant agency, or Congress, regulate the new technology? Or, given the inherently uncertain nature of the risk, should they do nothing?

[Several prominent voices](#) in the artificial intelligence (AI) regulation debate have made versions of the latter argument, contending that policymakers should ignore catastrophic AI risks given their inherently uncertain nature. The most recent and most thorough example of this uncertainty argument was laid out a few weeks ago by the computer scientists/pundits at [AI Snake Oil](#), and it’s worth considering in detail. Ultimately, their argument regarding the uncertainty of AI risk is flawed on its own terms, and several of its premises are mistaken, as I discuss below. But a primary issue is their failure to grapple with the history and literature of regulating risks in contexts of uncertainty. AI is not the first area where regulators or forecasters have had difficulty assigning exact numbers to the potential risks involved. And the idea of waiting until catastrophic risks are precisely quantified before addressing them is one only the most radical libertarian could approve.

¹¹⁷ Arvind Narayanan & Sayash Kapoor, [AI Existential Risk Probability Are Too Unreliable to Inform Policy](#), AI as Normal Technology (formerly AI Snake Oil) (Jul. 26, 2024) (first in a series of essays for policymakers concerned about AI x-risk).

Regulation Under Uncertainty

Proponents of uncertainty arguments against AI regulation generally start from a place of regulatory skepticism. For example, the AI Snake Oil pundits' starting point in assessing regulation is that "the state should not limit people's freedom based on controversial beliefs that reasonable people can reject." Taken seriously, this principle would invalidate a massive swath of current laws and regulations, and prevent most new regulations. The AI Snake Oil pundits are indeed skeptical of any restrictions on AI development, although they seem to recognize a substantial obstacle to their view—that, according to a [recent survey](#), a majority of AI experts provide a 10 percent or higher probability of humans becoming extinct or its functional equivalent in the future due to advanced AI systems. This is—to put it mildly—quite a lot of risk of a bad outcome when the bad outcome is essentially the wiping out of all human life forever.

However, the 10 percent figure is not a precise calculus of the risk involved; rather, it is an estimate in the context of substantial uncertainty. Instead of a simple mathematical output calculated via historical or experimental data, it is merely the best guess of published experts in the field. But this is no reason for policymakers to ignore it. Those who know AI technology best are all but shouting about its dangers from the rooftops. Regulators need not wait until those dangers are quantified to the last decimal place to start addressing them. Of course, regulation might be unjustified on other grounds, for example, if policymakers estimate that the (uncertain) benefits of AI innovation impeded by regulation likely outweigh AI's (uncertain) harms. The point is simply that the existence of uncertainty surrounding a new technology or harm should not foreclose regulation.

Indeed, there is a [substantial, venerable literature](#) on regulating under conditions of uncertainty. For over a century, scholars have recognized that policymakers must sometimes act to address risks even when those risks cannot be calculated precisely. In areas such as climate change, pandemics, and genetic modification, policymakers have addressed risks that can only be estimated or guessed at rather than quantified. To be sure, the level of uncertainty surrounding AI is currently greater than that surrounding these other areas (this is discussed further in the next section). But full quantification of risk remains difficult in each context. Nonetheless, regulation is often appropriate in these contexts despite uncertainty about the future.

A [widely held view](#) within the regulatory literature is that precautionary regulatory approaches may be justified in situations of uncertainty when a new technology or practice creates a nontrivial risk of catastrophic harm. In such situations, policymakers can maximize welfare by pursuing a “maximin” strategy—that is, choosing the policy approach with the best worst-case outcome. Under this approach, if regulating a new technology with uncertain costs and benefits reduces a meaningful catastrophic risk, regulation is justified and a better option than merely waiting for the costs and benefits to become more easily quantifiable.

Artificial intelligence is a prime example of a technology for which a maximin, precautionary regulatory strategy is appropriate. The path of its future development is uncertain, and—according to a [majority of experts in the AI field](#)—it poses a substantial risk of catastrophic harm. Given these characteristics, meaningful regulation, rather than a laissez-faire approach or even a flexible, wait-and-see approach to policymaking, is currently the optimal choice. Early laws directed at catastrophic AI risk, such as [California’s proposed SB 1047](#), which aims to impose basic safety requirements on developers of frontier AI models, are a small step in the right direction. The cautious, government-skeptical philosophy adopted by some pundits risks delaying meaningful restrictions on dangerous technologies until it is far too late.

Calculating AI Risks

While the premise of the AI Snake Oil pundits’ argument that policymakers should not regulate on the basis of uncertain risks is wrong, their specific arguments about the total unquantifiability of AI extinction risks are also flawed. Though the risks are genuinely uncertain and can only be roughly estimated, AI experts are not pulling their extinction risk estimates from thin air. They are estimating AI extinction risk the same way anyone might estimate an uncertain future risk: by reference to prior, observed events. The AI Snake Oil pundits contend that human extinction due to advanced AI is so unlike any prior event that history is no guide, at all, whatsoever. But the various components of most AI extinction scenarios—the development of supercapable AI, the failure to align that AI with human goals or safety, and the extinction of a displaced or disempowered species—are tractable, with historical precedents that are relevant and capable of providing a basis for estimation. One needs only to multiply these estimated

component probabilities to produce an estimated probability of AI-caused extinction risk.

For example, someone estimating the likelihood that researchers will someday develop generally supercapable or generally superintelligent AI—defined for present purposes as AI with general capabilities or general intelligence beyond that of humans—might look to areas where current AIs or related autonomous systems already exceed human capability or intelligence in narrow domains. There are many such domains, from strategy games such as [chess](#) and [Go](#), to [many mechanical or automated tasks](#), to [aspects of aerial and ground combat](#). Modern-day AIs, though sharply limited in their general capabilities, already achieve [standardized test scores that exceed human averages](#) on a vast array of cognitive and creative tests. These instances provide a relevant context from which inferences about future capability gains can be made. Likewise, previous breakthroughs in AI design, such as the development of the [transformer](#) that made modern large language models possible, provide a context for inferences regarding potential future breakthroughs. The same is true of the exponential growth of AI capabilities over the past several years. Whether such growth is predicted to continue to proceed exponentially, to slow, or to grind entirely to a halt, the fact of its existence and the history of other exponentially advancing technologies provide a basis for inferences about its future course.

Similarly, someone estimating the likelihood that a supercapable or superintelligent AI would become misaligned with human goals or interests can look to areas where existing AIs have proved to be misaligned with human goals or interests. They can also look, more broadly, at areas throughout law and society where human agents, corporations, and other entities have been unable to fully align with the goals of their principles, let alone with the interests of humanity writ large. Even confining oneself solely to the AI context, the examples are numerous and enlightening. For instance, ChatGPT [seems to go “insane”](#) from time to time, giving bizarre or nonsensical answers and [insulting or threatening its users](#). Then there’s the AI system trained to sort data quickly that realized the quickest way to sort data was to [delete it all](#). Or the system that [learned to pretend it was inactive](#) to avoid the scrutiny of the researchers in charge of it. My [recent article](#) (with Yonathan Arbel and Albert Lin) includes numerous other examples. In other words, there is ample relevant context for estimating the probability of AI alignment problems.

In assessing the threat to humanity from a supercapable or superintelligent AI unaligned with human interests—to the extent that this is not obvious on its face—a person might look to the history of extinctions caused by invasive species or by humanity itself to other species substantially less capable and intelligent than our own. Humans, currently the most generally capable species, affect habitats and threaten other species in countless ways. In the few hundred years since humanity came to dominate Earth’s ecosystems, the rate of extinctions of other species has increased to somewhere between [100 and 1,000 times](#) the historical rate. Similarly, there are [numerous recent examples](#) of invasive animal species causing the extinction of native animal species that may provide a relevant backdrop for inferences regarding potential human extinction.

None of these historical instances are the same as scenarios involving AI-driven extinction, of course. They merely share some characteristics and do not share others. But the same can be said of any novel future event that experts and forecasters might try to predict. History may provide an imperfect guide for estimation, yet estimation is hardly futile. Uncertainty inevitably surrounds the future course of phenomena like climate change, or pandemics, or artificial intelligence. But when a majority of experts in these fields warn that a catastrophe may easily occur if policymakers do nothing, those policymakers should listen.

Putting it All Together: Chapter 10

For this exercise, focus solely on the existential risk (“x-risk”) to humanity posed by further AI development.

1. What are the benefits of a precautionary approach to AI x-risks? What are the drawbacks?
2. Which entities are likely to argue for more regulation of AI to address its x-risks? Are they likely to be able to band together? Are they likely to have political power?
3. Which entities are likely to argue for less regulation of AI, despite the potential for x-risk? Are they likely to be able to band together? Are they

likely to have political power?

4. If you wanted to apply a precautionary approach to regulating AI, would you ban certain aspects of its development or use? Place a moratorium on its development or use? Create a new regulatory regime?
 - a. If you want to ban it, can you? How broadly or narrowly would you craft it? Would you carve out any exceptions?
 - b. If you want a moratorium, what would it cover? When would it expire?
 - c. If you want a new regulatory regime, what would it encompass? Safety standards? Licensing requirements? Transparency requirements? Reporting requirements?
5. Generally, when do you think that a precautionary approach should be adopted? Avoided?