FREE MARKETS *and* SOCIAL JUSTICE

CASS R. SUNSTEIN

Oxford University Press
New York Oxford

Introduction

We are in the midst of a period of mounting enthusiasm for free markets. This is of course true in many of the former Communist nations. It is true for much of the West as well, prominently including England and the United States.

Free markets are often defended as an engine of economic productivity, and properly so. But they are also said to be required for social justice, and here things become far more complex. Certainly there are connections between free markets and social justice. A system aspiring to social justice aspires to liberty, and a system of free markets seems to promise liberty, because it allows people to trade goods and services as they wish. In fact, a system of free markets seems to promise not merely liberty but equality of an important sort as well, since everyone in a free market is given an equal right to transact and participate in market arrangements. This form of equality should not be trivialized or disparaged. For example, race and sex discrimination has often consisted of exclusions of certain classes of people from the market domain. In both South Africa and the United States, discriminatory practices frequently took the form of incursions on free markets in employment.

An appreciation of the virtues of free markets has been an important part of the economic analysis of law, perhaps the most influential development in legal education in the last quarter-century, and a development with growing effects on public policy in the United States and abroad. As it operates in law schools, economic analysis is concerned above all with the consequences of legal rules. Often its practitioners have asked whether intrusions on free markets have desirable consequences. If the minimum wage is increased, what, exactly, will happen? What are the real-world consequences of bans on discrimination, legal rules for controlling air pollution, and rent control laws? There are limits to how much economic theory can say on such issues; empirical evidence is necessary. But by looking at the effects of law on incentives, economics can point in the right direction. Often it can show that the consequences of intrusions on markets will be unfortunate or even perverse. The economic analysis of law has produced significant advances, many of which are discussed in this book. Much remains to be learned.

This book is not, however, a simple celebration of free markets, and it raises a number of questions about economic analysis of law in its conventional form. Free markets can produce economic inefficiency and (worse) a great deal of injustice. Even well-functioning economic markets should not be identified with freedom itself. Freedom is a complex notion, to say the least, and free markets can sharply limit freedom as that term is usually understood. In fact, free markets depend on a range of coercive legal interventions, including the law of property, which can be a serious intrusion on the freedom of people who lack ownership rights. And it should not be necessary to emphasize that important forms of equality—including race and sex equality—can be undetermined, not promoted, by free markets. Race discrimination is often fueled by market forces.

Moreover, economics—at least as it is used in the conventional economic analysis of law—often works with tools that, while illuminating, may be crude or lead to important errors. Consider, for example, descriptive or "positive" economics. The economic analysis of law has been built on a certain conception of human rationality, in which people are seen as "rational profit-maximizers." For some purposes, this is a helpful foundation. Certainly it is true that most people try, most of the time, to find ways of promoting their own ends. But it is not always clear in what sense human beings can be said to be "rational" or "profit-maximizers." The motivational foundations of human behavior have enormous complexity. Sometimes people do not seem at all rational. Sometimes they are ignorant and sometimes they seem to defeat their own goals. People rely on rules of thumb, or heuristic devices, that cause them badly to misunderstand probabilities and facts; this can lead to irrationality for individuals and societies alike. Sometimes people undervalue their own futures, or suffer from weakness of will, or choose what they know, on reflection, they ought not to choose. Sometimes people care not just about social outcomes, but also about the "meaning" of such outcomes, that is, the values expressed in and by the outcomes. This point very much bears on law, in such areas as environmental protection, race and sex equality, and occupational safety and health (chapter 2).

Above all, social norms are an important determinant of behavior, and they have received far too little attention from those interested in free markets, economic analysis of law, and social justice. A number of puzzles and anomalies underlie human decisions. Analysis of law, economic or otherwise, would do well to incorporate an understanding of these puzzles and anomalies (chapters 2 and 4).

Thus far, I have been discussing descriptive or positive economics. If we turn to the evaluative side—to questions about what the law should be or do—economic analysis of law encounters equally serious problems. In its usual form, economics offers an inadequate understanding of social welfare. Often it is concerned with the satisfaction of existing preferences. This is far from an unworthy goal; the frustration of peoples' preferences can lead to misery and injustice. But in any society, existing preferences should not be taken as natural or sacrosanct. They are a function of context. Sometimes they are a product of deprivation, injustice, or excessive limits in available opportunities.

Moreover, the economists' conception of social welfare is too "flat," insofar as it evaluates diverse social goods along the same metric. People care about things not just in terms of amounts, but also in different ways. Some human goods, like cash, are simply for use. But people value things for reasons other than use. They respect other people; sometimes they love each other; they see some things, like a painting or a beach, as objects of awe and wonder (chapter 3). A well-functioning legal system attempts to make space for people's diverse valuation of diverse human goods. This point bears on the uses and the limits of free markets.

More particularly, this book develops seven basic themes:

1. *The myth of laissez-faire.* The notion of "laissez-faire" is a grotesque misdescription of what free markets actually require and entail. Free markets depend for their existence on law. We cannot have a system of private property without legal rules, telling people who owns what, imposing penalties for trespass, and saying who can do what to whom. Without the law of contract, freedom of contract, as we know and live it, would be impossible. (People in Eastern Europe are learning this lesson all too well.) Moreover, the law that underlies free markets is coercive in the sense that in addition to facilitating individual transactions, it stops people from doing many things that they would like to do. This point is not by any means a critique of free markets. But it suggests that markets should be understood as a legal construct, to be evaluated on the basis of whether they promote human interests, rather than as a part of nature and the natural order, or as a simple way of promoting voluntary interactions.

2. *Preference formation and social norms.* It is important not only to know what choices and preferences are, but also to know how they are formed, and whether they are really connected with human well-being. Unjust institutions can breed preferences that produce individual and collective harm. Severe deprivation—including poverty—can be an obstacle to the development of good preferences, choices, and beliefs. For example, a society in which people "prefer" to become drug addicts, or violent criminals, has a serious problem. Such preferences are likely to be an artifact of existing social norms, and those norms may disserve human freedom or well-being.

In this light a society should be concerned not simply and not entirely with satisfying the preferences that people already have, but more broadly with providing freedom in the process of preference-formation. Social practices, including law, will inevitably affect preferences. There is no way for a legal system to remain neutral with respect to preference formation. In these circumstances it is fully legitimate for government and law to try to shape preferences in the right way, not only through education, but also (for example) through laws forbidding racial discrimination, environmental degradation, and sexual harassment, and through efforts to encourage attention to public issues and to diverse points of view.

3. *The contextual character of choice.* Choices are a function of context. If someone takes a job that includes a certain danger, or chooses not to recycle on a Tuesday in March, or discriminates against a certain female job candidate, we cannot infer a great deal about what he "prefers" or "values." All of these choices might be different in a different context. Our discriminator may support a law banning discrimination; people who do not recycle in March may recycle in May or June, and they may well support laws that mandate recycling. Economists and economically oriented analysts of law sometimes think that they can derive, from particular choices, large-scale or acontextual accounts of how much people value various goods. This is a mistake, involving extravagant inferences from modest findings.

4. *The importance of fair distribution.* It is necessary to know not simply whether a society is rich in economic terms, but also how its resources are distributed. Thus a problem with gross domestic product, as a measure of social well-being, is its obliviousness to distributional concerns. Free markets can help fuel economic growth, and economic growth can improve people's lives. But many citizens are not benefited by growth, and at a minimum government should take steps to combat human deprivation and misery in the midst of growth. In any case, it is important to develop standards for measuring social well-being that allow people, in their capacity as citizens and voters, to focus on the issue of distribution.

5. *The diversity of human goods.* I have noted that human beings value things not just in different amounts, but also in different ways. They value a friend in one way; a park in another; a species in another; a spouse in another; an heirloom in another; a large check in another; a pet in still another. The way they value a funny movie is qualitatively different from the way they value a tragedy, a mountain, a beach, or a car. Insofar as economics uses a single metric or scale of value, it flattens qualitative differences. For some purposes the flattening is very useful, but for other purposes it is harmful. Such qualitative differences can be crucial to private life, public life, and social science.

These points help show the inadequacy, in some settings, of social ordering through markets. Some goods should not be sold on markets at all; consider the right to vote or the right to be free from discrimination on the basis of sex. Markets work best when (what we rightly treat as) consumption choices are at issue. In a liberal society, some choices should be understood not to involve consumption choices at all. For example, the choice to vote, and the decision how to vote, are not best understood as involving mere consumption. How a person votes should depend not on anyone's "willingness to pay," but on the reasons offered for and against a certain candidate. The right to vote is debased if it is understood as simply a matter of "buying" something.

This claim bears on a range of issues involving the use of markets for distributing such goods as reproductive capacity, sex, endangered species, and environmental amenities in general. Some goods are best allocated on the basis of an inquiry not into economic value, but into the reasons offered for any particular allocation.

6. *Law can shape preferences.* We have seen that no market can exist without legal rules. Legal rules must also allocate entitlements. In a system of private property, it is necessary to say who owns what, at least in the first instance. It is also necessary to create rules of tort law, saying who can do what to whom, and who must pay for injuries and harms. As a great deal of empirical work has shown, these legal rules, allocating basic entitlements, have effects on choices and preferences. Someone who has been given a legal entitlement —to chocolate bars, clean air, freedom from sexual harassment, environmental goods—may well value the relevant good more than he would if the good had been allocated to someone else in the first instance. The preference-shaping effects of the initial allocation via law raise important questions for the analysis of law and free markets. They suggest that government and law may not be able to leave preferences "as they are."

7. *Puzzles of human rationality.* Are human beings rational? What criteria should we use to answer that question? These questions have become all the more urgent in light of recent work attempting to apply economic models of rationality to race and sex discrimination, family choice, pollution problems, aging, even sexual encounters (and hence to the problem of AIDS). Such work has been highly illuminating, but it is important to know in exactly what sense people might be said to be (or not to be) "rational" when they choose spouses, recycle garbage, reject employees of certain kinds, or engage in sexual or risky activity. Any views on such matters may well depend on a controversial account of what it means to be rational.

I am especially interested in the consequences for law and policy of recent experimental and theoretical work on rationality, social norms, and individual valuations. This work suggests that people's choices and judgments are quite different from what traditional economists predict. People's choices are a function of the distinctive social role in which they find themselves, and we may act irrationally or quasi-rationally. In particular, this work points to the important role of social norms involving fairness, reciprocity, and cooperation in producing individual choice. If we do not uncover the relationship between rationality and social norms, we will make major mistakes in designing policies to accomplish our goals.

This book investigates these points in many different contexts, involving property, protection of nature, race and sex discrimination, broadcasting, occupational safety, and much more. I do not take these points to pose across-the-board challenges to the use of free markets. On the contrary, I argue that in some situations, constitutions themselves should protect free markets (chapter 7), and also that market-oriented policies are far better in the areas of health, safety, and the environment than is generally recognized (chapters 13 and 14). But free markets are a tool, to be used when they promote human purposes, and to be abandoned when they fail to do so. Moreover, economic analysis of law should be based on an understanding of how human beings actually behave. An inquiry into these topics has a substantial empirical direction, and I shall refer to empirical matters—involving smoking, polluting, dis-

criminating, and other issues—at many points. But it is also important to be clear on our underlying judgments about social justice.

The book comes in three basic parts. Part I deals with foundations—with the appropriate role of existing "preferences," the importance of social norms, the question whether human goods are commensurable, and issues of distributional equity. I claim that the term "preference" is highly ambiguous and that people's "preferences," as they are expressed in the market domain, should not be deemed sacrosanct. On the contrary, market "preferences" are sometimes a product of background injustice or of social norms that people do not really like. Acting as citizens, people should be permitted to change those norms. I also argue that human goods are not commensurable; there is no metric by which we can assess such goods as environmental quality, employment, more leisure time, less racial discrimination, and so forth. My discussion of "measuring well-being" is designed to respond to this problem; my general treatment of incommensurability shows how the absence of a unitary metric plays a large role in both law and daily life.

Part II deals with rights. The basic goal is to show that markets have only a partial and instrumental role in the protection of rights. In a claim of special relevance to current disputes in the United States, I argue that free markets are not likely to stop discrimination on the basis of race and sex. On the contrary, markets often promote discrimination. I also claim that in recent years, the free speech *principle* has been wrongly identified with free speech *markets*—as in the striking claim by a recent chairman of the Federal Communications Commission that television "is just another appliance," or "a toaster with pictures." Against this view, I argue that the free speech principle is mostly about democratic deliberation, not about free markets. This idea has a range of consequences for how we think about (for example) radio, the Internet, and cable television. This claim certainly bears on the subject of campaign finance regulation, to which a chapter is devoted.

But Part II shows that markets can play an important role in a regime of rights. Thus I argue that constitutions, at least in Eastern Europe, should attempt to protect private property and market ordering. This argument depends on the claim that constitutions should be "countercultural," in the sense that they should protect against those aspects of a country's culture and traditions that are most likely to produce harm. In Eastern Europe, there is a pressing need to establish the institutions of a civil society, including a degree of market exchange.

Part III deals with regulation, especially in the context of risks to life and health. A common theme is that reliance on "free markets" would be a huge mistake for regulatory policy and that deregulation would often be a foolish solution. On the other hand, many of the problems in federal regulation in the United States stem from Soviet-style command-and-control regulation, which produces billions of dollars in wasted money and laws that are far too ineffectual in improving health and saving lives. American government does not set sensible priorities; it does not use the best regulatory tools; risk regulation sometimes increases risks. Regulation is also an ineffective tool for redistribut-

wealth. Thus, the essays in Part III try to develop approaches that would promote both economic and democratic goals. A particular theme is the uses and limits of cost-benefit analysis. Balancing of costs and benefits is far better than absolutism. But everything depends on how "costs" and "benefits" are specified. I suggest that many statutes have goals other than economic efficiency, and that such statutes are entirely legitimate. With respect to statutes based on economic grounds, we should try to promote social commitments in the most cost-effective manner.

I also suggest that it is time to move beyond the increasingly tired and increasingly helpful question whether we should have "more" or "less" government or "more" or "less" in the way of free markets. These dichotomies are far too crude. As we have seen, markets depend on government. Sometimes government can improve existing markets by creating good incentives for socially desirable behavior. Sometimes markets should be supplemented by government services, like education, job training programs, and health care. There is no inconsistency in urging greater reliance on market instruments in some areas while insisting on a larger role for the public sector in others. In any case, future problems are not usefully approached by asking whether there should be "more" or "less" regulation. The real question is what kinds of regulations (emphatically including those that make markets possible) promote human well-being in different contexts.

In a book of this sort it would be far too ambitious to attempt to announce a theory of justice. But debates that seem intractable at the most abstract levels may admit of solutions when the question is narrowed and sharpened, and hence an inquiry into the relation between markets and justice may be most productive when we draw close attention to the setting in which market remedies are proposed. To the seven points I have listed above we may therefore add one more: Achievement of social justice is a higher value than the protection of free markets; markets are mere instruments to be evaluated by their effects. Whether free markets promote social justice is an impossible question to answer in the abstract. Far more progress can be made by examining the circumstances in which markets, adjustments of markets, and alternatives to markets are proposed as solutions.

example, leisure, consumer goods, vacation, and education for their offspring. But much depends on what the ambiguous word "implicitly" means. The fact that the trade-off is not made explicitly is hardly a matter of indifference. When the trade-off is made only "implicitly," it is not well described as an ordinary trade-off at all. The actor may be showing a commitment to a certain set of judgments about how relationships and prospects should be valued, and if the trade-offs were made explicitly, that commitment would be undermined or even violated. The explicit trade is not equivalent to the implicit one; the absence of explicitness maintains certain social norms. In an analogous context, Holmes wrote, "[e]ven a dog distinguishes between being stumbled over and being kicked";[15] the difference between a trip and a kick—identical in effect but conveying very different attitudes—is related to the difference between an implicit and an explicit trade. In these circumstances, it is hopelessly underdescriptive to claim that someone "implicitly" trades off (say) cash and friendship in making lunch decisions, even though this way of seeing things might be quite helpful for predictive purposes.

Economic or utilitarian descriptions of human behavior are unlikely by themselves to alter social norms or law in any significant way.[16] But social norms are in a constant process of evaluation and flux, and the description of kinds of valuation is one method by which norms are created and altered. The making of law reflects this process of description and valuation; law can have effects on kinds of valuation as well. Indeed, many of the sharpest disputes in law relate to the appropriate kinds of valuation to bring to bear on disputed problems. Here the pervasive question is what sorts of valuations the law ought to encourage or extinguish.

If kinds of valuation matter to human life, judgments about such kinds will matter a great deal to law as well. And in many areas of law, the question of commensurability occupies a surprisingly important place.

IV. Law in General

Preliminaries

These diverse kinds of valuation do not fit neatly into legal categories. We cannot say that, because people value different events in different ways, it follows that law should have a particular content. Nor can we draw lessons for law from the bare fact of incommensurability in any of the senses discussed here. This is so for several reasons.

First, we need an account of which kinds of valuation are appropriate in order to make recommendations for law. We are not and should not be agnostics or skeptics about appropriate kinds of valuation; they can be evaluated on the basis of reasons. Nor should we think that, because people value a certain thing in a certain way, the law ought to do so as well. Of course, democracy has its claims; but there may be good reason to challenge any particular popularly endorsed judgment on kinds of valuation. Indeed, some aspects of constitutional law represent an effort to discipline democratic discussion by limiting

possible kinds of valuation, on the ground that they are too sectarian or inconsistent with principles of civic equality. When there are no constitutional barriers, the democratic process is formally unconstrained, but an important part of democratic debate consists of challenges to the kinds of valuation that prevail among the democratic majority.

Any judgments about appropriate kinds of valuation are of course a complex matter. This is so not least because, in a heterogeneous society, the state ought to allow a wide range of diverse valuations. The regulation of valuations can be a stifling matter. At least as a presumption, the state ought to allow people to value in the way that they see best. These considerations, however, do not imply that the state should or can remain agnostic about proper kinds of valuation. Incommensurabilities—even of the form that block certain reasons for action—are not merely freedom-reducing, but constitutive of certain sorts of freedom, for they make possible certain valuable relationships and commitments. Moreover, a state will have a hard time if it seeks to be entirely neutral about valuations. Global neutrality is impossible. The state has to make decisions about how to allocate rights and entitlements; it has to decide what can be traded on markets and what will be subject to politics, and these decisions inevitably will take some sort of stand on appropriate valuations.

Second, a judgment about the appropriate kind of valuation, even if it can be reached and persuasively defended, need not entail a particular conclusion for law. For example, it might be shown that prostitution entails an improper valuation of human sexuality, but this need not mean that prostitution should be outlawed. To justify outlawing prostitution, we must make an additional set of arguments about both the likely effectiveness of the ban and the principle that supports use of the coercive power of the state. Or we might think that the best human life involves a certain way of valuing the environment, without thereby rejecting, for example, tradable emissions permits as a regulatory tool. Perhaps tradable permits do not affect valuations of the environment; perhaps any such effects are minimal and well justified by the various gains. Or we might believe that animals deserve consideration and therefore ought not to be eaten, but also believe that the law should not require vegetarianism. State-compelled vegetarianism might well be ineffectual and inconsistent with individual liberty, rightly conceived. These examples show that any general claim about the right kind of valuation needs a great deal of supplementation to result in concrete recommendations for law and policy. There are also recurring questions about the feasibility of various legal strategies.

More generally, we can defend one kind of valuation for law and government and other kinds for family, church, and civil society. The law might, for example, insist on calculating the value of human life through conventional economic measures. The calculation might be acceptable if there is an "acoustic separation"[17] between legal measures and private life. At least if the measures do not affect prevailing kinds of valuation and can operate autonomously in their own sphere, the law may be acceptable, even if it does not reflect an appropriate kind of valuation for other contexts.

Third, the legal system has crude remedial tools. Usually it must use monetary remedies. In view of this limitation, the fact that these remedies are not commensurable with some harms is often merely an interesting theoretical point. When someone has lost an arm, or when a river has been polluted, the legal system has to work with money. The tools of the legal system lack sufficient refinement fully to take account of diverse kinds of valuation. What can be said about personal valuations cannot be said about legal institutions. It may follow that public policy ought to operate on the assumption of commensurability even if human beings should not; or it may follow that the legal system must often put problems of incommensurability to one side, leaving those problems for ethics rather than for law.

These disclaimers are important. Nonetheless, the fact of incommensurability, and the existence of diverse kinds of valuation, do help illuminate a wide range of disputes about the substance of law, about legal institutions, and about legal reasoning. In the next section, I offer several examples of pertinent legal debates.

The Expressive Function of Law

A unifying theme for the discussion is *the expressive function of law*. When evaluating a legal rule, we might ask whether the rule expresses an appropriate valuation of an event, person, group, or practice. The point matters for two reasons. The first and most important is based on a prediction about the facts: An incorrect valuation may influence social norms and push them in the wrong direction. For example, if the law wrongly treats something—say, reproductive capacities—as a commodity, the social kind of valuation may be adversely affected. If the law mandates recycling, subsidizes national service, or requires mandatory pro bono work, it may have healthy effects on social valuations of the relevant activities. It is appropriate to evaluate the law on this ground.

We can go further. Some people seem to think that it is possible to assess law solely on the basis of consequences—that an open-ended inquiry into consequences is a feasible way of evaluating legal rules. But this is not actually possible. The effects of any legal rule can be described in an infinite number of ways. Any particular characterization or accounting of consequences will rest not on some specification of the brute facts; instead, it will be mediated by a set of (often tacit) norms determining how to describe or conceive of consequences. Part of the expressive function of law consists in the identification of what consequences count and how they should be described. Because any conception of consequences is interpretive and thus evaluative in character, "pure" or unmediated consequentialism is not a feasible project for law.

A description of the effects of some legal rule is a product of expressive norms that give consequences identifiable social meanings, including norms that deny legal significance to certain consequences. We can therefore see the expressive function both in the effects of law on social attitudes and in the use of law to decide what sorts of consequences matter for legal purposes. When

it seems as if we can talk about consequences alone, it is only because the mediating expressive norms are so widely shared that there is no controversy about them.

I have emphasized the possibility that the kinds of valuation reflected in law will affect social valuations in general. Sometimes this claim is right; valuations, like preferences and beliefs, are not a presocial given, but a product of a complex set of social forces, including law. But sometimes law will have little or no effect on valuations. Society is filled with legal provisions for market exchanges of goods and services—like pets and babysitting, for example—that are valued for reasons other than use. Market exchange need not affect social valuations; certainly, intrinsic goods are purchased and sold. The question therefore remains whether the claimed effect on social norms will occur. It is fully plausible, for example, to say that, although a law that permits prostitution reflects an inappropriate valuation of sexuality, any adverse effect of the law on social norms is so small as to be an implausible basis for objection.

But there is a second ground for endorsing the expressive function of law, and this ground does not concern social effects in the same sense. The ground is connected with the individual interest in integrity. Following the brief but suggestive discussion by Bernard Williams,[18] we might say that personal behavior is not concerned solely with producing states of affairs, and that, if it were, we would have a hard time in making sense of important aspects of our lives. There are also issues involving personal integrity, commitment, and the narrative continuity of a life. Williams offers several examples. Someone might refuse to kill an innocent person at the request of a terrorist, even if the consequence of the refusal is that many more people will be killed. Or a pacifist might refuse to take a job in a munitions factory, even if the refusal will have no salutary effects.

Our responses to these cases are not adequately captured in purely consequentialist terms.[19] Now it is possible that, for example, the refusal to kill an innocent person is consequentially justified on balance, for people who refuse to commit bad acts may cultivate attitudes that lead to value-maximizing behavior. But this is a complex matter. My point is only that consequentialist accounts do not fully describe our evaluative attitudes toward such acts.

Moreover, the expression of the appropriate evaluative attitude should be understood as a human good, constitutive of desirable characteristics. By making certain choices and not others, people express various conceptions both of themselves and of others. This is an independently important matter. We should agree on this point even if we also believe that consequences count (mediated as they are by expressive norms), and that people ought not to be fanatical.

There is a rough analog at the social and legal levels. A society might identify the kind of valuation to which it is committed and insist on that kind, even if the consequences of the insistence are obscure or unknown. A society might, for example, insist on an antidiscrimination law for expressive reasons, even if it does not know whether the law actually helps members of minority groups. A society might protect endangered species partly because it believes that the protection makes best sense of its self-understanding, by expressing

an appropriate valuation of what it means for one species to eliminate another. A society might endorse or reject capital punishment because it wants to express a certain understanding of the appropriate course of action after one person has taken the life of another.

This point bears on the cultural role of adjudication and especially of Supreme Court decisions. The empirical effects of those decisions are highly disputed. When the Supreme Court says that capital punishment is unconstitutional, that segregation is unlawful, that certain restrictions on hate speech violate the First Amendment, or that students cannot be asked to pray in school, the real-world consequences may be much smaller than we think. But the close attention American society pays to the Court's pronouncements may well be connected with their expressive or symbolic character. When the Court makes a decision, it is often taken to be speaking on behalf of the nation's basic principles and commitments. This is a matter of importance quite apart from consequences, conventionally understood. It is customary and helpful to point to the Court's educative effect. But perhaps its expressive effect, or its expressive character, better captures what is often at stake.

I do not claim that the expressive effects of law, thus understood, are decisive or that they cannot be countered by a demonstration of more conventional bad consequences. As noted earlier—and it is important here—we might insist on a sort of acoustic separation between the domain of law and other spheres, hoping and believing that the kind of valuation appropriate to government will not affect the generally prevailing kind. If OSHA engages in cost-benefit analysis, surely people will not begin to think of their spouses as commodities. But I do suggest that the expressive function is a part of political and legal debate. Without understanding the expressive function of law, we will have a hard time in getting an adequate handle on public views with respect to civil rights, prostitution, the environment, endangered species, capital punishment, and abortion.

V. Law in Particular

I now discuss a number of areas of law in which kinds of valuation are at stake. In some of these areas, issues of commensurability also move to the fore. Of course, there are important differences among these areas, each of which raises distinctive issues of its own. Some of the issues are connected to problems of individual or social choices amid scarcity; others are not. But I believe that it is impossible to obtain a full understanding of these areas of law without reference to ideas of a general sort. In this sense, these very diverse areas share a common subject. They are united by the presence of important questions about appropriate kinds of valuation and about commensurability.

Social Differentiation

A liberal society allows a high degree of social differentiation. It includes the political sphere, the family, markets, intermediate organizations—especially

religious organizations—and much more. Michael Walzer's influential book[20] offers an instructive discussion of these different "spheres." It would be especially valuable to be able to understand the social function or purpose of this sort of differentiation. Why might it be a good thing to carve up life in this way? And what exactly is the role of law in this endeavor?

The liberal commitment to social differentiation can be understood as an effort to make appropriate spaces for different kinds of valuation. Without accepting naive conceptions of the public-private distinction, we can see the family as the characteristic liberal sphere for the expression of love. At its best, politics embodies the forms of respect entailed by processes of reason-giving. In many intermediate organizations, people express affection and admiration. In religious organizations, the prevailing kind of valuation is usually one of worship and reverence. The market is typically the sphere for use. Things bought and sold on markets are typically valued in the way associated with pure commodities, although it is important and also true that many things sold on markets—music, vacations in beautiful places, art, childcare—are valued for reasons other than simple use.

Law plays an important and often overlooked role in the construction and maintenance of these various spheres, which are anything but natural. We know far too little to say that, in the state of nature, there is any such division. Markets are of course a function of the law of property, contract, and tort, without which voluntary agreements would not be possible. It is law that decides what can be traded on markets, and how trades can occur. Undoubtedly, families of various sorts would arise in the state of nature, but the particular families we have are emphatically a function of law. The law helps to create an independent familial sphere; it also determines who may be entitled to its protections and disabilities.

The state also insulates the public sphere from civil society, through, for example, the rule of one person, one vote and limitations on expenditures on campaigns. Intermediate organizations, like religious groups and labor unions, also receive various protections, insulations, and disabilities by laws designed to recognize independent social spheres. The creation of diverse social spheres, understood as a mechanism for allowing diverse kinds of valuation, is also an important social good, providing a form of liberty that is indispensable in modern society.

It would be foolish to idealize current practices. Institutional practice often deviates from institutional aspiration. Politics is often a realm for use; it has important marketlike features. It is hardly a simple process of exchanging reasons. The family is not only a place for the expression of love. Women and children have often been used without their consent, sometimes like commodities or objects (consider domestic abuse). People who believe in different kinds of valuation do not deny the gap between current practice and current aspiration. Much social criticism consists of an insistence on that gap, and of a claim for reform in the interest of producing conformity to the aspiration.

Religion, Civic Equality, and Political Liberalism

Our system is one of liberal republicanism. As such, it is committed to a principle of political equality and to a certain view about the relationship between political life and religious conviction. We can make some progress on these abstractions by observing that a liberal republic attempts to exclude certain kinds of valuation from public life. Those kinds of valuation may be too sectarian, or they may be inconsistent with the premise of political equality. A liberal republic thus bans particular "inputs" into politics either because they express a kind of valuation that is suited only to private life, or because they deny political equality, a commitment that entails certain evaluative attitudes toward fellow citizens. What I want to emphasize is that we can get a distinctive purchase on certain constitutional issues in this light.

Under the Constitution, the Establishment Clause is the key example of the ban on certain kinds of valuation. The Constitution rules out valuations that assume certain conceptions of what is sacred, at least those that invoke religious commitments. For example, a law making Easter a national holiday may not rest on the ground that it reflects the sanctity of Jesus Christ. The Establishment Clause generally rules sectarian justifications for statutes out of bounds, even if other, neutral factors could support the same laws. Political liberalism is constituted in part through the idea that certain kinds of valuation are too contentious to be a legitimate part of public life, even if they are a fully legitimate part of private citizenship.

A liberal republic also excludes certain kinds of valuation of human beings as inconsistent with constitutionally prescribed norms of political equality. I have referred to the prohibition on "titles of nobility," a key to the American tradition in this regard. The Equal Protection Clause is now understood very much in these terms. Thus, for example, a law cannot be premised on the ground that blacks are inferior to whites, or women inferior to men.[21] This idea may help make sense of the controversial claim that discriminatory intent is a necessary and sufficient condition for invalidating legislation.

Rights

In an influential if extreme formulation, fomer Judge Robert Bork asks how the Court can protect the right to sexual privacy if it does not protect the right to pollute. On Bork's view, both rights amount to "gratifications," and there is no essential difference between them.[22] A less extreme version of this view is the concern whether the Court can consistently protect sexual privacy without protecting property interests; who is to say that the former is more important than the latter?

We might be able to handle this question better if we insist on qualitative rather than merely quantitative differences among rights. It is not necessarily true that someone will value sexual privacy "more" than continued employment for longer hours at better wages, but it is possible to have a distinct concern about government regulation of the former, and it may be because of

the way that sexual privacy is valued that constitutional protection is appropriate. Once claims for constitutional protection are not aligned along a single scale and assessed in purely quantitative terms, it becomes easier to make progress in thinking about fundamental rights.

This general understanding may cast light on some familiar and somewhat mysterious notions in legal and social theory. Consider, for example, Ronald Dworkin's account of rights as "trumps,"[23] or John Rawls's claim that, under the appropriate theory of justice, the principle calling for equal basic liberty is lexically prior to better economic arrangements.[24] If we treat rights as "trumps," we may be taken to be saying not that they are infinitely or even extraordinarily valuable when viewed solely in terms of aggregate levels, but that they are valued in a distinctive way—a way quite different from, and qualitatively higher than, the way we value the competing interests. Because of the distinctive way that rights are valued, it is necessary that the competing interests be (a) of a certain qualitative sort and (b) extraordinary in amount or level, in order to count as reasons for abridgment.

Similarly, in speaking of the lexical priority of equal liberty, we may mean that a just society values this interest in a way that precludes its violation for social and economic advantages. The lexical priority of liberty thus represents an effort to restate Kantian ideals about the priority of equal dignity and respect. We can imagine cases in which this judgment could be fanatical. But it seems clear that, in some cases, a belief in lexical priority may well reflect claims about incommensurability. We might treat equal liberty as a reflection of the foundational commitment to equal dignity and respect and believe that we do violence to the way we value that commitment if we allow it to be compromised for the sake of greater social and economic advantages. On this view, the lexical priority of equal liberty is structurally akin to the refusal to allow a child to be traded for cash, or a promise to be breached as a result of mildly changed circumstances. It reflects a judgment that the prevalent kind of valuation forbids compromising the good for certain reasons, even though those reasons are legitimate bases for action in other contexts.

Blocked Exchanges

An understanding of diverse kinds of valuation helps explain why liberal regimes generally respect voluntary agreements. If people value things in different ways, the state should allow them to sort things out as they choose. If values were commensurable, perhaps we could seek to block certain voluntary exchanges simply because we had better information about relevant costs and benefits than the parties themselves. In the face of diverse kinds of valuation, it is best to permit people to value as they like.[25]

But even a system that generally respects freedom of contract may block exchanges on several grounds. Typically, such grounds involve some form of market failure: A party may lack relevant information, a collective action problem may exist, third parties may be affected, a party may be myopic. But an additional and distinct reason is that some exchanges involve and encourage improper kinds of valuation.

I think that more common arguments stressing distributive considerations or unequal bargaining power often depend, on reflection, on an unarticulated claim about inappropriate valuation. The claim is that to allow purchase and sale of a good will mean that the good will be wrongly valued in a qualitative sense. Consider, for example, the right to vote. Vote trading is objectionable in part because it would allow inappropriate concentration of political power in the hands of a few. The prohibition therefore overcomes a collective action problem. But perhaps the ban on vote-trading also stems from a concern about kinds of valuation. If votes were freely tradable, we would have a different conception of what voting is for—about the values that it embodies—and this changed conception would have corrosive effects on politics.

Some reactions to Judge Richard Posner's argument on behalf of a form of "baby selling" stem from a similar concern. Judge Posner contends that a market for babies would serve most of the relevant policies better than does the current system. In some ways his argument is persuasive. Certainly the desire of infertile couples for children would be better satisfied through a market system. But part of the objection to free markets in babies is not quite engaged by Judge Posner. Instead, the objection is that a system of purchase and sale would value children in the wrong way. This system would treat human beings as commodities, a view that is itself wrong, and a practice with possible harmful consequences for social valuation in general. At most, this is a summary of a complex argument, based partly on uncertain empirical judgments, and it is hardly by itself decisive. But we cannot get an adequate grasp on the problem without seing this concern.

Or consider a possible application of the widely held view that market incentives are preferable to command-and-control regulation (a view endorsed in chapter 13). Might it not be preferable, for example, to allocate tradable racial discrimination rights, so that discriminators, or people who refuse to act affirmatively, can purchase rights from people who do not discriminate or who engage in affirmative action?[26] Suppose it could be shown that an approach of this sort would be more efficient than the approach in the current civil rights laws, and also that it would produce equal or better outcomes in terms of aggregate hiring of minority group members. Even if this were so, it might be thought unacceptable to permit employers to discriminate for a fee, because the way that we do or should value nondiscrimination is inconsistent with granting that permission. This argument might be educative: If discriminators could buy the right to discriminate, perhaps discrimination would not be stigmatized in the way we want. Or it might be expressive: Perhaps society would like to condemn discrimination quite apart from consequential arguments.

In coming to terms with the issue of blocked exchanges, some further points are important. As noted, the recognition of diverse kinds of valuation might well be made part of a conception of law that places a high premium on individual liberty and choice. At least as a general rule, the state should not

say that one kind of valuation will be required. Within broad outlines, people ought to be permitted to value as they wish, though we should recognize that any kind of valuation will inevitably be affected by social norms and by law.

Moreover, there is an important difference between urging changes in social norms and urging changes in law. There may be a justified norm against asking a neighbor to mow one's lawn without a law to that effect. Similarly, one might speak on behalf of a powerful norm against certain attitudes toward animals without also recommending laws that, for example, forbid people to eat meat. It is possible to urge people to value animals in a certain way without thinking that the law should require that kind of valuation.

We might understand the distinction between norms and laws in two different ways. First, the distinction might be largely pragmatic. A law imposed on people who do not share the relevant norm may breed frustration and resentment. It may be counterproductive or futile. Second, the difference might be rooted in a principle of liberty, one that calls for a strong presumption in favor of governmental respect for diverse kinds of valuation. A legal mandate may be simply too sectarian or intrusive for a heterogeneous society. Both pragmatism and principle seem to argue in favor of a sharp distinction between norms and law.

There is a final point. Markets are filled with agreements to transfer goods that are not valued simply for use. People purchase music, even if they regard the performances as deserving awe and wonder. They buy human care for their children. They trade the right to see beautiful areas. They purchase pets for whom they feel affection or even love, and whom they hardly regard as solely for human exploitation and use. The objection to the use of markets in certain areas must depend on the view that markets will have adverse effects on existing kinds of valuation, and it is not a simple matter to show when and why this will be the case. For all these reasons, opposition to commensurability, and insistence on diverse kinds of valuation, do not by themselves amount to opposition to market exchange, which is pervaded by choice among goods that participants value in diverse ways.

Legal Reasoning

A belief in diverse kinds of valuation has consequences for current debates about the actual and appropriate nature of legal reasoning. If it can be shown that a well-functioning system of law is alert to these diverse kinds of valuation, theories of legislation, administration, and adjudication may be affected.

Consider, for example, economic analysis of law. It is clear that this form of analysis has produced enormous gains in the positive and normative study of law. The approach has been criticized on many grounds, most familiarly that it is insufficiently attuned to distributive arguments. But perhaps the resistance to economic analysis stems from something quite different and less noticed. In its normative form, economic analysis depends on too thin, flat, and sectarian a conception of value, captured in the notion that legal rules

should be designed so as to maximize wealth. The problem with this idea is that the word "wealth" elides qualitative distinctions among the different goods typically at stake in legal disputes. Instead of *maximizing* wealth, it is desirable to have a highly disaggregated picture of the consequences of legal rules, a picture that enables the judge to see the various goods at stake. This is true not only of the law of free speech and religious liberty, but also of the law of contract and tort. At least under ideal circumstances, it would be good to have a full sense of the qualitatively distinct interests at stake before reaching a decision.

We still lack a full account of practical reason in law. But what I have said here offers some considerations in support of analogical reasoning, the conventional method of Anglo-American law. Analogical reasoning, unlike economic analysis, need not insist on assessing plural and diverse social goods according to a single metric. The analogical thinker may be alert to the manifold dimensions of social situations and to the many relevant similarities and differences. In picking out relevant similarities, the analogizer does not engage in an act of deduction. Instead, she identifies common features in a way that helps constitute both legal and cultural categories, rather than being constrained by some particular theory of value given in advance. Unequipped with or unburdened by a unitary theory of the good or the right, she is in a position to see clearly and for herself the diverse and plural goods that are involved and to make choices among them without reducing them to a single metric. This is one description of the characteristic style of Anglo-American law in its idealized form.

A distinctive feature of analogical thinking is that it is a "bottom-up" approach, building principles of a low or intermediate level of generality from engagement with particular cases. In this respect, it is quite different from "top-down" theories, which test particular judgments by reference to general theory.[27] Because analogy works from particular judgments, it is likely to reflect the plural and diverse goods that people really value. Of course, there are nonanalogical approaches that insist on the plurality of goods, and those approaches also have a "top-down" character.

This approach has disadvantages as well. The use of a single metric makes things simple and orderly where they would otherwise be chaotic. In certain areas of the law, this may be a decisive advantage, all things considered. Certainly, it has been taken as an important advantage in many areas of contemporary law, in which the displacement of common law courts by legislation has been designed to promote rule-of-law virtues inaccessible to ordinary adjudication.

VI. Choices Again

I have not yet dealt with two major questions. The first involves the choice among different kinds of valuation and, more particularly, the decision whether or not to make things commensurable. The second involves the issue of how to make choices among incommensurable goods. The two questions

are obviously related. They are also large and complex. To answer them, we would need to answer many of the major questions in ethical and political theory. As I have emphasized, most answers must be developed in the context of particular problems. I offer only a few notations here.

Which Kinds of Valuation?

To come to terms with the question of appropriate kinds of valuation, we should begin by asking whether diversity in kinds is desirable at all. Perhaps we should adopt a unitary kind of valuation, or seek to obtain some sort of commensurability. In many ways, this project unites such diverse thinkers as Plato, Bentham, and, in law, Richard Posner. It seems clear that this approach would have many advantages. It would simplify and order decisions by placing the various goods along a single metric. Often there is indeed pragmatic value in this enterprise (see chapter 12). According to one familiar conception of rationality, this step would increase rationality. Would it not be a large improvement if commensurability could be obtained?

Let us put law to one side for the moment and think about claims for unitariness and commensurability in human life generally. It seems obvious that an answer to this question should turn on what leads to a better conception of the human good. We might make some progress by considering what the world would be like if kinds of valuation really were unitary or if commensurability really did obtain. In such a world, for example, a loss of friendship or the death of a parent would really be like a loss of money, though undoubtedly a lot of it. An achievement in something that one prizes—like art or music—would be valued in the same way as an increase in net worth, or the birth of a new child, or falling in love, or the relief of human suffering, or the victory of a favorite sports team. Offers of cash exchange would really be evaluated solely on the basis of their amount. The distinction between instrumental and intrinsic goods would collapse, in the sense that many intrinsic goods would become both fungible and instrumental.

A great deal would be lost in such a world. A life with genuine commensurability would be flat and dehumanized. It would be inconsistent with an appropriately diversified approach to a good human life. It would eliminate desirable relationships and attitudes. In fact, it would be barely recognizable.

On the other hand, nothing in this brief account is fatal to the view that law and public policy sometimes ought to rest on unitary kinds of valuation or on assumptions of commensurability. It is possible to think that under ideal conditions, the best system of valuation is diverse and plural, but that in light of the weaknesses of human institutions and the constant prospect of bias, confusion, and arbitrariness, public choices should assume a single kind or a unitary metric as the best way of promoting all of the relevant goods. This question cannot be answered a priori or in the abstract. But those who favor legal approaches based on unitary kinds of valuation and commensurability should understand that their approach is best defended as a means of overcoming certain institutional obstacles, and not as reflecting a fully adequate understanding of the relevant problems. I have suggested as well that even if, say, government officials align the diverse effects of regulation along a single metric, they should also provide a disaggregated picture, so that citizens can be aware of qualitative differences and of the various goods at stake.

All this leaves much uncertainty. But it does suggest that, institutional issues aside, we should approve of a large degree of diversity in kinds of valuation. From this it does not follow that any particular kind of valuation is appropriate for any particular sphere. To evaluate kinds of valuation, we also have to think very concretely about what kinds in what places are parts of a good life or a good political system.

Choices Among Incommensurable Goods

How are choices made among incommensurable goods? How can those choices be assessed? The first point is that there is no algorithm or formula by which to answer this question. If we are looking for a certain sort of answer—the sort characteristic of some believers in commensurability—we will be unable to find it. Relatively little can be said in the abstract. Instead, we need to offer detailed descriptions of how such choices are made, and how to tell whether such choices turn out well. Here there are many possible criteria for public and private action, and a mere reference to the existence of incommensurabilities will be unhelpful standing by itself.

In law and politics, a diverse set of standards—liberty, equality, prosperity, excellence, all of these umbrella terms—will be brought to bear on hard cases. A particular goal is to find solutions that will minimally damage the relevant goods. Perhaps an approach that promotes a recognizable form of liberty will only modestly compromise a recognizable conception of equality. Here, too, there is no escape from close examination of particular cases.

Consider, for example, the issue of surrogacy. It would be plausible to conclude that the legal system ought not to criminalize surrogacy but also ought not to require surrogate mothers to hand over their children to the purchasing parents, at least during a brief period after birth. Criminalization would entail difficult enforcement problems and might also discourage arrangements that do much good and little harm. We can understand why surrogacy arrangements might be thought to be connected with sex inequality, but the contribution to discrimination is probably too attenuated to justify criminalization. On the other hand, it may be damaging to the birth mother to force her to hand over a child against her will, especially because before the fact it may be quite difficult for her to know exactly what this action entails. The failure of the legal system to order specific performance is unlikely to have serious adverse consequences on prospective participants in such arrangements. This is, of course, an inadequate treatment of a complex problem. All I mean to do is to suggest some of the lines along which the inquiry might occur. (Compare an approach that would insist on commensurability. Such an approach would be laughably inadequate; it would prevent us from seeing what is really at stake.)

Much of the relevant work here is done in two ways: through analogies and through understanding consequences, mediated as these are through expressive norms. When incommensurable goods are at stake, it is typically asked: What was the resolution of a previous case with similar features? Through this process, people seek to produce vertical and horizontal consistency among their various judgments. This system of testing is designed not to line up goods along a single metric, but to produce the sort of consistency and rigor that characterizes the successful operation of practical reason. The inquiry into consequences avoids monism while still examining the real-world effects of different courses of action. How, for example, will a particular result compromise or promote the relevant goods? This is a characteristic part of practical reasoning in law.

In the context of emissions trading in environmental law, for example, it seems hard to support the empirical claim that the shift from command-and-control government to financial incentives will have serious adverse effects on people's thinking about pollution. Many claims about the educative function of law are actually claims about real-world consequences, understood through expressive norms on which there is no dispute. It is possible to think well about those consequences, and sometimes we can see that the feared effects will not materialize. In this way, it is possible to explore whether there will be damage to some of the goods that are allegedly threatened by certain initiatives. Of course, it is true that, on some issues, we will lack relevant data.

A Note on Tragedy

There is a final point. A recognition of incommensurability is necessary to keep alive the sense of tragedy, and in certain ways this is an individual and collective good, perhaps especially in law. Recognition that all outcomes "impose costs" is not quite the same thing as a sense of tragedy.[28] Though I can hardly discuss this complex matter here, the very notion of tragedy seems to embody a commitment to an understanding of the uniqueness of certain goods, or the irreversibility and irreplaceability of certain losses. If tragedy were understood to mean instead high costs, or the existence of losses that accompany benefits (Kaldor-Hicks rather than Pareto improvements, for example), the sense of tragedy would be dramatically changed. Here, too, we would not have a simple redescription of the problem.

A sense of tragedy is an individual good because it accompanies and makes possible certain relationships and attitudes that are an important part of a good life. It is a collective good for this reason and also because it focuses attention on the fact that, even when the law is doing the right thing, all things considered, much may be lost as well. This is valuable, for example, in current thinking about the environment and occupational safety. In the presence of tragedy, there is a large incentive to create social arrangements so that people do not face that prospect.

At its best, the Anglo-American legal system is alert to the fact that diverse goods are at stake in many disputes. Judges know that not all of these goods

can be simultaneously preserved. This awareness is itself desirable because of the pressure that it tends to exert on judges and legislators as well. A redescription of tragedy in terms that assume a monistic theory of value would not create the same sort of pressure.

Conclusion

Human beings value goods, events, and relationships in diverse and plural ways. Sometimes money itself is not fungible. Often we face serious problems of commensurability. This problem does not entail paralysis, indeterminacy, or arbitrariness. Decisions are made all the time among incommensurable goods, at the personal, social, and legal levels, and these decisions may well be rational or irrational. It might even be possible to convert our kinds of valuation for use on a unitary metric, or to make goods, events, and relationships commensurable. But if what I have said here is correct, this would be a sort of tragedy, not least because it would make the very fact of tragedy puzzling or even incomprehensible.

The fact of diverse kinds of valuation, and the existence of incommensurable goods, have not yet played a major role in legal and economic theory. But these issues underlie a surprisingly wide range of legal disputes. No unitary "top-down" theory can account for the complexities of most controversies in law. Perhaps most dramatically, the liberal insistence on social differentiation—markets, families, religious groups, politics, and more—is best justified as an attempt to make a space for distinct kinds of valuation and to give each of them its appropriate place in human life. The Establishment and Equal Protection Clauses are centrally concerned with regulating kinds of valuation. In environmental law, the major issue of contestation is frequently the appropriate kind of valuation of environmental amenities; if beaches, species, and mountains were valued solely for their use, we would not be able to understand them in the way that we now do.

The same issue arises in the law of contract, especially in the award of specific performance remedies. It also plays a role in thinking about the point of damage remedies in tort. In both settings, it is wrong to ignore the highly contextual nature of choice and to act as if a particular decision—not to take out insurance for a certain danger, to accept a job at a certain risk premium—reflects some global judgment simply adaptable for policy use. No global judgment need underlie particular choices. Many blocked exchanges, moreover, attest to social resistance to commensurability. We might also think that attention to these issues shows that cost-benefit analysis is obtuse, because it aligns qualitatively distinct goods along a single metric. In addition to cost-benefit analysis—which might of course be helpful in our at most second-best world—what is desirable is a disaggregated picture of the effects of different courses of action, so that officials and citizens can see those effects for themselves.

In disputes over free speech, large questions are whether speech ought to be valued in the same way as commodities traded on markets, and whether free speech values are unitary or plural. We might think that we ought not to

treat free speech as an ordinary commodity and that we should recognize the diverse ends it embodies. Various arguments for revision of existing understandings—market thinking for families and politics, extension of the metaphor of the family, the view that reason-giving ought to occupy all social spheres—might be seen as attempts to renovate current forms of valuation in favor of a new or even unitary kind.

An especially large task for legal theory is to offer an adequate description of how, in legal contexts, choices should be made among incommensurable goods and among different possible kinds of valuation. I have at best started to undertake that task here. There are limits to how much can be said in the abstract; a close inspection of particular contexts will be indispensable to this endeavor. But I conclude with two suggestions. An insistence on incommensurability and on diverse kinds of valuation is one of the most important conclusions emerging from the study of Anglo-American legal practice, and an appreciation of those diverse kinds will yield major gains to those seeking to understand and to evaluate both public and private law.

Notes

1. I owe much help in the discussion here to Elizabeth Anderson, *Value in Ethics and Economics* (Cambridge: Harvard University Press, 1993). Anderson uses the term "modes of valuation" to refer to the same basic idea. I should emphasize that I am discussing our best self-understandings, and that more parsimonious assumptions may be best for predictive purposes (see chapter 12).

2. For an especially instructive discussion of this point, see *id.* at 8–11.

3. Note in this regard the social norm requiring that, in some circumstances, gifts should not take the form of cash, which is regarded as excessively impersonal. As compared with a gift of (say) $30, a gift of (say) a tie costing $30 establishes a distinctive and often preferable relationship between the recipient and the giver—even if in some contexts $30 in cash would be worth more to the recipient than the $30 tie, and even if in some contexts a gift of $30 would be less costly, to the giver, than the gift of a $30 tie. This point is interestingly missed in Joel Waldfogel, "The Deadweight Loss of Christmas," 83 *Am. Econ. Rev.* 1328, 1336 (1993), which finds $4 billion in annual deadweight losses from noncash gifts and which assumes that cash gifts are always more efficient and therefore generally preferable.

4. The example comes from Joseph Raz, *The Morality of Freedom* (New York: Oxford University Press, 1986), pp. 348–49. I am generally indebted to Raz's discussion of incommensurability and especially constitutive incommensurabilities.

5. See Steven Kelman, *What Price Incentives?: Economics and the Environment* (Boston: Auburn House, 1981), pp. 54–83; Margaret J. Radin, "Market-Inalienability," 100 *Harv. L. Rev.* 1849 (1987).

6. See Richard A. Posner, *Sex and Reason* (Cambridge: Harvard University Press, 1992), pp. 409–17; Elisabeth M. Landes and Richard A. Posner, "The Economics of the Baby Shortage," 7 *J. Legal Stud.* 323 (1978); Richard A. Posner, "The Regulation of the Market in Adoptions," 67 *B. U. L.Rev.* 59 (1987).

7. See Viviana A. Zelizer, *The Social Meaning of Money* (New York: Basic Books, 1994); Steven Croley and John Hanson, "The Nonpecuniary Costs of Accidents," 108 *Harv. L. Rev.* 1785 (1995); Richard Thaler, "Mental Accounting Matters" (unpublished manuscript, 1995).

11

Paradoxes of the Regulatory State

By "paradoxes of the regulatory state," I mean self-defeating regulatory strategies—strategies that achieve an end precisely opposite to the one intended, or to the only public-regarding justification that can be brought forward in their support. An example of a regulatory paradox would be a Clean Air Act that actually made the air dirtier, or a civil rights law that increased the incidence of racial discrimination. This definition excludes a number of pathologies of the regulatory state that are clearly related to the phenomenon of regulatory paradoxes, such as strategies whose costs exceed their benefits, or that have unintended adverse consequences. I am interested only in self-defeating regulatory strategies here.

A large literature, inspired by public choice theory and welfare economics, has grown up around the theory that purportedly public-interested regulation is almost always an effort to create a cartel or to serve some private interest at the public expense. Although I shall be drawing on much of that literature here, I do not conclude, as some of that literature appears to, that the appropriate response to regulatory paradoxes is to abandon modern regulation altogether and rest content with the operation of private markets. In many cases, the market itself produces harmful or even disastrous results, measured in terms of efficiency or justice. The appropriate response to the paradoxes of regulation is not to return to a system of "laissez faire," but to learn from past failures. To this end, I outline the lessons, for legislators, judges, and administrators, that are to be drawn from the omnipresence of regulatory paradoxes. My general goal is to describe some reforms by which to restructure regulatory institutions so as to achieve their often salutory purposes, while at the same time incorporating the flexibility, respect for individual autonomy and initiative, and productive potential of economic markets.

I. The Performance of the Regulatory State: A Prefatory Note

Empirical assessments of the consequences of regulation remain in a primitive state; but it is possible to draw several general conclusions. I outline some of them here.

In many ways things are getting better, not worse. The average American's life expectancy at birth was 47 years in 1900; it rose to 78 in 1990, partly because the death rate from infectious diseases is less than 5 percent, down from 20 percent in 1900. Consider the data in Table 11.1.

Undoubtedly many of these improvements come from social changes having nothing to do with regulation. But the view that regulation has generally proved unsuccessful is far too crude. Some of the progress comes from more and better immunizations, better nutrition, and improved water and sanitation systems, some of which is a result of governmental activity and regulatory controls. In fact, environmental protection has produced enormous gains.[1] Regulatory controls have helped to produce substantial decreases in both the levels and emissions of major pollutants, including sulfur dioxide, carbon monoxide, lead, and nitrogen dioxide. Lead is an especially dangerous substance, and ambient concentrations of lead have decreased especially dramatically, declining 96 percent since 1975; transportation emissions of lead decreased from 122.6 million metric tons in 1975 to 3.5 in 1986. The EPA's phase-down of lead in gasoline was firmly supported by cost-benefit analysis. So, too, carbon monoxide levels fell 57 percent from 1975 to 1991; nixtrous oxide levels fell about 24 percent; and particulate levels fell 26 percent. Most important, the vast majority of counties in the United States are now in compliance with air quality goals. In 1982, about 100 million people lived in nonattainment areas; in 1994, the number was about 54 million.

Water pollution control has shown significant successes as well. Perhaps most important, all sewage in the nation is treated before it is discharged, and the treatment ordinarily brings water to a level safe for swimming. There are many successes with respect to particular water bodies. The Great Lakes are much cleaner than they were in 1965. Phosphorus loading is, for example, substantially down, somewhere between 40 and 70 percent less than in peak years. Levels of PCBs in the Great Lakes are almost 90 percent lower than they were in 1973.

In addition, a number of harmful nutrients have been reduced by nearly 50 percent in national rivers. In 1970, about 25 percent of U.S. river miles were safe for fishing and swimming; the number is now about 56 percent. Industrial pollutants in Chesapeake Bay fell by 53 percent in the brief period from 1988 to 1992. Governmentally required lead and nitrate reductions have produced huge improvements in water quality. Ocean dumping of sludge is now substantially eliminated in the United States. All in all, both air and water are substantially cleaner than they would have been without regulatory controls, and despite a wide range of errors, the American experience serves in some respects as a model for the rest of the world. People continue to be very much concerned about toxic waste sites. But it is notable that since the enactment of the Solid Waste Disposal Act, nearly all toxic byproducts are disposed of in safe ways, and new toxic waste "dumps" are almost entirely a thing of the past.

The risk of a fatal accident in the workplace was reduced by 50 percent between 1970 and 1990, partly because of workers' compensation, occupational safety and health, and other programs. Similarly, automobile safety regulation has significantly reduced deaths and serious injuries.[2] Automobiles are much safer for occupants, and many of the safety-improving features of automobiles are a product of regulatory controls supported by cost-benefit balancing. For example, the requirement that all cars have center-high mounted stop lamps was shown to be highly effective in reducing rear-end collisions. More generally, highway fatalities would have been about 40 percent higher in 1981 if not for governmental controls.[3] Between 1966 and 1974, the lives of about 34,000 passenger car occupants were saved as a result of occupant safety standards. The annual benefits from regulation exceed ten billion dollars. For automobile regulation, in general, the ratio of benefits to costs is extremely high. Indeed, some of the regulations pay for themselves in terms of health and related savings, and the large number of deaths actually prevented is of course a bonus.

Table 11.1. Principal Death Risk Trends

Years	Annual Rate of Increase in Death Rates: Work (per 100,000 population)	Home (per 100,000 population)	Motor Vehicle (per 100,000 population)
1930–40	−1.8	−0.2	−3.3
1940–50	−2.3	−2.2	−4.0
1950–60	−2.8	−2.1	−3.5
1960–70	−1.2	−1.7	−0.8
1970–80	−1.6	−2.7	−3.4
1980–90	−3.2	−2.4	−4.3

Source: W. Kip Viscusi, *Fatal Tradeoffs: Public and Private Responsibilities for Risk* (New York: Oxford University Press, 1992), p. 285. Reprinted by permission.

More broadly, studies of the costs and benefits of regulatory initiatives show that a number of other measures have produced health and other benefits at especially low costs. Of course, it would be foolish to suggest that government regulation in its current form passes a global cost-benefit test. There is much indeterminacy in the data, and certainly many regulatory efforts have been futile, self-defeating, counterproductive, or much worse. But every detailed study shows a number of regulations that have saved lives at comparatively low cost.

Consider, for example, W. Kip Viscusi's influential discussion, which shows that thirteen of the most important federal regulations pass a cost-benefit test.[4] These include OSHA's hazard communication regulation, saving 200 lives per year; the Consumer Product Safety Commission's unvented space heaters regulation, saving over 60 lives per year; the Federal Aviation Administration's seat cushion flammability regulation, saving 37 lives per year; OSHA's oil and gas well service regulation, saving 50 lives per year; and the National Highway Traffic Safety Administration's (NHTSA) passive restraints/belts regulation, saving no fewer than 1,850 lives per year. OSHA's regulation

of asbestos prevents an estimated 396 deaths per year, and it does so at relatively low expense. EPA's regulation of trihalomethanes saves a life at only $300,000 per year; the National Highway Traffic Safety Administration's fuel system integrity controls, also $300,000; the Consumer Product Safety Commission's (CPSC) mandatory smoke detector rule, between $0 and $85,000; NHTSA's roadside hazard removal rule, $0. Consider in this regard the good results shown in Table 11.2.

Finally, regulatory successes are not limited to the areas of safety and health. The most important civil rights initiative—the Voting Rights Act of 1965—appears to have broken up the white monopoly on electoral processes in a number of states. Only two years after passage of the Act, the number of registered blacks in the eleven southern states increased from 1.5 million to 2.8 million—an increase of nearly 90 percent. This is an impressive illustration of regulatory success.

Now turn away from voting rights and toward discrimination in general. Though the evidence is more disputed, the better view is that the Civil Rights Act of 1964 has also had important beneficial consequences. The most careful and impressive study shows that this Act had major effects on the manufacturing sector in South Carolina. "Suddenly in 1965 the black share in employment begins to improve when Title VII legislation becomes effective and the Equal Employment Opportunity Commission begins to press textile firms to employ blacks and when Executive Order 11246 forbids discrimination by government contractors at the risk of forfeit of government business."[5] Many other studies reach a broadly similar conclusion: The federal effort to redress race discrimination has had many favorable consequences. There have been gains in the area of sex discrimination as well.

Finally, the Endangered Species Act has saved a number of species from extinction and endangerment.[6] Thus the American bald eagle was, in 1963, thought to be on the verge of extinction; it is now doing much better, and in 1995 its status was shifted from "endangered" to merely "threatened." The number of gray wolves in Minnesota has grown from 500 to 1750 in a relatively short time; there are now about 10,000 adult peregrine falcons in the United States, compared with very few in the early 1970s; sea lions had declined to a population of 30,000 in 1972 and are now up to 180,000; California gray whales recovered to a population of 21,000 in 1994, when they were removed from the list under the Endangered Species Act. In fact, of the 71 species initially listed under the Act in 1973, only seven are gone, and 44 are stable or improving.

On the other hand, regulation has frequently failed. Sometimes it has imposed enormously high costs for speculative benefits; sometimes it has accomplished little or nothing; sometimes it has aggravated the very problem it was designed to solve. For example, the United States spent no less than $632 billion for pollution control between 1972 and 1985, and some studies suggest that alternative strategies could have achieved the same gains at less than one-fifth the cost. There is no doubt that environmental and safety regulation have taken an unnecessarily high toll on the productivity of the American economy.

Table 11.2. Cost-Effective Regulations

Regulation	Agency	Cost Per Premature Death Averted ($ millions 1990)
Unvented Space Heater Ban	CPSC	0.1
Aircraft Cabin Fire Protection Standard	FAA	0.1
Auto Passive Restraint/Seat Belt Standards	NHTSA	0.1
Steering Column Protection Standard	NHTSA	0.1
Underground Construction Standards	OSHA-S	0.1
Trihalomethane Drinking Water Standards	EPA	0.2
Aircraft Seat Cushion Flammability Standard	FAA	0.4
Alcohol and Drug Control Standards	FRA	0.4
Auto Fuel-System Integrity Standard	NHTSA	0.4
Standards for Servicing Auto Wheel Rims	OSHA-S	0.4
Aircraft Floor Emergency Lighting Standard	FAA	0.6
Concrete and Masonry Construction Standards	OSHA-S	0.6
Crane Suspended Personnel Platform Standard	OSHA-S	0.7
Passive Restraints for Trucks and Buses (proposed)	NHTSA	0.7
Side-Impact Standards for Autos (dynamic)	NHTSA	0.8
Children's Sleepwear Flammability Ban	CPSC	0.8
Auto Side Door Support Standards	NHTSA	0.8
Low Altitude Windshear Equipment and Training Standards	FAA	1.3
Electrical Equipment Standards (metal mines)	MSHA	1.4
Trenching and Excavation Standards	OSHA-S	1.5
Traffic Alert and Collision Avoidance (TCAS) Systems	FAA	1.5
Hazard Communication Standard	OSHA-S	1.6
Side-Impact Standards for Trucks, Buses, and MPVs (proposed)	NHTSA	2.2
Grain Dust Explosion Prevention Standards	OSHA-S	2.8
Rear Lap/Shoulder Belts for Autos	NHTSA	3.2
Standards for Radionuclides in Uranium Mines	EPA	3.4
Benzene NESHAP (original: Fugitive Emissions)	EPA	3.4

Source: Stephen G. Breyer, *Breaking the Vicious Circle: Toward Effective Risk Regulation* (Cambridge: Harvard University Press, 1993), pp. 24–27. Reprinted by permission.

Studies show reduced output in the economy as a whole as 2.5 percent, with a cumulative impact through 1986 of 11.4 percent from the level that would have been achieved without OSHA and EPA regulation.[7]

The fuel economy standards for new cars appear to have produced no substantial independent gains in fuel economy; consumer demands for fuel efficient cars, in response to gas shortages and high gas prices, were leading in the same direction. Worse, the fuel economy standards have led manufacturers to produce smaller, more dangerous cars; an estimated 2200–3900 mortalities were expected over a ten-year period as a result of regulatory changes in 1989 alone.[8] There is little question that the administration of the Natural Gas Act helped produce the energy crisis of the late 1970s—with huge attendant costs to investment and employment—by artificially restraining the price of gas.[9]

Some of OSHA's carcinogen regulations impose enormous costs for uncertain gains. Indeed, the pattern of OSHA regulation of carcinogens is a crazy quilt; regulations costing up to $40 million per life saved exist in some areas, with no regulations at all in others (Table 5.4). By delaying the entry of beneficial drugs into the market, the Food and Drug Administration has, in many settings, dramatically increased risks to life and health.[10]

Ironically, a large source of regulatory failure in the United States is the use of Soviet-style command-and-control regulation, which dictates, at the national level, technologies and control strategies for hundreds, thousands, or millions of companies and individuals in a nation that is exceptionally diverse in terms of geography, costs and benefits of regulatory controls, attitudes, and mores. A valuable perspective on this problem can be obtained by examining the paradoxes of regulation, which pose a particular dilemma for the administrative state. A government that eliminated self-defeating regulatory strategies would eliminate a significant source of regulatory failure. And although the paradoxes are numerous, six of them have been of major importance in the last generation.

II. The Paradoxes

I have defined a regulatory paradox as a self-defeating regulatory strategy; but whether a strategy is self-defeating depends on how its purposes are described. Any statute that fails to produce a net benefit to society can be described as self-defeating if its purpose is described as the net improvement of the world. But if the statute's purpose is to benefit a particular group or segment of society, and that purpose is achieved, then the statute is not self-defeating at all. For example, a statute benefiting the agricultural industry at the expense of the public will not be self-defeating if its purpose is described as helping farmers. Throughout this discussion I describe the relevant statutory purposes as public-regarding rather than as benefiting special interest groups.

Moreover, I mean to assess whether a statute is self-defeating by comparing the result it has produced with the likely state of affairs had Congress either enacted a different and better statute or enacted no statute at all. Measured against either of these benchmarks, regulation has produced a wide range of paradoxes.

Importantly, nearly all of the paradoxes are a product of the government's failure to understand how the relevant actors—administrators and regulated entities—will adapt to regulatory programs. The world cannot be held constant after regulations have been issued. Strategic responses, the creation of perverse incentives for administrators and regulated entities, unanticipated changes in product mix and private choice—these are the hallmarks of the paradoxes of the regulatory state. The adoption of strategies that take account of these phenomena would produce enormous savings in both compliance costs and safety and health gains.

Paradox 1: Overregulation Produces Underregulation

The first paradox is that especially aggressive statutory controls frequently produce too little regulation of the private market. This surprising outcome arises when Congress mandates overly stringent controls, so that administrators will not issue regulations at all or will refuse to enforce whatever regulations they or Congress have issued.

The imposition of extremely stringent controls on regulated industries is a common strategy in Congress. Such controls typically ban cost-benefit balancing or indeed trade-offs of any sort. The expectation is that these controls will bring about safety in the workplace, or clean air and water, even if both the agency and industry are reluctant to act, and even if the costs of regulation are high. This strategy was especially popular during the dramatic growth of regulation in the 1960s and 1970s. It both fueled and was fueled by the notion that a safe workplace, or clean air and water, should be treated as involving a right to be vindicated rather than a risk to be managed. Consider President Nixon's proclamation: "Clean air, clean water, open spaces—these should again be the birthright for every American."[11] This form of rights-based thinking was also inspired by evidence that recalcitrant agencies, suffering from inertia or immobilized by the power of well-organized private groups, frequently refused to enforce regulatory controls.

The strategy of imposing stringent regulatory controls or banning cost-benefit balancing is not impossible to understand. It might seem natural to think that if air pollution is a severe problem, the correct response is to reduce it as much as possible; and this idea quickly translates into a command to the EPA to reduce dangerous substances in the atmosphere to a level that will not adversely affect human health. Certainly this seems like a natural thought if health or safety is conceived as a right. Similarly, an obvious method for controlling toxic substances in the workplace is to tell OSHA to eliminate these substances "to the extent feasible." Such strategies might produce too much regulation, but this might be thought a small price to pay for (finally) reducing pollution in the air or deaths in the workplace. In addition, a prohibition on "balancing" might be thought desirable by those fearful that any effort to balance would be distorted by the enforcement agency's undervaluation of life and health, especially in the context of seemingly permanent political divisions between the executive and the legislative branches of government.

But consider the record of both the EPA and OSHA in these settings. Of the several hundred toxic substances plausibly posing significant risks to human health, the EPA had for a long period regulated only seven—five as a result of court orders. The Clean Air Act's severe provisions for listed pollutants, operating in rule-like fashion, led the Environmental Protection Agency to stop listing pollutants at all. Thus, the Act's strict duties of response to danger prompted officials not to recognize the dangers in the first instance. Of the many toxic substances in the workplace, OSHA had for a long time controlled only ten. Stunningly, this is so even though the private organization

that once performed some of OSHA's functions had recommended lower exposure limits for hundreds of chemicals.[12] To be sure, those substances that EPA and OSHA regulate are stringently controlled. The regulatory pattern, however, includes not only substantial overregulation of the substances that are subject to federal standards, but also, and possibly more serious, substantial underregulation of dangerous substances, such as chromium, perchloroethylene, and trichloroethylene.

Despite the stringency of statutory standards, many activities in the United States are entirely free from regulatory controls. There is no evidence that the United States generally does a better job than England in protecting workers and citizens from occupational and environmental hazards, even though the English system consciously allows balancing in most contexts and the American system consciously rejects it.[13] Compare in this regard the pattern under rigid criminal statutes, including the "three strikes and you're out policy" for felons and mandatory minimum sentences. These have often produced underregulation by encouraging prosecutors to indict for lesser charges or by leading prosecutors not to enforce at all.

Statutes containing stringent regulatory requirements have thus yielded no protection at all in many settings. What is responsible for this astonishing outcome? It might be tempting to find answers in the power of regulated industries or in the intransigence and deregulatory zeal of government officials. But the pattern of underregulation can be found in the Carter Administration as well as the Reagan Administration, even though President Carter's appointees, drawn in large number from the consumer and environmental movements, were hardly eager to prevent the government from curbing the proliferation of toxic substances. Elaborate and costly procedural requirements for the promulgation of federal regulations undoubtedly provide some explanation, since the process, including judicial review, has built into it enormous delays and perverse incentives. These requirements surely slow down and deter rule making. Industry has every opportunity and every incentive to fend off regulation by making plausible claims that additional information is necessary before regulation can be undertaken. Here we have a subparadox that might explain the overregulation–underregulation paradox: Procedural protections designed to improve agency rule making may lead to little or no agency rulemaking. This explanation is not in itself adequate, however, because organized interests have not prevented agencies from being far more aggressive in other settings.

A large part of the explanation lies in the stringency of the regulatory standard itself. A stringent standard—one that forbids balancing or calls for regulation to or beyond the point of "feasibility"—makes regulators reluctant to act. If, as is customary, regulators have discretion not to promulgate regulations at all, a stringent standard will provide them with a powerful incentive for inaction. Their inaction is not caused by venality or confusion. Instead, it reflects their quite plausible belief that the statute often requires them to regulate to an absurd point. If regulators were to issue controls under the statute, government and private resources would be less available to control other toxic substances; domestic industry costs would increase; ultimately, industries

competing in world markets would face a serious risk of shutdown. Under these circumstances, a stringent standard will mobilize political opposition to regulation from within and without government. It will also increase the likelihood of judicial invalidation. Finally, it will require agencies to obtain greater supporting information to survive political and judicial scrutiny, while at the same time making it less likely that such information will be forthcoming from regulated class members. All the incentives are therefore in the direction of issuing fewer regulations.

It is thus unsurprising that a draconian standard produces underregulation as well as overregulation. A crazy quilt pattern of severe controls in some areas and none in others is the predictable consequence of a statute that forbids balancing and trade-offs (Table 5.3 in chapter 5).

The problem goes deeper still. Even if the resistance of the agency has been overcome, and some or many regulations have been issued under a statute calling for stringent regulatory controls, the risk of underregulation does not disappear. Levels of enforcement—inspections and fines—will reflect the agency's reluctance. This has in fact been the pattern with OSHA's safety and health regulations, some of which have been effectively unenforced by Democratic as well as Republican administrations. This, then, is the first paradox of the regulatory state: Stringent regulatory standards produce underregulation.

Paradox 2: Stringent Regulation of New Risks Can Increase Aggregate Risk Levels

Congress is often presented with a risk or problem that can be found in both existing entities and in potential new entrants. For example, automobiles produce carbon monoxide; modern electricity plants emit sulfur dioxide; many existing buildings are inaccessible to the handicapped; and drugs currently on the market pose health hazards to consumers. In such situations, a common strategy has been to impose especially severe limitations on new sources but to exempt old ones. Indeed, such exemptions might be a political prerequisite for enactment of the regulation, since existing companies might otherwise fight very hard against enactment. Congress might require that new automobiles be equipped with pollution control devices, that new plants emitting pollution meet stringent regulatory controls, that new buildings be accessible to the handicapped, and that new drugs survive special safety requirements.

This strategy is a pervasive one in current regulatory law, and it has obvious advantages.[14] Retroactive application of regulatory requirements can be extremely costly; the expense of altering existing practices is often high. Requiring the specified approach only prospectively can achieve significant savings. In addition, it may seem unfair to impose costs on people who would have ordered their affairs quite differently had they been informed beforehand of the regulatory regime.

As a control technique, however, the strategy of imposing costs exclusively on new sources or entrants can be self-defeating. Most important, it will dis-

courage the addition of new sources and encourage the perpetuation of old ones. The problem is not merely that old risks will continue, but that precisely because of regulatory programs, those risks will become more common and last longer—perhaps much longer—than they otherwise would.

Two different phenomena underlie the old risk–new risk paradox. First, those who plan regulatory programs often assume that the programs will not influence private choices. Private choices are, however, a function of current supply and demand. If the program raises the price of new products, it will shift choices in the direction of old risks. Second, a focus on new risks reduces the likely entry of potentially superior sources or technologies and thus perpetuates old ones. Regulatory controls eliminate possibilities that might have turned out to be substantially safer than currently available options. The result is to increase the life of those options.

Examples are not difficult to find. The EPA's program requiring the installation of antipollution technology in new automobiles belongs in the first category. This program has prolonged the use of old, dirty vehicles, retarding the ordinary, salutary retirement of major sources of environmental degradation. Command-and-control regulation of new pollution sources creates incentives to use existing facilities longer, with harmful consequences for the environment. Imposition of high, safety-related costs on new airplanes may well encourage airlines to retain (and repair) old, risky planes.

We might put the EPA's requirement of costly "scrubbing" strategies for new sources of sulfur dioxide in the second category. This rule has perpetuated the existence of old sources of sulfur dioxide, thus aggravating in many parts of the country the very problem it was designed to solve.[15] So too, the imposition of stringent barriers to nuclear power plants has perpetuated the risks produced by coal, a significantly more dangerous power source. And perhaps worst of all, the FDA's stringent regulatory standards for approving new drugs have forced consumers to resort to old drugs, which are frequently more dangerous, more costly, or less beneficial than the new drugs being kept off the market.

A final example of the old risk–new risk paradox is the Delaney Clause, which prohibits manufacturers from using food additives containing carcinogens. Ironically, this provision has probably increased safety and health risks. The Clause forces manufacturers to use noncarcinogenic, but sometimes more dangerous, substances. In addition, it makes consumers resort to substances already on the market that often pose greater risks than new entrants would. Since the newest and best detection equipment is used on proposed new additives, the statutorily prohibited additive may well pose fewer risks to consumers than substances already on the market that were tested with cruder technology. The Delaney Clause defeats its own purpose.[16]

The phenomenon of careful regulation of new risks and lenient or no regulation of old ones may not simply reflect legislative myopia or confusion. Public choice theory provides a plausible explanation for the phenomenon. A system of regulation that imposes controls solely on new products or facilities should have considerable appeal for those in possession of old ones. If new

sources will face regulatory costs, the system of government controls will immunize existing producers from fresh competition. Indeed, the regulatory statute will create a partial cartel, establishing a common interest among current producers and giving them a significant competitive advantage over potential new entrants. The victims of the old-new division, however, often do not yet exist. They are usually hard to identify, do not perceive themselves as victims, and are not politically organized.

It may be for this reason that the careful regulation of new risks is such a popular strategy. It is apt to be favored both by existing industry and by many of those who seek to impose controls in the first instance. The potential victims—consumers and new entrants—often have insufficient political strength to counter the proposals. When this phenomenon is combined with the apparently sensible but sometimes self-defeating idea that a phase-in strategy is better than one that requires conversions of existing producers, it is no surprise that the old risk–new risk division remains so popular. It is a perverse strategy, but it is likely to continue, at least so long as there is no shift from command-and-control regulation to economic incentives.

Paradox 3: To Require the Best Available Technology Is to Retard Technological Development

Industry frequently fails to adopt the best technology for controlling environmental or other harms. The technology exists or can be developed relatively cheaply, but polluters simply refuse to use it. Congress and the EPA have often responded by requiring that all industries use the best available technology (BAT). The BAT strategy is pervasive in federal environmental law and may indeed be its most distinctive characteristic.

The BAT strategy is motivated by a desire to produce technological innovation, and here it has a surface plausibility. As discussed previously, recent years have witnessed large decreases in air and water pollution, and these decreases are partly attributable to the use of emission control technologies. requiring the adoption of the best available control technology seems a sensible way to ensure that all industries are doing their utmost to prevent pollution. This strategy also appears inexpensive to enforce. The government simply decides on the best technology and then requires all industries to comply.

The BAT approach, however, can defeat its own purposes and thus produce a regulatory paradox. It is an extremely clumsy strategy for protecting the environment. To be sure, the approach is a plausible one if the goal is to ensure that all firms use currently established technology. But a large goal of regulation should be to promote technological innovation in pollution control. Regulation should increase rather than decrease incentives to innovate. Government is rarely in a good position to know what sorts of innovations are likely to be forthcoming; industry will have a huge comparative advantage here. Perversely, requiring adoption of the BAT eliminates the incentive to innovate at all, and indeed creates disincentives for innovation by imposing

an economic punishment on innovators. Under the BAT approach, polluting industries have no financial interest in the development of better pollution control technology that imposes higher production costs. Indeed, the opposite is true. The BAT approach encourages industry to seek any means to delay and deter new regulation. Industry will have the information as well as the incentive to persuade administrators, courts, and other authorities that a suggested technology is not "feasible" and should not be required.

If government requires whatever technology is available, then, industry has no economic incentive to develop new mechanisms for decreasing safety and health risks. Moreover, the BAT approach, applicable as it is only to new sources, raises the cost of retiring old facilities, which delays capital turnover and in that way aggravates environmental degradation. The paradox, in a nutshell, is this: Designed to promote good control technology, the BAT strategy actually discourages innovation. It is therefore self-defeating.

Paradox 4: Redistributive Regulation Harms Those at the Bottom of the Socioeconomic Ladder

A common justification for regulation is redistribution. The legislature imposes controls on the market to prevent what it sees as exploitation or unfair dealing by those with a competitive advantage. In principle, the claim for redistribution is often a powerful one. People often enter into harsh deals, simply because their options are so few. In fact, market wages and prices depend on a wide range of factors that are morally irrelevant: supply and demand curves at any particular point; variations in family structure and opportunities for education and employment; existing tastes; and perhaps even differences in initial endowments, including talents, intelligence, or physical strength (chapter 6). So long as the regulation can be made effective and does not produce high ancillary costs (an important qualification), government should not always take these factors as "natural," or let them be turned into social disadvantages.

The inspiration for minimum wage legislation, for example, is easy to identify. Such legislation prevents workers from having to settle for market wages that do not even approach the poverty level and thus offer minimal incentives to work. So, too, occupational safety statutes protect workers against extremely hazardous workplaces; rent control legislation prevents tenants from being subject to unanticipated price increases and perhaps thrown into significantly inferior housing; and implied warranties of habitability protect tenants from living in disgraceful and indeed dangerous apartments.

In all these cases, however, regulation is a poor mechanism for redistributing resources, precisely because it is often self-defeating. The problem is that if everything else is held constant, the market will frequently adjust to the imposition of regulation in a way that will harm the least well off. It is a mistake to assume that regulation will directly transfer resources or create only ex-post winners and losers—an idea exemplified by the assumption that the only effect of the minimum wage is to raise wages for those currently working. An important consequence of the minimum wage is to increase unemployment

by raising the price of marginal labor; and those at the bottom of the ladder—the most vulnerable members of society—are the victims.[17] In the same vein, rent control legislation and implied warranties of habitability create incentives for producers (landlords) to leave and disincentives to enter the housing market, with perverse redistributive consequences and especially harsh results for the poor, who may be left without housing at all.[18] It should not be surprising to find that Cambridge, Massachusetts has experienced a dramatic growth in new housing since rent control has been eliminated by state referendum.

Laws forbidding discrimination or requiring affirmative action will to some extent have the same perverse effect, since they will make it more expensive to hire blacks, women, and older people by increasing the likelihood that employers will be subject to a lawsuit in the event of a discharge.[19] Similarly, occupational safety and health regulation does not unambiguously promote the interests of workers. By raising costs, it may depress wages and increase unemployment, thus harming the least well off. In each of these cases, the group that is harmed is likely to be poorly organized and incapable of expressing itself through the political process.

In sum, redistributive regulation will have complex distributive consequences, and the group particularly disadvantaged by the regulation will typically consist of those who are already most disadvantaged. Efforts to redistribute resources through regulation can therefore have serious perverse results.

Three qualifications are necessary here. First, the redistributive regulation, though in some ways perverse, might be part of a system of redistribution that is effective overall. A minimum wage law might be justified as a means of protecting the working poor, if it is accompanied by a welfare system to take care of those who cannot work at all. For this reason, plausible arguments can be made for the minimum wage. It has been argued, for example, that an increase in the minimum wage is necessary to guarantee that work will be sufficiently remunerative to keep people out of poverty and to send a signal about the importance and value of work, thereby increasing the supply of and demand for labor. These effects might outweigh any unemployment effect.

A second qualification of the redistribution paradox relates to the fact that preferences are not static. Preferences are usually taken as exogenous to and independent of the legal rule, but sometimes this is a mistake (chapter 1). If the statute in question transforms preferences and beliefs, the self-defeating effect just described will not occur. For example, laws forbidding sexual harassment aim to alter the desires and beliefs of would-be harassers; and if the laws succeed in this goal, any perverse side effects may be minimal or nonexistent. The same argument may apply to antidiscrimination laws generally. If such laws change attitudes, they may not on balance harm the least well off. There is, however, little empirical evidence on the effects of law in changing preferences and beliefs, and, in any case, this is not likely to result from such redistributive regulation as minimum wage legislation.

The third qualification is that laws that seem to be justified on redistributive grounds are best understood as a response to inadequate information. Occupational safety and health legislation does not transfer resources from em-

ployers to employees; but it may be justified as a response to the fact that workers lack information about workplace risks. In the face of inadequate information, a regulatory response may well make sense. In fact, a well-tailored response will make workers better off. But this point leads directly to our fifth paradox.

Paradox 5: Disclosure Requirements May Make People Less Informed

Sometimes markets fail because people are deceived or lack information. Regulatory agencies commonly respond by requiring correction or full disclosure. Congress and agencies have imposed disclosure regulations in many areas, ranging from occupational and environmental risks to potentially deceptive advertising (chapter 13). Here the rationale is straightforward. Whether or not ignorance is bliss, it is an obstacle to informed choice. Surely, it might be asked, regulation cannot be condemned for increasing information?

Disclosure strategies are indeed valuable in many circumstances. But for two reasons, they can be self-defeating. The first is that people sometimes process information poorly. After being given certain data, they actually "know" less than they did beforehand. In particular, when people receive information about probabilities, especially low ones, they frequently rely on heuristics that lead to systematic errors (chapter 10). Thus, for example, people assess probabilities by asking if the event was a recent one, by seeing if an example comes readily to mind, and by misunderstanding the phenomenon of regression to the mean. In addition, disclosure or corrective language can help straighten out one form of false belief but at the same time increase the level of other kinds of false beliefs.[20] Finally, there is a risk of information overload, causing consumers to treat a large amount of information as equivalent to no information at all.[21] All this suggests that with respect to information, less may be more. Additional information can breed confusion and a weaker understanding of the situation at hand.

The second problem is that a requirement of perfect accuracy will sometimes lead producers or other regulated entities to furnish no information whatsoever. For example, if producers are prohibited from advertising, unless they eliminate all potential deception or offer strong substantiation for their claims, they might not advertise at all. The result will be the removal from the market of information that is useful overall.[22] If advertisers must conduct extensive tests before they are permitted to make claims, they will be given a strong incentive to avoid making claims at all. More generally, almost all substantive advertisements will deceive at least some people in light of the exceptional heterogeneity of listeners and viewers. If this is so, efforts to eliminate deception will significantly reduce advertising with substantive content.

These various difficulties suggest that the recent enthusiasm for disclosure requirements is at least in some settings a mistake, for the simple reason that such requirements may defeat their own purpose. Disclosure requirements sometimes ensure that people are less informed; a good deal of thought needs to be given to their content (chapter 13).

Paradox 6: Independent Agencies Are Not Independent

The distinctive institutional legacy of the New Deal period is the "independent" regulatory agency. An agency is independent if Congress has provided that its members can be discharged by the president only for specified causes. If Congress has so provided, it is ordinarily understood that the president cannot discharge independent commissioners simply because he disagrees with their views, and that his supervisory authority is sharply limited. Independent agencies, some of them antedating the New Deal, include the Federal Trade Commission, the Federal Communications Commission, the Interstate Commerce Commission, and the National Labor Relations Board. The paradox at issue here is one of institutional design rather than substantive regulatory policy.

The argument for the independent agency stems largely from a belief in the need for expert, apolitical, and technically sophisticated administration of the laws. Even if independent agencies achieved this end, we should question the goal itself. Independent agencies often must make important judgments of value and principle, and on those judgments expertise is never decisive. Consider, for example, the decisions of the National Labor Relations Board defining what constitutes an unfair labor practice; the judgment of the FCC about whether licensees are obliged to present programming on public issues, or whether diversity on the basis of race or sex counts in favor of an applicant for a license; and the safety requirements of the Nuclear Regulatory Commission. None of these policies is based solely on technocratic judgments, and so may properly belong—at least in part—in the political rather than the regulatory sphere.

But even if we accept the premise that political independence is necessary, the fact is that independent agencies are not independent at all. Indeed, such agencies are responsive to shifts in political opinion and even to the views of the president.[23] But the problem is even worse than that. The independent agencies have generally been highly susceptible to the political pressure of well-organized private groups—perhaps even more susceptible, on balance, than executive agencies.[24]

Many of the most egregious illustrations of agency vulnerability to pressure groups can be found in precisely this area. Thus, the Interstate Commerce Commission created and enforced cartels in the transportation industry; the Federal Trade Commission has sometimes behaved in an anticompetitive manner, capitulating to losers in the marketplace; and the FCC has been dominated by the communications industry.[25] Far from acting as disinterested experts, independent administrators often are, in practice, subject to parochial interests.

Why would agencies independent of the president be so susceptible to factional power? The phenomenon might be explained at least in part by the fact that executive agencies, precisely because they are subject to presidential control, are able to withstand the parochial pressures imposed on "independent" agencies that lack the buffer of presidential oversight. The absence of this presidential buffer leaves agencies vulnerable both to individual members and committees of Congress, which sometimes represent narrow factions and well-organized private groups with significant stakes in the outcome of regulatory decisions. Executive agencies are at least sometimes immunized from those pressures precisely because of the protective, insulating wing of the president. Ironically, independence from the president often appears to be a mechanism for increasing susceptibility to factionalism.

The susceptibility of the independent agencies to factionalism does not of course imply that executive agencies are invulnerable to similar forces. The notion that independent agencies are systemically more susceptible to factions than their counterparts within the executive branch seems much too broad. But if Congress wants to ensure independence in the execution of its laws, the independent agency device is a most unlikely way to achieve that goal.

Other Paradoxes, in Brief

I have described some prominent regulatory paradoxes, but there are others as well. For example, it has been argued that the pursuit of the "best interests of the child"[26] in custody determinations in fact disserves the best interests of children, because of the enormous time spent in resolving the complicated factual question.[26] Protectionist legislation is sometimes justified on the theory that it will help domestic industries develop into potent competitive forces, but in fact, protectionism may induce flabbiness and in the end defeat the goal of promoting international competitiveness. As we have seen, extensive procedural protections, designed to improve agency rule making, may lead to little or no agency rule making. And restrictions on the availability of abortion, defended as a means of protecting human life, appear to have resulted in the death of many women per year and at the same time not to have protected a large percentage of fetuses from the practice of abortion.[27]

Many more paradoxes can be found. There is evidence that mandatory prescriptions for drugs have increased health risks by limiting the availability and raising the cost of prescription drugs; this in turn has decreased self-treatment and encouraged people to use possibly less effective over-the-counter drugs.[28] Product safety regulation may have a "lulling effect" on consumers, leading them to take fewer precautions and to miscalculate risks.[29] Health regulation can increase health risks, if agencies focus on one kind of risk and ignore others that can be aggravated by regulation itself (chapter 12). The government prohibition on cigarette advertising on television, designed to decrease smoking, may have increased smoking because it (1) reduced competition among firms, thus cartelizing the industry over the advertising issue; (2) eliminated the application of the fairness doctrine to cigarettes, which would have ensured a vigorous anticigarette campaign; and (3) saved the industry substantial sums of money.[30]

A final paradox can be found in the law of sex discrimination, where principles of "formal equality" have been invoked to forbid consideration of sex in custody, alimony, and divorce disputes. It is quite possible that equality principles, understood as prohibitions on any form of sex differentiation in law, have in some contexts produced less rather than more in the way of real equality between men and women.[31] When two groups are differently situated, a legal requirement that they be treated the same seems a perverse method of promoting equality between them. Here too, then, legal controls have been self-defeating.

III. Two Questions: What We Don't Know

Causation

It is tempting to react to the regulatory paradoxes by suggesting that the relevant strategies are not self-defeating at all. On the contrary, they might represent a conscious governmental choice and even, on one view, regulatory success. Public choice theory suggests that legislative outcomes are frequently a product of pressure applied by well-organized private groups. It is not difficult to find "cartels in the closet" to account for many or all of the paradoxes and to make them seem far less mysterious.

For example, the apparently perverse effects of redistributive regulation may be actively sought by the benefited groups. On this account, the purpose of minimum wage legislation might not be to help the poor, but rather to immunize union members from competition by people who are willing to work for low wages by limiting entry into the labor market. Far from being unintended consequences, the harmful effects on those at the bottom of the economic ladder may be actively sought. Looked at from this perspective, minimum wage legislation creates a cartel among those not threatened by unemployment, benefiting them at the expense of new entrants into the labor market.

So, too, independent agencies might be created at the behest of groups that know they will have particularly strong influence over public officials not subject to presidential oversight; or Congress might create an independent agency not to ensure technocracy or neutrality, but to increase the power of its members and committees over agency decisions. Similarly, existing industry, in a bid to reduce competition, might acquiesce in or actively seek regulations distinguishing between old and new risks. It is hardly unusual for companies to enlist regulatory law in the service of cartelization.

The overregulation–underregulation phenomenon has a similar explanation. By adopting a draconian standard, legislators can claim to support the total elimination of workplace hazards or dirty air; but legislators and regulated industries know that administrators will shrink from enforcing the law. A "deal" in the form of a stringent, unenforceable standard benefits the politi-

cally powerful actors. Hence, the political economy of overregulation is similar to that of open-ended delegations of administrative authority: In both cases, legislative incentives incline Congress toward broad and appealing statutes that will not in practice harm politically powerful groups. The public is the only real loser.

Explained in this manner, the paradoxes of the regulatory state are not mysterious at all. On the contrary, they are perfectly predictable responses to electoral self-interest and to disparities in political influence.

While explanations of this sort have power in some settings, clear evidence on their behalf is often unavailable. It is of course possible that the seemingly paradoxical effects of regulatory programs actually account for their enactment. But this is only a possibility. To explain a phenomenon by reference to its consequences is bad social science, even though it is pervasive in such widely diverse disciplines as neoclassical economics, Marxism, and sociobiology.[32] In the context of the regulatory state, whether public choice explanations are good ones, rather than merely plausible stories, depends not just on the consequences of regulation, but also on a careful investigation into the actual forces that lead to regulation. In the regulatory sphere, such investigations are infrequent.

The most we can say is that the regulatory paradoxes might reflect the influence of well-organized private groups, and that in some settings there is direct or indirect evidence to support that conclusion. At least thus far, any more global conclusion is not supported by what we know about the facts.

Magnitude

Whether the regulatory paradoxes should cause major concern depends on their magnitude. Here too, much of the relevant information remains to be developed. For example, a decision to focus on new sources of pollution would be understandable if that decision would have only a minor effect in perpetuating old sources. But if the effect is substantial, the regulatory policy would almost certainly be ill-considered. Everything depends on the question: How large is the relevant effect? Similarly, the minimum wage might well be justified if its effect is the unemployment of only a few additional people. The relevant question is the elasticity of the demand for labor. Even if some people are misled by compulsory disclosure of risks, perhaps there will be sufficient gains through reducing others' ignorance to justify the regulation. And even if some producers refuse to advertise at all in the face of a substantiation requirement, perhaps the overall level of information will increase.

Critics of regulation sometimes treat the existence of unintended side effects or partly self-defeating strategies as a reason to abandon regulatory controls altogether. But in order to justify that conclusion, it is necessary to gather detailed evidence on the magnitude of the relevant effects in particular regulated markets and overall. In some contexts, regulation having some self-defeating results will on the whole make things better rather than worse. The

American regulatory state, though pervaded by paradoxes, has had a number of substantial successes.

From both theory and experience, it is possible to conclude that the regulatory paradoxes will arise frequently and thus to prescribe efforts to avoid them. Certainly we have far too little information to say, as a general matter, that regulatory programs embodying the paradoxes are by virtue of that fact a bad idea on balance, at least when compared with the preregulatory status quo. Total elimination of such regulatory programs is hardly warranted. Nevertheless, a system that avoided the paradoxes would bring about major improvements.

IV. Lessons

Congress

The paradoxes of regulation provide a number of concrete lessons for Congress. At the most general level, they suggest that legislators should be attentive to the incentive effects of regulatory statutes and the possibility of strategic or self-interested adaptation by administrative agencies and members of regulated classes. Statutes embodying an assumption that the preregulatory world can be held constant—that existing prices, wages, choices, and so forth will endure—are particularly likely to be confounded when implemented.

More specifically, the paradoxes suggest that the legislature should generally avoid best available technology strategies; be concerned with old risks as well as new ones; decline to attempt to redistribute resources through regulation; be attentive to the possibility that disclosure requirements will simply confuse people or chill information in the first instance; create incentives for regulation when regulation is desired; as a rule, place agencies under the control of the president; and call for some form of balancing between the costs and benefits of regulation. Ideas of this sort have direct implications for modern regulatory reform (chapters 12 and 13).

Judges and Administrators

The regulatory paradoxes provide important lessons for judges and administrators as well as legislators. These officials are of course bound by legislative enactments, and to the extent that regulatory statutes unambiguously call for self-defeating strategies, officials have no choice but to honor them. But frequently the interpretation of a statute, or the filling of statutory gaps, is based on an understanding of the real-world consequences of the alternative possibilities. Administrators exercise considerable discretion in giving content to ambiguous laws, and the legal judgment about whether an agency's decision is "arbitrary" within the meaning of the Administrative Procedure Act should be informed by an accurate understanding of the paradoxes of the regulatory state. Attention to the often unanticipated systemic effects of regulatory con-

trols is an imperative for administrators and judges as well as for legislators. I offer three examples here of how these officials can use the knowledge of regulatory paradoxes to inform their actions.

The Overregulation–Underregulation Paradox

In two important cases, the Supreme Court was asked to interpret the provisions of the Occupational Safety and Health Act that regulate exposure to toxic substances. The pertinent language directs the secretary of labor to promulgate the standard that "most adequately assures, to the extent feasible . . . that no employee will suffer material impairment of health or functional capacity even if such employee has regular exposure to the hazard . . . for the period of his working life." The statute also defines "occupational safety and health standard[s]" as measures that require "conditions . . . reasonably necessary or appropriate to provide safe or healthful employment and places of employment."

In *Industrial Union Department, AFL-CIO v. American Petroleum Institute*,[33] the Supreme Court was confronted with an OSHA regulation of benzene. Though the consequences of the regulation were sharply contested, there was reason to believe that the regulation would impose enormous costs for small or speculative gains. A plurality of the Court concluded that the secretary of labor must establish that a toxic substance posed a "significant risk" to health before she could regulate it. There was little direct support for the plurality's conclusion in the language or history of the Act. Unable to point to a solid textual basis for its "significant risk" requirement, the plurality invoked a clear statement principle:

> In the absence of a clear mandate in the Act, it is unreasonable to assume that Congress intended to give the Secretary the unprecedented power over American industry that would result from the Government's view. . . . Expert testimony that a substance is probably a human carcinogen . . . would justify the conclusion that the substance poses some risk of serious harm no matter how minute the exposure and no matter how many experts testified that they regarded the risk as insignificant. That conclusion would in turn justify pervasive regulation limited only by the constraint of feasibility. . . . [T]he Government's theory would give OSHA power to impose enormous costs that might produce little, if any, discernible benefit.[34]

The plurality went on to suggest that the government's interpretation would give the secretary of labor "open-ended" policy-making authority that might amount to an unconstitutional delegation of legislative power. In a concurring opinion advocating an interpretation of the Act that would permit cost-benefit balancing, Justice Powell suggested that "a standard-setting process that ignored economic considerations would result in a serious misallocation of resources and a lower effective level of safety than could be achieved under standards set with reference to the comparative benefits available at a lower cost."

The "significant risk" requirement cannot be found explicitly in the statute; indeed, the text of the relevant provisions suggests that no such requirement was imposed on the secretary of labor. But the plurality's conclusion was nonetheless sound. Realistically speaking, the language of the statute need not be considered dispositive. It is simply a myth to suggest that the Congress that enacted OSHA even considered the question whether regulation could require enormous expenditures to redress minimal risks. Despite the broad language of the toxic substances provision, Congress never focused on that problem. It is a legitimate interpretive strategy to understand broad words not to require an outcome that Congress could not plausibly have contemplated.

In the context of *American Petroleum*, the plurality was therefore correct in considering itself free to read an implicit "significant risk" requirement into the statute. In light of the overregulation–underregulation paradox, it would make little sense to interpret the statute so as to allow—indeed, require—the secretary to regulate to the point of "feasibility" merely because one or a few employees might suffer "material health impairment" as a result of a lifetime of exposure. Such an interpretation would make the Department of Labor reluctant to embark on a course of regulation at all, and as we have seen, would result in less, not more, protection of workers. It would ensure that there would be less regulation of carcinogens or less enforcement of those regulations that were promulgated—or, most likely, both.

In *American Textile Manufacturers Institute v. Donovan*,[35] the Supreme Court decided a question left open in *American Petroleum:* Whether the Occupational Safety and Health Act required cost-benefit analysis. In arguing that it did, the industry contended that the word "feasible" meant that the secretary must show not only a significant risk, but also that the benefits of regulation justified the costs. "Feasibility," in the industry's view, contemplated a balancing of costs and benefits. The government contended that once OSHA had shown a significant risk, it could regulate to the point where the survival of the regulated industry would be endangered by additional controls. For the government, the term "feasibility" connoted not cost-benefit balancing, but instead regulation to the maximum extent "possible."

In accepting the government's argument, the Court relied on the dictionary definition of "feasible," concluding that the term meant "capable of being done, executed, or effected," rather than justified after balancing costs and benefits. This approach to statutory interpretation was not entirely unreasonable. But the same principles that support the plurality view in *American Petroleum* cast doubt on *American Textile Manufacturers*.

First, notwithstanding the statute's language, it is probably unrealistic to believe that Congress actually focused on, and resolved, the question whether the government's approach was to be favored over some kind of balancing of costs and benefits. That question never arose during the debates. Second, a system requiring the secretary to identify a significant risk, but prohibiting her from undertaking cost-benefit analysis, seems utterly irrational. Whether a risk is "significant" depends not only on its magnitude, but also on the costs of eliminating it. A risk that is relatively small might call for regulation if the costs are also small, whereas a large risk might well be best left unregulated if the costs of regulation are enormous. A rational system of regulation looks not

at the magnitude of the risk alone, but assesses the risk in comparison to the costs. Finally, a law requiring the secretary to regulate all significant risks to the point of endangering the industry would be a recipe for both overregulation and underregulation.

These considerations could not have controlled the Court's decision if the statute dictated a contrary result, but the word "feasible" was probably capacious enough to accommodate a kind of proportionality requirement. To be sure, the case was a difficult one. But by its reading of the statute, the Supreme Court contributed to the irrationality of the Occupational Safety and Health Act—irrationality that has harmed workers, employers, consumers, and the public at large. An understanding of the overregulation–underregulation paradox might have prevented this result.

The Old Risk–New Risk Paradox

A number of judicial decisions might have been different if courts had been attuned to the old risk–new risk paradox. Consider, for example, one district court's creation, in the face of an ambiguous text, of the "prevention of significant deterioration" (PSD) program in *Sierra Club v. Ruckelshaus.*[36] In that case, the court ruled that state implementation plans under the Clean Air Act must include provisions not merely complying with national air quality standards, but also designed to prevent the degradation of air currently cleaner than those standards require. The consequence of the PSD program is to ensure that especially clean areas remain especially clean. They are not permitted to become dirtier, even if they would continue to provide a safe and healthful environment.

One of the court's goals was to ensure that federal environmental policy protected beauty and visibility in currently pristine areas. While the PSD program has to some degree promoted that goal, it has also had perverse side-effects. For example, it has delayed the salutary substitution of clean, low-sulfur western coal for dirty, high-sulfur eastern coal; at the same time, it has protected dirty existing plants in the East against replacement with cleaner new ones in the West.[37] To protect the atmosphere in Aspen from degradation is, almost inevitably, to perpetuate the existence of old, particularly dirty producers in New York. The foreclosure of new risks has thus increased the magnitude of old ones. It is far from clear that the environment is better off as a whole.

Indeed, it should come as no surprise that the PSD program has become a primary means of protecting eastern industry and eastern states against western interests. States in the West seeking to attract industry have found, perversely, that an environmental program can be used to create a cartel against new entry. A PSD program based on an understanding of the adverse effects of that cartel for the prevention of environmental degradation would take a quite different form.

The court that decided the *Sierra Club* case was unaware of these effects. Because the statutory basis for the decision was quite thin, an understanding of the environmental and nonenvironmental costs associated with the PSD program might well have led to a contrary result.

The Independent Agency Paradox

The precise constitutional status of the independent agency remains an uncertain question. In *Humphrey's Executor v. United States,*[38] the Supreme Court, affirming the constitutional validity of the independent agency, held that Congress could constitutionally prevent the president from removing a member of the Federal Trade Commission simply because it pleased him to do so. Recent decisions have reaffirmed the authority of *Humphrey's Executor* insofar as it recognizes that some degree of independence from the President is permissible.[39] But suppose that members of the Nuclear Regulatory Commission or the Federal Trade Commission act in ways that consistently reject the president's views about public policy. May the president discharge the relevant commissioners? It is frequently assumed that he may not. But neither *Humphrey's Executor* nor any other case explains what "independence" precisely means, or whether it extends to such situations.

The problem might be solved through statutory interpretation that takes account of the independent agency paradox. The relevant provisions allow the president to discharge a commissioner "for cause," defined as "inefficiency, neglect of duty, or malfeasance in office." Although ambiguous, these words do not entirely immunize commissioners from the control of the president; instead, they allow him to remove officials under certain circumstances. For those attuned to the independent agency paradox, it might seem that the words are best read to grant the president something in the way of supervisory and removal power—allowing him, for example, to discharge as inefficient or neglectful of duty those commissioners who have frequently or on important occasions acted in ways inconsistent with his policy wishes.

This result might seem counterintuitive in light of the frequent understanding that independent agencies are to be immunized from presidential policy making. But there is a plausible precedent for precisely this conclusion in a recent Supreme Court decision, *Bowsher v. Synar.*[40] In that case, the Court held that Congress could not delegate power to administer the Gramm–Rudman statute to the comptroller general, because the comptroller was subject to congressional will. In the Court's view, those who execute the law must not be subject to the policy-making authority of Congress, except insofar as legislative instructions are embodied in substantive law. The relevant statute allowed Congress to discharge the comptroller for "inefficiency, . . . neglect of duty, . . . [or] malfeasance." The Court said that these words conferred on Congress "very broad" removal power and would authorize Congress to remove the comptroller for "any number of actual or perceived transgressions of the legislative will."

Thus, the words governing congressional power over the comptroller general and presidential power over independent agencies are essentially identical. If those words have the same meaning in these admittedly different contexts,

the president has "very broad" removal power over the commissioners of the independent agencies, with correlative powers of supervision and guidance. It would follow, then, that the independent agencies are in fact subject to a considerable degree of presidential control. They are not, as a matter of statutory law, "independent" of him at all. In view of the independent agency paradox, courts would do well to invoke a clear statement principle that grants the president broad supervisory power over independent agencies unless Congress has expressly stated its will to the contrary. Such an approach would minimize the risks inherent in the independent agency form and promote coordination and accountability in government. It would require Congress to speak unambiguously if it wants to compromise those goals.

V. Conclusion

There are multiple breakdowns in private markets, and government regulation is sometimes an effective response. The administrative state has not been a universal failure; we have seen a wide range of successes. But regulatory programs have not always done what they should, and the paradoxes of the regulatory state have been a pervasive source of its problems. Self-defeating regulatory strategies take many forms. I have discussed six such paradoxes and referred to several others; still others undoubtedly exist.

In proposing reforms for the regulatory state, little can be gained from generalities that point to the frequent problems created by either government regulation or private markets (which are themselves of course a form of regulation). These problems are too particular and too dependent on the context to allow for global prescriptions. It is far more helpful to rely on particularized understandings of how both markets and regulation tend to break down—in short, to learn from the past. The experience of the regulatory state includes many self-defeating regulatory strategies. But enough information is in place to help legislators, administrators, and judges to minimize their adverse effects, and even to prevent their occurrence. The result would be a small but firm step in the direction of an American-style perestroika—a system that is entirely unembarrassed by the use of government to reflect democratic aspirations, to promote economic welfare, and to foster distributional equity, while at the same time insisting on strategies that embody the flexibility, adaptability, productive potential, and decentralization characteristic of private markets.

Notes

1. See John Graham and Jonathan Wiener, eds., Risk *versus* Risk, pp. 8–11 (Cambridge: Harvard University Press, 1995); Gregg Easterbrook, *A Moment on the Earth* (New York: Viking, 1995). Air and water pollution data can be found in *id.*; CEQ, 24th Annual Report, Environmental Quality (1993).

2. See Robert W. Crandall, et al., Regulating the Automobile (Washington, D.C.: Brookings Institution, 1986), pp. 44–74.

Afterword

What, then, is the relationship between free markets and social justice? In answering that question, we should recognize that markets, free or otherwise, are not a product of nature. On the contrary, markets are legally constructed instruments, created by human beings hoping to produce a successful system of social ordering. As I have emphasized throughout, there is no opposition between "markets" and "government intervention." Markets are (a particular form of) government intervention. Hence the interactions promoted by markets include coercion as well as voluntary choice. Markets should hardly be identified with freedom. The law of property, for example, coerces people who want access to things that they do not own. And like all instruments, markets should be evaluated by asking whether they promote our social and economic goals.

Often markets do promote the basic goals of a well-functioning social order, and a social order that aspires to be well functioning will not dispense with markets. In the area of free speech, markets are very important, since they facilitate the exchange and the production of information—scientific, political, medical, and much more (chapter 7). In constitutional law generally, property rights and market ordering can diminish unwarranted political interference with the production of social wealth (chapter 8). Environmental protection in the United States has abandoned markets too readily; it should take far more advantage of market thinking than it now does. In the area of risk regulation, economic incentives can promote environmental and safety goals in a cost-effective way. They can do this by channeling private behavior in the right directions (chapters 13 and 14).

For all these reasons, American and indeed Western governments should enlist markets more regularly than they now do. In the aftermath of the New Deal, American government has been much too willing to use rigid, bureaucratized solutions to economic and social problems. It should turn instead to flexible incentives, allowing private adaptation for the sake of public goals. Indeed, this step ranks among the most promising routes for reforms in the twenty-first century. Many creative possibilities can be imagined.

These are important points; but they are not really points about justice. We have seen that markets typically reward people on the basis of factors that

are irrelevant from the moral point of view. These morally irrelevant factors include not simply race and gender, which can play a large role in markets, but many other factors that account for market success. Achievements within markets come from the innumerable accidents that allow people to develop the characteristics that markets reward—or that prevent people from developing those characteristics.

These accidents are pervasive. If, for example, you are born to an average family on 57th Street and Dorchester Avenue in Chicago, your life prospects will be very good, and altogether different from what they will be if you are born to most families ten blocks south. If you are born into some families, you will be unlikely to be healthy, strong, well mannered, hard working, or well educated, and these are the characteristics that you may need in order to do well in markets. People from a diverse range of theoretical positions ought to agree that markets will not promote justice unless they are made part of a system that offers minimally decent opportunities to all. In existing societies that use markets, the ideal of equal or even decent opportunities is violated on a daily basis.

Markets are also accompanied by a large network of social norms, and existing norms—involving self-destructive or uncooperative behavior, discrimination, pollution—may produce inefficiency or injustice (chapter 2). Much of our conduct is a product of norms, for which we are not responsible and which we may wish, on reflection, to change. Collective and even governmental action may be necessary to improve norms or generate new ones. Far too little attention has been given to the harmful effects of social norms on individual freedom and well-being. This point bears very much on the role of government. Democratic efforts to promote well-being by improving norms—and changing choices—are fully legitimate. Of course, rights should operate as constraints on this process, and institutions must be developed to reduce the risk of abuse of government power.

As I have emphasized throughout this book, a just society should be closely attentive to the background conditions against which markets proceed. Existing distributions and preferences should not be taken as inevitable or as given. Extreme deprivation is unjust in large part because it denies people the opportunity to form preferences and beliefs that lead to good lives. It is thus important to attend to existing distributions of entitlements—distributions that are a function not of nature but of law—and to the effects of those distributions on the development of people's life prospects and even their desires and beliefs. Many of the chapters in this book challenge market thinking on the ground that it is insufficiently attuned to the harmful effects of unjust background conditions. This is true not only in the context of extreme poverty and deprivation, but also in the area of discrimination on the basis of race and sex (chapter 6), where people's preferences can be formed by background injustice, and even in the area of environmental protection (chapter 10).

To say this is not to say that societies that reject or try to reject markets are just. Usually they are especially unjust, because they use especially unfair mechanisms—various forms of political favoritism—for producing social re-

wards. But we need not enter into deep philosophical territory in order to recognize that much of the time, markets help or hurt people for reasons that are unfair, in the sense that they are ill connected with any plausible conception of justice.

Identification of injustice does not, of course, lead to any particular set of proposals for change. Disruption of markets may be futile or counterproductive. Markets, rearrangements of markets, and alternatives to markets should be assessed pragmatically and in terms of their actual consequences for those who live with them. We have seen many areas in which the effects of markets are good. We have seen other areas in which government interferences with market ordering, especially through command-and-control regulation, make things worse rather than better. Here there is enormous room for more empirical work and for substantive reforms. It is possible to imagine a wide range of programs that would improve human lives—in part simply by lengthening them—and do so largely through attending to public uses of market incentives. Thus a government attentive to the existence of background injustice should refuse to take existing practices, norms, and distributions as given; but it might well invoke market incentives in the support of social aspirations.

Through this possibility we may begin to see the place of market ordering in a system committed to social justice. Such a system is likely to favor a particular conception of democratic deliberation, one that embodies a belief in political equality and in reason-giving in the public domain. If that system is also committed to social justice, it will see free markets as instrumental goods to be evaluated by their effects. This is hardly a reason to abandon markets. But it is a reason to insist on the priority of democratic goals, including social justice, to market ordering—while enlisting, much of the time, the latter in the service of the former.