GENERAL CHARACTERISTICS OF RULES

Louis Kaplow

Discussion Paper No. 191
8/96

Harvard Law School
Cambridge, MA 02138

The Center for Law, Economics, and Business is supported by a grant from the John M. Olin Foundation.
General Characteristics of Rules
Louis Kaplow*

Abstract

This entry addresses two fundamental characteristics of rules. The first concerns the degree of precision, detail, or complexity they embody: how finely are different sorts of behavior to be distinguished? A second aspect of legal commands concerns when a given level of detail is provided -- at the time of promulgation ("rules") or subsequent to individuals' actions, in the context of an adjudication ("standards"). These aspects of rules are considered from a perspective that focuses upon information costs and dissemination: different sorts of legal commands involve differing costs of formulation and application by private parties (deciding upon their own conduct) and adjudicators, and the character of laws also influences how well parties actually will understand the law and conform their conduct accordingly. The discussion encompasses related questions involving the role of precedent, the evolution of the law over time, legal uncertainty, and accuracy in adjudication. This entry also addresses the separate problem of how changes in legal rules should apply to prior behavior or pre-existing investments -- issues of retroactivity and transition.

(forthcoming in the Encyclopedia of Law & Economics)

JEL No. K40

---

*Harvard Law School and National Bureau of Economic Research. I am grateful for comments from Steven Shavell, research assistance from Enrique Arana, Judson Berkey, and Jerry Fang, and support from the John M. Olin Center for Law, Economics, and Business at Harvard Law School.
1. Introduction

Legal rules serve many functions: channeling behavior (e.g., tort law, environmental law, criminal law), providing background rules (much of contract, commercial, and corporate law), and defining obligations and eligibility for benefits (tax law, social welfare provisions). This entry, following existing research, focuses on the first type of law. The relevance (or irrelevance) of the analysis to other types of law will usually be apparent, and some comments about the differences in contexts will be noted.

The most commonly noted characteristic of rules concerns the degree of precision, detail, or complexity they embody: how finely are different sorts of behavior to be distinguished? Such matters may arise in defining the scope of a legal command, providing for exceptions, adjusting sanctions based upon aggravating and mitigating circumstances, and so on. A related aspect of legal commands concerns when a given level of detail is provided -- at the time of promulgation ("rules") or subsequent to individuals' actions, in the context of an adjudication ("standards"). These aspects of rules are considered from a perspective that focuses upon information costs and dissemination: different sorts of legal commands involve differing costs of formulation and application by private parties (deciding upon their own conduct) and adjudicators, and the character of laws also influences how well parties actually will understand them and conform their conduct accordingly. This approach is used to illuminate choices about the degree of precision (or complexity) with which legal commands are formulated and between rules and standards. These subjects, in turn, encompass related questions involving the role of precedent, the evolution of the law over time, legal uncertainty, and accuracy in adjudication.
After outlining the main ideas concerning these topics, this entry addresses the separate problem of how changes in the law should apply to prior behavior or pre-existing investments -- issues of retroactivity and transition -- and considers some other aspects of rules that have been addressed in the economic analysis of law.

2. Precision of the Law

The problem. The precision of a legal command is taken here to refer to the degree of detail or differentiation involved. For example, an environmental regulation is more precise if more types of pollutants or sources of pollution are distinguished. A negligence regime is more precise if the standard of care is more finely tailored to different types of actors in different contexts. A damages scheme is more precise the greater the extent to which damages are designed to reflect the particular harm done to a particular victim.

The greater the degree of precision, the greater will be the costs of formulating legal commands and applying them in adjudication and of parties interpreting them for purposes of deciding how to conform behavior to the law. But it does not follow that the costs of precision -- any more than costs of litigation or of pollution control -- should be minimized: the simplest rules might permit all acts, require equal reduction of pollutants regardless of their toxicity, or require the same speed limits on all roads. Rather, the degree of precision should be optimized, which makes an economic analysis particularly useful.

Before proceeding, it is worth noting that greater precision in legal commands -- higher levels of detail, more distinctions recognized in rules and exceptions -- is a primary source of legal complexity, which is broadly condemned. The subject of the next section, whether a legal command is formulated as a rule or a standard (the latter leaving more open to subsequent
adjudication and thus being more difficult to interpret), also may bear on what is often described as legal complexity. A number of other sources of complexity could be considered, but they have received little attention in the literature. Most obviously, poorly drafted laws and legal opinions will be more confusing; it will be more costly to extract less information. Although this problem is often serious, analysis of it is straightforward. More subtle are problems like those confronted under tax laws wherein parties may devise complicated schemes to circumvent simple legal commands. In this setting, complexity may be unavoidable. Moreover, it is possible that greater complexity will reduce costs: if a more complex regime successfully discourages transactional manipulations, parties might in the end spend less in learning the intricacies of legal rules and in making convoluted arrangements.


First, we can consider the effects of precision on legal costs. It is straightforward that greater precision tends to increase them: more detailed laws tend to be more costly for the government to promulgate, for parties to interpret, and for enforcers to apply. The second factor -- private interpretation costs -- involves some subtlety, because parties will incur such costs only if they find it in their interest to do so. For example, extremely complex details may be ignored by almost everyone.

Second, more precise commands -- those that differentiate behavior more precisely -- will generally result in better behavior. If pollutants are differentiated and subject to particularized sanctions, rather than grouped together, subject to a single sanction (reflecting, say, the average harm they cause), there will be two beneficial results. Activities that are more harmful will be
subject to higher sanctions than otherwise, which will efficiently deter additional undesirable activity. Similarly, less harmful activities will face lower sanctions, avoiding inefficient overdeterrence.

These behavioral benefits are related to private legal costs, because the benefits can arise only if private actors expend the resources necessary to learn whether their particular activity is of a more or less harmful type, and thus subject to a higher or lower sanction. If the cost of determining the relevant sanction -- which may involve learning about the content of legal commands or about the nature of one's own activity (such as by testing the chemical content of discharges) -- is too high, individuals will not learn about the legal rules. Individuals with rather low benefits may simply forgo activities, even though there is some chance that they may be subject to low sanctions or no sanction, in which case their activities would be worth undertaking. Individuals with high benefits may simply act, even though there is some chance that they may face a very high sanction, in which case they will regret their decision. Individuals with intermediate benefits, however, will tend to acquire information and then to act accordingly -- that is, to proceed if the sanction is low and abstain if the sanction is high.

Given the nature of this problem, it is natural to ask whether private parties' incentives to acquire information will be socially appropriate. It turns out that in the simplest case, in which liability is strict and there are no predictable errors in adjudication, private and social incentives are aligned (for the familiar reason that private actors bear all the social costs of their activities, here including the cost of acquiring information). More generally -- if there is legal error, if individuals may misestimate the value of information, if information facilitates circumvention of legal sanctions, or if a negligence regime governs -- private incentives may be excessive or
inadequate. See Kaplow (1990), Kaplow and Shavell (1992), Shavell (1988a).

Having stated how parties will behave and what costs will be incurred when laws are more detailed, it is possible to characterize when greater precision is desirable. At the broadest level, the lower are the information costs and the greater is the variation in harm caused by different individuals' actions, the larger will be the net benefit of precision in legal commands. To offer more insight, it is helpful to consider some special cases.

First, suppose that formulation costs are low (perhaps because a single rule will apply to a large number of actions) and that adjudication costs are low (perhaps because harm occurs only with a low probability or because violations are enforced with a low probability/high sanction strategy, as suggested by Becker (1968)). In this case, more precise laws tend to be desirable as long as information is not prohibitively costly for actors to obtain. The reason is that the actors have the option of not obtaining information. Those who do make expenditures to become informed will be only those for whom the expected behavioral benefit is large -- individuals whose acts will be significantly desirable if they are not in fact very harmful but significantly undesirable if they are of the more harmful type. For example, consider a complex rule scheme that regulates routes over which hazardous substances may be transported. Most drivers -- those who never or rarely carry significant amounts of such substances -- will ignore the rules. Those who regularly transport hazardous materials, however, will undertake the necessary effort to learn the rules and conform their behavior. Their information costs may be high, but the benefits from better controlling their activity will be even larger.

One implication of this analysis is that the existence of high total compliance expenditures are not an unambiguous sign of excessively detailed or complex rules. Low total expenditures
may indicate that few individuals bother to learn the details of the law, in which case any additional formulation or adjudication costs would largely be wasted. High total expenditures, on the other hand, reflect that many individuals contemplate actions the desirability of which is contingent upon the type of their act; in this case, saving compliance costs by eliminating the law’s precision may be a mistake because significant behavioral benefits will be sacrificed.

Another implication of the analysis is that widespread noncompliance need not indicate that rules are unduly complex; as long as the few for whom the rules are most important do learn the details, the benefits of precision will be achieved at little cost with regard to others.

As a second case, suppose that private information costs are low (perhaps most individuals already know the details of their actions and how they might be sanctioned) but that the costs of differentiating acts in adjudication would be high (because it is hard for the state to learn the details of individuals' behavior). In this case, more refined legal commands may or may not be desirable. For example, additional costs may be repeatedly incurred in adjudication even though few individuals' behavior would be affected as a result of greater detail in the law. The distinction between this case and the foregoing one is that the state must make additional expenditures on adjudication even when the ex ante benefit of improved behavior would be small, because the state will not know when this is true; by contrast, when only ex ante information costs are high, private parties will incur them selectively.

Relationship of the precision of the law to the accuracy of adjudication. The precision of the law has a direct connection to the accuracy of adjudication, a topic explored in Kaplow (1994a) and Kaplow and Shavell (1994, 1996a). The former subject can be taken to refer to the precision of the law on the books and the latter question to the precision with which the law
actually is applied. At the time they act, individuals will be induced to become more informed about the details of their situation only if, in a subsequent adjudication, their behavior will in fact be differentiated from that of others. If, instead, individuals' acts will be assigned to different legal categories largely at random or if no effort will be made to refine the classification of acts, the situation is much as if the law itself did not make the refined distinctions in the first place. Many topics of procedure -- discovery, the use of experts, evidentiary rules, summary judgment, juries, appeals, burdens of proof -- and many substantive rules -- notably, concerning which categories of damages may be established and whether damages are determined by individualized evidence or by standardized tables -- may thus be illuminated by analysis similar to that presented here. A central concern when assessing these features of adjudication is parties' incentives in litigation to make expenditures on their own behalf: it often may be the case that parties, if freely permitted to undertake further efforts in litigation, will spend excessively because the private benefits from favorably influencing the outcome will exceed any social benefits from enhanced precision in applying legal rules.

*Related issues and literature.* The preceding discussion may be extended in a number of ways. Spier (1994) incorporates settlement bargaining. The interesting complication is that application of a more precise rule in adjudication may make relevant information that is possessed by only one party (asymmetric information), in which case greater precision may be accompanied by increased litigation costs due to lower settlement rates. Kaplow (1995a) briefly considers costly sanctions -- monetary payments made by risk-averse individuals and nonmonetary sanctions; few generalizations are possible. For risk-averse actors who do not become informed, greater differentiation may increase risk-bearing costs, although, as noted in Kaplow and Shavell
more precise damage awards may reduce the risks borne by uninsured victims. Kaplow (1990) examines the case in which some individuals know how the law differentiates among acts but others do not. In this case, there is no general deterrence rationale favoring lower sanctions for individuals who mistakenly commit violations, but reducing their sanctions may be optimal when sanctions are socially costly, because the lower rate of compliance by mistaken individuals results in sanctions being imposed more often. Kaplow (1995a) extends the analysis to regimes of self-reporting, as are common with much environmental and safety regulation as well as with the income tax. Self-reporting tends to be efficient when individuals' information costs are much lower than the government's; increasing precision in a self-reporting regime has the confounding effect of inducing some individuals to acquire information not to determine how better to behave but solely to determine how best to report their conduct.

Most applied work on complexity and compliance costs is in the field of taxation. Empirical work is extensive. See, e.g., Arthur D. Little (1988), Blumenthal and Slemrod (1992), Long and Swingen (1987), Slemrod (1989). Analysis of the benefits of precision in the tax context is qualitatively different from that presented here because the benefits of fine-tuning taxpayers' payments typically involve matters of distributive equity and minimizing adverse incentive effects of taxation rather than of optimally controlling harm-causing behavior. See, e.g., Kaplow (1994c, 1996), Yitzhaki (1979). The most obvious simplifying devices in the United States income tax code are the standard deduction (which taxpayers may take instead of itemizing particular deductions) and the use of floors (which forbid small amounts of deductions); they are analyzed in Kaplow (1994b) and Slemrod and Yitzhaki (1994).

The discussion here, like much analysis in law and economics, emphasizes the role of legal
commands in providing incentives to govern parties' behavior ex ante. Modifications to the analysis would be required when adjudication concerns future behavior, as in the granting of licenses, enjoining harm-causing behavior, incarcerating dangerous individuals, or determining eligibility for government benefits. In particular, the benefit of precision in adjudication will be greater than otherwise because the outcome will affect future activity directly (perhaps in addition to influencing ex ante incentives), and concerns about whether individuals will learn the relevant details of the law may be less relevant.

Finally, it is worth noting that the benefits of more detailed rules are relevant to wholly private ordering, as in contractual arrangements (contingencies can be distinguished more precisely, and provisions may provide reasonably complete specifications or be more open-ended) or in the internal organization of firms. As a result, work on the economics of contracting and organizations, wholly apart from analysis of the law, can illuminate issues concerning the design of legal rules, and conversely.

3. Rules versus Standards

The choice between rules and standards has long received attention from legal commentators and more recently has been addressed from an economic perspective. The presentation here will follow Kaplow (1992a); a number of the points, particularly concerning the effects of rules and standards on legal costs and on behavior, can also be found in the prior economic analyses by Ehrlich and Posner (1974) and Diver (1983).

The problem. An initial obstacle in analyzing rules and standards involves matters of definition. Usage varies greatly (which is no surprise since most dictionaries offer definitions of each term that use the other), and logically distinct issues are often combined or confused. For
purposes of economic analysis, it is useful to define the difference between rules and standards as involving exclusively the distinction between whether the law is given content ex ante or ex post. Thus, a rule might list hazardous substances that may not be released into the water supply whereas a standard may only proscribe releases of hazardous substances, leaving the determination of which substances are hazardous to adjudication, after releases have occurred.

This distinction between rules and standards is obviously one of degree, although it often is convenient to discuss it as an all-or-nothing choice. A law is more rule-like the greater the extent to which it provides advance guidance: it may, for example, provide a partial list or state criteria to be used, either of which will ease the application of the legal norm to particular behavior. Moreover, even a given formal specification will have a different character depending upon what is understood about the mode of adjudication and upon what other information is available in advance. For example, if everyone knows that an adjudicator is likely to rely upon a recent government study that indicates the degree of danger posed by potentially hazardous substances, the situation will be almost the same as that under a rule that incorporated the study’s results. For present purposes, such a situation will be viewed as essentially rule-like, as the economic analysis depends upon the extent to which content -- as understood by private parties and adjudicators -- has actually been provided in advance, not upon the wording of legal commands. (It thus can be seen that the generation and dissemination of information, even when not explicitly embodied in a legal rule, is an important aspect of the law.)

Finally, it is important to distinguish the question of the degree to which content is determined ex ante from the preceding subject, which concerns the degree of precision with which a law is formulated or applied. Often, commentators assume that standards are inherently more
precise or subtle than rules because standards leave open the possibility of considering all possible factors whereas rules, no matter how detailed, provide a definitive resolution that is not open to reconsideration. But this assumption implicitly contemplates that adjudicators will devote virtually unlimited resources in applying standards. Such an outcome is irrational, implausible, and inconsistent with the actual operation of legal systems. For example, tax codes and many safety regulations are far more detailed and context-dependent than, say, the results that would be produced by a jury that was presented with the open-ended question of what tax payment or safety regime was appropriate. As a result of law, custom, or the application of common sense, an adjudicator may well consider far less than all conceivably relevant factors. In an ordinary negligence dispute involving modest stakes, parties will not commission numerous expert studies; nor will the adjudicator commit substantial time to consider all possibilities.

Thus, it is a matter of choice how much detail will be incorporated in rules or standards, and it is hardly the case that rule-like systems tend to be particularly simple and standard-like schemes extremely precise in actual operation. The main point for present purposes is that the question of how much refinement is optimal when formulating and applying the law -- legal precision -- is logically and, to a great extent, practically distinct from the question whether a given degree of precision is best specified in advance (rules) or left to adjudication (standards). In the end, one can combine the following analysis of rules versus standards with the preceding analysis of the optimal precision of the law to determine what sort of legal command is best.

*Effects of rules and standards on legal costs and on behavior.* Rules are defined as legal commands that provide greater specification in advance. Hence, the initial cost of formulating rules will be greater than the cost of formulating standards. It will be necessary, for example, to
undertake studies to determine which substances are hazardous and to determine appropriate thresholds of safety if a rule listing hazardous substances is to be promulgated. Such efforts can be avoided (for the moment) under a standard that merely proscribes the release of hazardous substances and leaves their determination to subsequent proceedings.

On the other hand, rules will be cheaper for private parties to interpret when deciding upon their own conduct and for adjudicators to apply to past behavior. Precisely because more has been determined in advance, there will be less need for parties to conduct their own studies and to predict legal outcomes and for adjudicators to devote effort to determine whether a violation should be deemed to have occurred. If a rule specifies which substances may be discharged, actors and adjudicators need merely consult the list to apply a rule, whereas substantial inquiry may be necessary under a standard.

The same reasoning concerning the ease of application suggests that rules will tend to produce behavior more in conformity with legal norms than would equivalent standards. Because parties will be able to learn the governing norms more cheaply and accurately in advance, they will tend to behave more in accord with such norms. (This raises the question, explored in Kaplow and Shavell (1992) and noted above, whether private parties' incentives to learn about the law will be socially appropriate.) When guidance is easier to obtain, fewer parties will be deterred from behavior that actually is not objectionable, because they will have greater confidence that such behavior will not be subject to legal sanctions. More parties will be deterred from truly undesirable behavior because they are less likely to proceed in the mistaken belief that the prospect of being sanctioned is remote. It is important to observe that the desirable effect of rules on behavior may hold even when standards turn out to be more precise in actual adjudication, for
any refinement embodied in the application of standards will influence behavior only to the extent parties are able to predict the relevant details in advance, and often parties will not make such precise predictions. (The point that details tend to be helpful only when parties find it in their interest to learn them was emphasized in the preceding section on the precision of legal rules.)

Combining these effects, one can see that rules tend to be preferable when particular activities are frequent, and standards do best when behavior varies so greatly that any particular scenario is relatively rare. The reason is that the added cost of rules, at the formulation stage, is borne once, whereas the relative benefits of rules -- lower costs and greater compliance for private actors and lower costs of adjudication -- may arise often or almost never, as the case may be. Thus, for dry cleaning and automotive fluids, it would be best to specify permitted disposal techniques in advance, for the many small establishments would be unlikely to be able to determine the effects on their own and, even if they could, the costs of each obtaining information on the matter would be great; also, redundant costs would be incurred in inevitable enforcement proceedings. But for rare substances, it may not be worthwhile to specify the range of permitted and prohibited disposal techniques in advance because most scenarios involving potentially hazardous action may never arise.

In this regard, it should be noted that the frequency of activity may be great even if reported adjudications in a field are rare. For example, rules might lead countless individuals to comply with the law at very low cost. Even when compliance is imperfect (or when harm nonetheless arises and is subject to a regime of strict liability), there may be little adjudication despite frequent activity because the probability of harm is low, detection of harm is unlikely, or most legal disputes are quickly settled. Conversely, disputes may appear to be frequent but not
indicate that conditions are favorable to the use of rules. For example, there are many negligence disputes but some types may involve unique situations that, ex ante, were unlikely ever to arise.

Finally, the costs and benefits of rules and standards are are important for determining the optimal effort to devote to formulating rules and to applying rules and standards in adjudication. All else equal, the greater the frequency of the behavior to be governed, the more valuable it will be to formulate the law with greater care. Conversely, to the extent that an adjudication is relevant to only a single case, it will not be very valuable to expend substantial resources to consider myriad factors in order to reach a precise determination, unless a precedent it to be set, as discussed below. That individuals may have difficulty predicting refined decisions with any accuracy will strengthen this point. These considerations reinforce the previous observation that, in fact, rules often will be more precise and subtle than standards, taking into account how the latter will actually be applied.

_Precendent, other legal interpretations, and the evolution of legal commands._ Precendent may be seen as the (perhaps partial) conversion of a standard into a rule. Whether after accumulated experience or as a result of hearing a single case, a tribunal may issue a statement that will govern future adjudication. At that point, and to that extent, a rule will have been established. Similarly, even less authoritative legal interpretations, ranging from discussions in judicial opinions to public statements of enforcement agencies, will tend to guide future behavior and adjudications to some degree. Once the analysis of standards -- the state of affairs before the precedent -- and of rules -- after the precedent -- is understood, the study of precedent (and other legal interpretations) is straightforward. A few particular factors, however, deserve some elaboration.
It is worth asking why the use of precedent is superior to an earlier announcement of rules. After all, the primary burden of using a rule, incurring additional formulation costs, is ultimately borne when the precedent is established. But in the interim, before the precedent is announced, the benefits -- of cheaper and better guidance of behavior and simpler adjudication -- are forgone. This point is particularly striking when an important, open legal question is identified many years or decades before a precedent is finally announced. The primary advantage of delay is that more information may be collected in the interim -- information about the optimal legal command or about the importance of resolving an issue. But many precedents are announced with little more information than initially existed. Moreover, the most efficient way to collect information may be to commission a study in the first place rather than to wait until a case happens to appear before the relevant tribunal (when all that is then learned is the single data point from that case or perhaps a handful of data points). Also, once more information does become available through whatever means, it is not clear why the legal system should not make an announcement immediately rather than waiting still further.

It should be noted that making initial legal commands more rule-like does not imply that they cannot be changed if new information becomes available. The prospect of change does reduce the benefit of rules, but it does not follow that there will be no benefit from providing guidance in the interim. Indeed, many detailed regimes that are designed by regulatory agencies have a more rule-like character, with occasional revision, and detailed statutes, such as tax codes, are made precise but often revised later. (And, of course, precedents themselves are extended, narrowed, and reversed.) With standards, change may occur implicitly, but the prospect of change will accordingly increase the cost and difficulty of prediction. Substantial developments
that call for significantly different legal commands may be apparent, making prediction under standards easier, but also presenting a clear case for adjusting rules. It is hardly the case that operative legal commands in heavily rule-governed domains (like tax and environmental regulation) and the interpretation of important precedents change slowly compared, say, to applications of the negligence rule.

For prior economic analyses of precedent, emphasizing that the body of precedent can be analogized to a capital stock that depreciates over time, see Landes and Posner (1976) and Posner (1992). See also Landes and Posner’s (1994) discussion of the related subject of anticipatory adjudication, in which a party may seek a legal decision before undertaking an activity, so as to avoid the need to wait for a precedent.

Related issues and literature. Other law and economics work that is addressed directly to the subjects considered herein includes Ogus (1981) (discussing rules formulated in quantitative terms), and Ogus (1992) (emphasizing private actors greater ability to predict rules than standards).

The choice between rules and standards is closely related to the problem of legal uncertainty. Indeed, if rules are defined as the extent to which legal commands are given content ex ante and thus the ease with which private parties can predict how the law will apply to their conduct, legal uncertainty might be viewed as simply a measure of the extent to which a legal command is standard-like. Uncertainty about the negligence rule has received the greatest attention in the literature. See Calfee and Craswell (1984), Craswell and Calfee (1986), and Shavell (1987). Because the negligence rule has an all-or-nothing (discontinuous) character -- taking slightly more care will exonerate an actor from what otherwise would be significant liability
-- uncertainty about the legal standard may tend to produce excessive care (although it is also possible that great uncertainty will produce too little care). The suggestion that malpractice liability results in overly defensive medicine is one example. For empirical evidence, see Kessler and McClellan (1996). Kahan (1989) shows that these issues would not arise if negligent actors were not responsible for all the harm they caused, but only for the incremental increase in harm due to their negligence. Also, in considering legal uncertainty, it would be appropriate to extend the present analysis and that in prior work on the negligence rule to take explicit account of risk-bearing costs.

As noted in discussing legal precision, the discussion here focuses on the role of law in providing incentives to control parties' ex ante behavior. When future behavior (licenses, injunctions, incarceration, benefit eligibility) is involved, the analysis will differ. Indeed, the difference between ex ante and ex post formulation may be moot, as a definitive resolution will in any event precede the relevant effect of a legal decision. There are, nonetheless, similar questions about the benefits of wholesale decisionmaking, as in stating a rule for a class of cases rather than adjudicating the same question repeatedly. And the foregoing analysis remains applicable to the extent adjudication will influence ex ante incentives in addition to regulating future behavior (such as in the case of incarceration).

There are other functions of law as well. Much of contract law and related fields is concerned with providing rules of form or background rules that leave parties free to transact as they wish. See, e.g., Baird and Weisberg (1982). Because the purpose of such law is different -- law is largely facilitating rather than constraining -- the analysis would have to be different. Nonetheless, it seems that in contexts where transactions are frequent, the benefit of greater ex
ante formulation -- through rules -- would be larger.

Yet another concern with some bodies of law is to constrain fraud. In some instances (notably, tax evasion) greater specificity in advance might be problematic because parties will better be able to evade legal norms. Both uncertainty and complexity, discussed previously, may be desirable precisely because parties will be less able to predict the law.

Apart from normative analysis, one can inquire into lawyers' and legal institutions incentives to formulate legal commands. In this regard, conflicts between private and social interest may be noted. For example, lower legal costs save social resources but enhance the demand for lawyers. See, e.g., White (1992). And legislation might be designed with an eye to attracting campaign contributions rather than serving the public interest. See, e.g., Doernberg and McChesney (1987). Regulators may be more concerned with their own costs in formulating and applying rules than in private costs. See, e.g., Diver (1983).

Finally, the present discussion may be situated in the large literatures in the areas of legal philosophy and public choice which examine the rule of law and in particular address the use of rules to constrain official discretion. For a reasonably comprehensive treatment of rules that attends to functional concerns, see Schauer (1991).

4. Changes in Legal Rules: Transitions and Retroactivity

The problem. Transition rules, including a determination of whether a law is to apply retroactively, concern not the form and content of substantive legal commands but instead how the problem of legal change is to be addressed. The most common problem of legal change concerns the fact that investments and other actions often have effects in the future, at which time the governing legal regime may be different. Changes in the law will affect, positively or
negatively as the case may be, the returns from prior actions. This raises the question whether losses from legal change should be compensated or otherwise mitigated (as through grandfather provisions that exempt prior investments or phase-ins that implement the change only gradually) and, analogously, whether gains from legal change should be subject to windfall taxation or otherwise reduced.

As first analyzed carefully in the context of tax reform by Graetz (1977), these problems arise even when new legislation is nominally prospective. For example, if a tax subsidy is repealed, a prior investment that benefited on a continuing basis from the subsidy will become less valuable even if none of the prior subsidy payments must be returned to the government. The possibility of explicit retroactivity does, however, expand the potential scope for transition issues: even completed transactions might subsequently be rewarded or penalized when laws change.

The general problem of transition, analyzed in Hale (1927), Samuels (1974), and Kaplow (1986, 1987, 1992b), arises much more widely and with more common elements than is usually appreciated. Basically, whenever there is uncertainty concerning government policy -- about statutory change, new regulation or deregulation, court decisions, government spending priorities, takings of private property for public use, tariffs, and so on -- the resolution of the uncertainty will create winners and losers. Thus, all government action (unless fully anticipated), government inaction (that is not fully anticipated -- that is, when change had a positive probability that was not realized), and even changes in the prospect for action in the future will tend to affect the value of prior investments. The potentially affected investments include physical and financial assets, investments in human capital, and choices of location and occupation, among others.

As this general description suggests, an important aspect of the transition problem
concerns the ex ante choice of private actors -- individuals and firms -- under uncertainty. The prospect of legal change (like that of any other change) will affect both investment incentives and the imposition of risk, and these incentive effects and risk-bearing costs will in turn be influenced by whether transition relief is anticipated and provided. The immediately following analysis considers the question of optimal government transition policy, taking as given the actual change in the law that might occur and its desirability. The focus is on how transition relief affects the behavior and well-being of private parties and, accordingly, its effect on social welfare. Subsequent discussion will take up concerns about possible government misbehavior in changing the law and how different transition policies may influence government activity.

*Effects of transitions and transition relief on private actors' incentives and risk-bearing costs.* Begin by considering the case in which private actors are risk-neutral (which may be approximately true for large, widely owned enterprises that are the subject of much legal regulation). Suppose that there is some prospect, say, a fifty percent chance, of an adverse legal change: a firm's product might be banned if it turns out that the product is dangerous. Anticipating such a possibility -- and assuming for the moment that there will be no compensation, exemption, or other relief to those investing prior to the transition -- private parties will invest less than they would otherwise. Half the time they will earn some positive return and the other half of the time they will earn no return (and, indeed, they may lose their initial investment).

Now suppose instead that investors anticipate that there will be complete transition relief, say, full compensation. In this case, investors expect a positive return for sure, so they are more likely to invest -- or to invest at a larger scale -- than if there were to be no relief. As Blume, Rubinfeld, and Shapiro (1984) first showed formally in the takings context, the prospect of such
relief induces inefficient overinvestment. The key point is that a private loss in such instances corresponds to a real social loss. If a product is too dangerous and must be banned, sunk investments that are valuable only for producing such a product will be of no subsequent social value. Hence, insulating investors from such risks is inefficient, just as it would be inefficient to insulate investors from the prospect that their products might become obsolete or lose out to better products of competitors. As noted by Hale (1927), Samuels (1974), Graetz (1977), and Kaplow (1986, 1987, 1992b), from the perspective of investors, the risk of future government action is similar to various market risks and natural risks (fire, flood). Private decisions are efficient when actors bear the full costs and benefits of their actions. Transition relief thus does not correct an externality; it creates one.

This analysis of transition relief may be extended to determine when retroactive application of legal changes is efficient. See Kaplow (1986, 1987, 1992b). If a product is banned because it is discovered to be dangerous, it is probably the case that past production of the product resulted in danger as well. Thus, imposing retrospective liability will tend to be efficient, in order for investors ex ante to have had the proper incentives to take safety into account. By contrast, if a pollutant is to be regulated because it damages a newly introduced crop strain, there is no reason to impose retrospective liability, for the decision to regulate the pollutant does not suggest that the pollutant previously caused some social harm. To generalize, when the government discovers new information about conditions that have long been in existence, retroactive application in some form will often be necessary to provide the right ex ante incentives, but when there is a change in circumstances that makes future use undesirable without affecting the desirability of past conduct, retroactivity would be inefficient.
Similar reasoning suggests that transition relief is inefficient even when private actors are risk averse. See Kaplow (1986, 1987, 1992b). In this case, relief has the advantage of reducing risk-bearing costs. But private arrangements -- diversified ownership, insurance, contractual risk allocations, and the like -- also may be available. Moreover, these mechanisms have systematic advantages over transition relief. First, they often are implemented in ways that preserve incentives, in full or in part. For example, insurance premiums are greater the more assets are covered, so investors contemplating expansion in the face of some risk will bear the expected loss ex ante through higher insurance payments. In other cases, private risk mitigation is partial precisely to preserve incentives; absent other market imperfections (notably, adverse selection), private risk-sharing arrangements tend to be efficient. Transition relief generally lacks these features. Second, even if transition relief were only partial, private parties would undertake actions to further mitigate risk; such actions may worsen their incentives (moral hazard), but some of the resulting cost would be borne by the government. Again the analogy to market risks can be made: just as government relief (as in the case of disaster relief) tends to be inefficient for ordinary risks, see Arnott and Stiglitz (1991) and Kaplow (1991), so government relief against government-created risks tends to be inefficient with regard to private investment and risk-mitigation incentives.

Related issues. The analysis assumes that transition policy will be anticipated by investors and actually implemented by the government. To the extent such policy is embodied in constitutional provisions or strong norms that guide legal reform, the assumption is plausible. More generally, a government is unlikely to succeed in the long run in getting investors to expect one transition policy while often implementing another. Problems of credible commitment may
make it difficult to establish a desired transition policy. To consider a familiar analogy, some governments regularly offer ex post relief when natural disasters strike, which has the effect of encouraging individuals to, say, invest excessively in building on flood plains. A possible solution to this political commitment problem is for the government either to regulate such activity (limiting investment in high-risk areas) or to compel the purchase of insurance in advance at actuarially fair rates (the insurance premiums would discourage excessive investment and the fact that victims have paid for compensation in advance would eliminate the pressure for ex post relief).

Most of the discussion here, like most in various literatures that address transition issues -- ranging from tax reform to deregulation -- emphasizes the case in which investors suffer transition losses. But it was noted at the outset that transition gains arise with similar frequency. The analysis of both possibilities is symmetric: just as providing relief for losses will encourage overinvestment, imposing windfall taxes on gains (or otherwise reducing them) will produce underinvestment. Few if any windfall gains -- or losses -- are really pure windfalls. It tends to be best for investors to bear all the gains and losses they cause, with any risk mitigation being contractually arranged to provide for the most effective limitation of adverse incentive effects.

Another question of transition policy concerns the means of relief. One can compensate losers (and tax winners) in whole or in part, grandfather prior investors (that is, provide some exemption from new rules), or curtail the new policy (as by delay or phase-in or implementing reform only partially). With regard to ex ante incentives and mitigation of risk, all of these approaches are similar, to the extent they provide the same degree of effective relief. (For example, a phase-in that reduces gains and losses by half has effects like that of fifty percent
compensation and fifty percent windfall taxation.) With regard to future behavior, however, compensation and windfall taxation are generally superior to other forms of relief. Grandfathering, for example, might exempt existing producers of dangerous products, allowing the danger to continue (but perhaps not to grow); delaying implementation would similarly permit some future harm. By contrast, immediate elimination of the danger accompanied by compensation would avoid this difficulty.

As in the discussion of rule formulation, the present analysis is most applicable to laws that regulate behavior. Contract rules that serve as defaults are quite different. If a default rule is changed, it may well be that the prior rule should govern pre-existing contracts, for parties may have failed to make special provision in their contracts precisely because they knew they could rely upon background rules. On the other hand, if a default is changed because new analysis suggests that most parties have a different preference than previously suspected, and if many contracts are silent simply because of the cost of making special provision, it could be best to apply the new default to pre-existing agreements.

*Effects of requiring transition relief on government behavior.* The foregoing analysis takes as given a particular legal change, assumes that it is well conceived, and asks how transition relief affects private behavior. But it cannot be assumed that government always acts appropriately, and transition relief may affect the behavior of government.

One familiar argument is that transition relief -- in particular, requiring compensation -- provides the government proper incentives because then government bears the full costs of its actions. One can make the analogy to tort liability when both injurers' and victims' behavior affects accident costs. For analysis of this tort problem, see Shavell (1980), and for application to
the transition context see Kaplow (1986, 1987); for a similar discussion, see Blume, Rubinfeld, and Shapiro (1984). No liability -- or schemes that provide no compensation when injurers behave well, a negligence-type rule, or that deny compensation when victims behave improperly, contributory or comparative negligence rules -- is best for victims' incentives (like investors' incentives, above), whereas full liability -- strict or otherwise -- is best for injurers' incentives. If one can assume that the government (the "injurer") always behaves properly, then a regime that denies compensation is best for the incentives of investors (the "victims"). But if one cannot make this assumption about government behavior, one may wish to impose liability on the government -- for example, by requiring compensation. Unless some tribunal can determine when the government or investors misbehaved, which is likely to be very difficult in most transition contexts (and which involves courts second-guessing legislative, executive, agency, or their own actions), it must be recognized that any solution will be a compromise: if some relief is required, government's incentives may improve but investors' incentives will be worse.

This argument, however, is incomplete. See Kaplow (1986, 1987). The implicit assumption is that the government, like a private investor, receives the benefits from its actions. But the benefits of most government action -- banning hazardous products, improving tort rules, building highways -- are received by private citizens. Thus, making the legislature or an agency pay for transition losses (in addition to paying for direct costs, as in building a highway) when few if any of the benefits are directly received might systematically skew decisions against any change. Once benefits as well as costs are considered, it is not obvious that even a highly imperfect government is generally biased in favor of change when no compensation or other relief is required. For any bias -- whether for or against change -- to exist, there must be some asymmetry
in the government's treatment of costs and benefits. Some theories of government action, see Downs (1957) and Olson (1965), suggest that the most effective political interest groups are those on whom gains or losses are concentrated. Other theories are concerned with the majority exploiting the minority. Still others, see Niskanen (1971), emphasize bureaucrats desire to enhance their budget or prestige. Thus, whether requiring compensation would make government behavior better or worse depends upon which theories of government behavior are most plausible, whether it is the benefits or the costs that are more concentrated, and other factors. See also Quinn and Trebilcock (1982).

Another concern with government behavior involves the problem of pure expropriation. See Kaplow (1987). It is familiar from public economics that the most efficient government tax is a one-time capital levy, in which capital (more broadly, savings, assets in any form) is taken and the government promises never to repeat the action. As long as the expropriation in wholly unanticipated, there will be no adverse ex ante incentive effects; likewise, if it will never be repeated, there will be no future distortions in behavior. Levhore (1993) offers this as a basis for a more favorable view toward retroactive taxation. But it has long been recognized that governments would find it difficult to engage in such inconsistent behavior. Rather, there needs to be some commitment -- whether through constitutional means or long-established norms -- that the government will not expropriate, and governments that deviate from such a path will be unlikely to be able to provide investors with confidence for the future.

In principle, expropriation can be distinguished from other government action. As noted when addressing private incentives, it is efficient for investors to anticipate and respond to the prospect that new information or changed circumstances will affect what government regulation is
optimal, just as new developments will influence the market directly in ways that investors should take into account. Anticipating expropriation, however, is inefficient ex ante. The benefit arises from transferring resources to the government, not from providing proper incentives for private behavior. To be sure, some taxation is required, to fund public goods and services and to accomplish redistribution. Particularly with regard to redistribution, it would improve efficiency if taxpayers did not anticipate having to pay taxes, as such anticipation results in distortion of behavior. But this problem is largely unavoidable.

Successful economies seem to have succeeded, for the most part, in avoiding expropriation. The mechanism by which this is accomplished is not that clear; no system prohibits taxation or legal changes, and constitutional or other strong compensation requirements are limited. (For example, in the United States, the government must compensate a landowner whose building is taken for a highway, but a high capital income or wealth tax is not legally proscribed.) The development of strong norms seems to play a greater role in successful economies, but such norms cannot succeed unless the relevant audience can understand which legal reforms are improper in the relevant respect.

A variant of the expropriation problem arises when the government is involved with ordinary contractual behavior and related activity. In fact, most governments accomplish many of their purposes -- building roads, procuring military equipment, hiring employees -- through ordinary market transactions. (In some countries, state enterprises are common; these too, however, often use ordinary contractual means of paying workers and purchasing other necessary inputs.) Moreover, most governments, despite the availability of sovereign immunity from private suits, voluntarily subject themselves to a legal regime under which their contracts having a binding
effect similar to that of private contracts. This avoids problems of opportunism, both large and small. Some commentators have analogized legislation and other government law-making to government contracts. See, e.g., Graetz (1977), Kaplow (1986, 1987), Ramseyer and Nakazato (1989), Logue (1996). Thus, sometimes the government uses subsidies to induce certain behavior rather than explicit contracts under which it pays for a promise to undertake certain activity. If the subsidy is terminated, this might be viewed as a breach of an implicit contract. Of course, subsidies are routinely terminated. When governments use ongoing subsidies to encourage an activity rather than entering into a long-term contract to pay for a given amount of the activity, they are choosing to use a more flexible mechanism, allowing the possibility of change without compensation. To impose a norm or constitutional requirement that all law be treated as if it were a binding contract would not add to the government’s options; it would eliminate one. (Actually, it may have no effect at all, as laws could simply be written to expire unless extended or with some other loophole providing the government a way to avoid any binding obligation.)

Related issues. Transition relief is sometimes favored because it may make it politically feasible to implement desirable reform. For example, efficient deregulation might be opposed by protected incumbents, who could be bought out with compensation. See Tullock (1978), see also Usher (1995) (addressing compensation for takings). The problem is that, ex ante, special interest groups would have a greater incentive to seek inefficient rules, knowing that even if the rules were subsequently repealed they would benefit from transition relief. See Kitch (1977), Quinn and Trebilcock (1982), McKenzie (1986), Kaplow (1987, 1992b). Thus, it is not obvious that a policy of buying out such political opposition would in the long run produce better or worse laws and lead to more or less resources being wasted on rent-seeking activity.
Another problem concerns changes in political power that may produce legal changes that do not reflect real changes in information or circumstances (although legal reforms may reflect shifts in perceptions or in value judgments). Small changes in beliefs may produce large legal changes (as when an opposition party gains a few percentage points in support and comes into power), and such change may be frequent. Ex ante behavior that is desirable from the perspective of one party's policies may be undesirable from the perspective of the other's. Use of transition relief will tend to cause investors to put more weight on the status quo, something unlikely to be favored by a subsequent regime seeking change. Analysis of social welfare in such instances is difficult.

_Literature._ The first careful analysis of many issues concerning legal transitions appears to be Hale (1927). For a general analysis of changing legal rules in a variety of contexts, see Kaplow (1986, 1987, 1992b).

Most writing by legal academics on the subject focuses on takings or on tax reform. With respect to takings, Michelman (1967) offers an important early analysis, although it does not take the economic approach as far as subsequent work. Baxter and Altree (1972) advance a "first-in-time" approach to address the problem of airport noise; they are among the first to pursue directly the problem of the adverse incentive effects of compensation. Blume, Rubinfeld, and Shapiro (1984) analyze incentive effects of compensation in a formal model and Blume and Rubinfeld (1984) consider risk, although these authors do not integrate the two analyses or systematically consider private risk mitigation. Kaplow (1986, 1987) discusses takings in the context of analyzing the transition problem more generally. For additional work, with emphasis on the problem of the government's incentives, see, e.g., Fischel and Shapiro (1989), Hermelin (1995).
For an empirical analysis of the effect of requiring compensation on government behavior, see Cordes and Weisbrod (1979).

Feldstein (1976a,b) and Graetz (1977) offer the first substantial treatments of tax transitions, followed by Kaplow (1986, 1987, 1992b), who more explicitly analyzes effects of relief on investor's incentives and private risk mitigation, compares alternative transition mechanisms, and addresses institutional problems with government decision making. This latter problem is also considered by Ramseyer and Nakazato (1989) and Logue (1986), who are both favorable toward relief because of concerns about government misbehavior. Levmore (1993) is more favorable to (uncompensated) retroactivity because of the possibility of an efficient capital levy. The particular problem of transition to a consumption tax is addressed in Graetz (1979), Kaplow (1986, 1987, 1995b), Sarkar and Zodrow (1993), and Bradford (1996).

Another subject receiving increased attention is deregulation and other regulatory changes. See Quinn and Trebilcock (1982), Kolbe, Tye, and Myers (1993), Kaplow (1993).

5. Additional Topics

There are a number of other dimensions along which many legal rules vary that are sufficiently general to warrant independent study. Because these subjects are addressed in other fields within law and economics, the present discussion will be brief. Each of the topics considered involves the general question whether laws should identify certain conduct as permitted or prohibited or should instead impose fines, provide for private liability, or use other means of aligning private and social interests so as to induce parties to make efficient choices on their own.

*Strict liability versus negligence.* The choice between strict liability and negligence has
received much attention. Similar issues arise involving defenses based upon victim behavior. And other areas of law present analogous questions: namely, is a party responsible for all harm caused, or is the party responsible only if behavior was improper (and, then, only to the extent of any additional harm caused by the inappropriate action)? The latter type of law requires both defining a norm of behavior (which may be done ex ante, as a rule, or ex post, as a standard, and which may be done in a simple or precise manner), and both types of law require provisions governing the measurement of harm. This sort of choice involves many of the issues considered here, as well as concerns for the frequency of litigation, victims’ incentives, and other factors. See Brown (1973), Shavell (1980), and the entry on strict liability versus negligence in tort.

*Property versus liability rules.* The distinction between property and liability rules involves remedies: if one party wishes to infringe the legal rights of another, must the acting party first obtain permission (or else be subject to extreme sanctions and the possibility of injunction, if feasible) -- a property rule -- or may the party simply act and pay damages -- a liability rule? This scheme is first explored in detail by Calabresi and Melamed (1972) and has subsequently been expanded in Kaplow and Shavell (1996b). Other relevant work includes Ellickson (1973), Krier and Schwab (1995), Polinsky (1979, 1980). The problem of bargaining under asymmetric information has attracted disproportionate attention, beginning with Samuelson (1985), followed, e.g., by Farrell (1987), Shavell (1988b), Johnston (1995), Ayres and Talley (1995a, 1995b), Kaplow and Shavell (1995, 1996b). See the entry on property versus liability rules.
References


USA Tax System”, *Tax Notes* 68, 1109-1118.


