Revised September 7, 2009

REGULATION (AGENCIES) VERSUS LITIGATION (COURTS): AN ANALYTICAL FRAMEWORK

Richard A. Posner*

I. INTRODUCTION

Economic analysis of law treats common law fields, especially tort law, which provides legal remedies for physical, mental, or financial injuries caused by negligence, medical malpractice, nuisance (which includes pollution), defamation, defective products, misrepresentation, or other wrongful conduct, as forms of regulation. The emphasis is thus on the deterrent effect of the threat of liability, rather than on the compensatory role of liability; compensation is thought better provided for by insurance. Common law is thus conceived of as regulation by judges—by judges not only because common law remedies are obtained by means of lawsuits against injurers but also because common law doctrines are made by judges.

My purpose in this paper is to compare common law (including federal common law, that is, common law made by federal judges—indeed my primary interest is in federal regulation) with administrative regulation as methods of social control: more precisely, to compare the common law type of regulation with the administrative type, recognizing however that administrators often use common law methods of regulation and that judges sometimes use methods similar to those of administrative agencies. Nevertheless, judges are considerably more comfortable with the common law approach, and agencies that rely on common law methods to regulate are generally thought to have forgone the dis-

* Judge, U.S. Court of Appeals for the Seventh Circuit; Senior Lecturer, University of Chicago Law School. I thank Daniel Kessler for helpful comments on an earlier draft of this piece.

tinctive advantages of administrative regulation. So there is some utility to contrasting "litigation" with "regulation" as alternatives method of social control, while recognizing the overlap.

My analysis is normative; the question I address is what the better method—litigation or (administrative) regulation—would be, from the standpoint of overall economic efficiency, for regulating a particular activity. I leave to other work (some in this volume) positive questions about the choice between litigation and regulation, such as the political and cultural forces (including legalistic and individualistic traditions, and the influence of the legal profession, which has been said to be the American counterpart of European aristocracy and elite bureaucracy) that shape American government.

I assume that the concern is with regulatory problems that can't be left to the market to sort out—problems that involve large externalities that market forces cannot eliminate (that is, that transaction costs are too great for the Coase Theorem to apply). Even so, it is still necessary to consider whether public control is justified, because the costs may exceed the benefits in internalizing externalities, or because an intermediate form of regulation between pure market forces and public control may be superior to both. I refer to industry self-regulation, illustrated by board certification of physicians, hazing-type medical education to instill norms and create a "high commitment" environment ("professionalism"), contracts between patients and physicians and between consumers and producers, rulemaking and standards-setting by trade or professional associations, and arbitration or mediation to resolve disputes. If public control is not superior to private ordering, the next question—the positive question—is why the private alternative has been rejected.

II. Characterizing the Differences between Regulation and Litigation

Regulation and litigation tend to differ along four key dimensions: (1) regulation tends to use ex ante (preventive) means of control, litigation ex post (deterrent) means;(2)regulation tends to use rules, litigation standards; (3) regulation tends to use expert (or at least supposed experts) to design and implement rules, whereas litigation is dominated by generalists (judges, juries, trial lawyers), though experts provide input as witnesses; and (4) regulation tends to use public enforcement mechanisms,whereas litigation more commonly uses private ones (other than the judges themselves).

(1) *Ex Ante versus ex Post*

The first method is illustrated by speed limits, the seond by personal-injury suits for negligence. As in this example, the two types of regulation are frequently conjoined. The regulation of highway safety is a complex mosaic of ex ante regulation (speed limits and other safe-driving rules, federal safety design standards, standards for design and maintenance of highways, licensure of drivers) and ex post regulation (suits for negligent driving, product liability suits for defects in design or manufacture of cars, criminal prosecutions for drunk or other reckless driving).

To repeat an earlier point, ex ante regulation can be judicial as well as administrative, as in preventive detention, injunctions, and especially "regulatory [injunctive] decrees" (such as judicial administration of school systems with the aim of remedying racial discrimination), and ex post regulation can be administered by agencies as well as courts, such as the Federal Trade Commission and the National Labor Relations Board, which operate mainly by trial-type proceedings conducted after a violation of the laws administered by the agency has occurred.

Ex ante: pros. The ex ante approach promotes clarity by laying down rules in advance of regulated activities; it is activated before there is a loss, unlike a lawsuit; it can be centrally designed and imposed (for example, by a single agency such as the FDA, rather

than by a decentralized judicial system); and it is enforceable by means of light penalties, because the optimal penalty for creating a mere risk of injury is normally lighter than the optimal penalty for causing an actual injury. (Rules do require enforcement proceedings, as compliance is never 100 percent; so ex ante and ex post regulation are inseparable.) Rules involve heavy fixed costs (designing the rule in the first place) but, if they are very clear and carry heavy penalties, low marginal costs—compliance is achieved with few proceedings. So rules are attractive when the alternative would be vague standards, resulting in frequent actual or arguable violations and hence frequent enforcement proceedings.

As this discussion shows, ex ante regulation and rules have an affinity. Ex ante regulation enables exploitation of the economizing properties of rules as preventives. With vague standards, the regulatory emphasis inevitably shifts to seeking deterrence by proceeding against suspected violators.

But the affinity between ex ante regulation and rules requires a qualification. Consider the criminal penalties for the sale of illegal drugs. The underlying criminal prohibition is a flat, clear rule, but compliance is achieved almost entirely by threat of punishment, which is ex post. Contrast that with the regulation of legal drugs, where, although there is ex post enforcement, including products liability suits, the emphasis is on testing new drugs in advance for safety and efficacy, and refusing to allow drugs to be sold that flunk the tests.

Ex ante: cons. Ex ante regulation narrows the information base, because, when it takes the form of rules, it buys precision at the cost of excluding case-specific information that the promulgators of the regulation either did not anticipate or excluded in order to keep the regulation simple (i.e., a rule). Standards, such as negligence (versus rules, such as a numerical speed limit), allow much more information to be considered in particular cases, but in doing so not only reduce predictability but also, as noted above,

merge into ex post regulation. In addition, ex ante regulation, like preventive care in medicine, can burden much harmless activity, such as safe driving in excess of the speed limit. (Compare screening the entire population for medical conditions that afflict only a small percentage of the population.) This is related to the fact that rules exclude relevant circumstances for the sake of clarity.

Ex post: pros. Ex post regulation may require only rare interventions (again compare screening for medical conditions with treatment if and when a condition produces symptoms)—zero in the limiting case in which a rule or standard achieves 100 percent compliance, though there may of course be costs of compliance. Ex post regulation economizes on administrative apparatus because intervention is sporadic, and utilizes both case-specific information (including information about causation and victim fault, and other information obtained after regulation is promulgated and in the context of a particular injury) and adversary procedure, which may increase regulatory accuracy.

The earlier example of illegal drugs illustrates the case in which ex post regulation does not refine a pre-existing rule or standard. Drug enforcement is purely punitive—and largely ineffectual, because, although the penalties are stiff, the expected cost of punishment is for many potential offenders low relative to the expected profits of drug trafficking, because of the ease of concealment of illegal activity—a general problem in "victimless" crimes, since there is no one to complain to the authorities.

Ex post: cons. Ex post regulation, typified by common law adjudication with its heavy emphasis on standards, such as negligence and good faith, in preference to rules, involves high costs per case, compared to adjudicating a speeding ticket. This is partly because of the additional information generated by a proceeding focused on a specific injury inflicted in particular circumstances. More information can not only make a proceeding more costly but also create more uncertainty and as a result more variance in outcome; uncertainty also makes it more difficult to moni-

tor the judge or other regulator. And, a point related to the fact that the optimal penalty when an injury has occurred is greater than when a risk has been created that has not yet materialized, the injurer may not have sufficient resources to pay the penalty. There are also problems of proof when a causal relation to injury must be proved, illustrated by cases in which exposure to radiation increases the incidence of cancer but it is impossible to determine whose cancers were due to the radiation and whose would have occurred anyway. This problem can be solved, however, at least in principle, by class actions that amalgamate claims of probabilistic injury of all persons who had been exposed to the hazard in question.

Since deterrence is unlikely to be 100 percent effective, ex ante regulation is strongly indicated when the regulated activity can give rise to catastrophic injury. The greater the injury if deterrence fails and the likelier deterrence is to fail, the stronger the case for ex ante regulation. Even if 99 percent of building collapses, but only 10 percent of drug offenses, can be prevented by ex post regulation (suits for negligent design or construction in the first case, criminal punishments in the second), the social cost of the 1 percent of building collapses may greatly exceed the social cost of the 90 percent of drug offenses, and if it also exceeds the cost of prevention by building codes, then ex ante regulation is justified in the first case, and probably not in the second (no one has good ideas about how to prevent the sale of illegal drugs, other than by prosecution of sellers or buyers). And the gravity of the injury and the likelihood that deterrence will fail are positively correlated, because the limited solvency of potential injurers may cause the expected cost *to them* of the injury to be far below the expected social cost.

This points helps to explain the different regulatory systems for new drugs and for medical procedures. A drug sold to millions of people can, if it is unsafe, wreak enormous harm, whereas individual cases of medical malpractice injure only one patient.

Moreover, it is feasible to test every new drug, and thus determine safety in advance; it is infeasible to require physicians to seek approval from a regulatory agency for every procedure that they perform. So ex ante regulation is the dominant mode of regulation of new drugs, while ex post regulation in the form of medical malpractice suits is the dominant mode of regulation of medical treatment. Obviously, medical education and apprenticeship (residency) play a big role in preventing malpractice, but that is not the focus of the training.

(2) *Rules versus Standards*

I elaborate here on the comparison made earlier between rules and standards as regulatory techniques, where I noted the (loose) association between rules and ex ante regulation and standards and ex post regulation.

A rule abstracts from a number of relevant facts (as in a numerical speed limit, which ignores other circumstances bearing on the danger caused by driving). A standard is open-ended because it directs the judge or jury or other regulator to consider the particular circumstances in which a violation is alleged.

Rules: pros. They tend to be simple and clear, which reduces enforcement costs and facilitates monitoring of the court or other agency that applies the rules to particular facts. The simplicity of rules and the ease of monitoring compliance with them make them especially attractive for societies in which the judiciary is prone to incompetence and corruption.

Rules: cons. Yet often rules are not really simple and clear, because of pressure for exceptions and the boundary issues created by exceptions; it may be unclear whether a particular case falls within the general rule or within one of its exceptions. The answer to such a question is usually found by considering the purpose behind the rule and the exception in question, and that is the sort of analysis employed when standards are being applied.

Rules tend also to be crude, because they exclude relevant facts (such as, in the speed-limit example, traffic conditions, weather and time of day, emergencies, and driver skills); they thus rest on a narrower information base than standards. That exclusion also makes them somewhat arbitrary, and as a result counterintuitive. "Being careful" is intuitive; driving below 50 m.p.h is not, which is why speed limits have to be posted. Rules, in contrast to standards, tend also to separate rule creation from application: legislature promulgate rules, courts apply them. Common law courts both create and apply standards, and there are some efficency gains from vesting both functions in the same organization.

Standards: pros. Standards are the inverse of rules, so that the disadvantages of rules become the advantages of standards. They are flexible; intuitive; consider more information, including information generated after the standard was initially adopted (that is a serious problem with rules—they exclude from consideration factors the significance of which was not realized when a rule was promulgated); and they facilitate merger between the maker and the applier of the standard—it is often the same entity, namely the same court.

Standards: cons. Similarly the advantages of rules show up on the other side of the ledger as the disadvantages of standards. They are vague; costly to administer because open-ended; and difficult to monitor compliance with by the court or other body that enforces the standard.

Notice that despite the association of rules with legislatures and rulemaking with administrative agencies, rules can be judicial (an example is the judge-made rule entitling a criminal suspect to a probable-cause hearing within 48 hours of arrest), and standards can be administrative (examples are police discretion in enforcing speed limits and the use of broad standards such as "unfair labor practice" and "unfair or deceptive acts or practices" by the NLRB and the FTC, respectively). Also, standards can be ex ante, as in

the safety and efficacy standards used by the Food and Drug Administration to decide whether to approve a new drug. And rules can be ex post, for example when a rule is declared by a court for the first time in a case in which the parties did not anticipate the rule; nevertheless it binds them, as well as others who may have violated the rule before it came into existence. Indeed judicial rulemaking is characteristically ex post.

(3) *Agencies versus Courts*

Agencies: pros. Agencies are specialized, and this facilitates the development of expertise in technical subject matters (examples are traffic safety departments prescribing speed limits and the Food and Drug Administration regulating pharmaceuticals). They usually have large staffs and flexible powers—often they are authorized to engage in both ex ante and ex post regulation. They are less hobbled by precedent than courts are. The reason is that agency members have more political legitimacy than judges and thus have less need to avoid being "activist" and demonstrate continuity with past political settlements. Judges are reluctant to innovate, or at least to seem to innovate, lest they be accused of crossing the line that separates applying law from making law, the latter orthodoxly considered a legislative rather than judicial function.

Agencies: cons. Agencies are subject to far more intense interest-group pressures than courts. The agency heads are political appointees and their work is closely monitored by congressional committees. As a consequence of their large staffs, they are bureaucratized and exhibit the pathologies of that form of organization. And generally agency regulation is really dual agency-court regulation, because agency rulings are appealable to courts.

Courts: pros. Courts are relatively immune to interest-group pressures (at least federal courts, whose judges have secure tenure, and some state courts); nonbureaucratic; decentralized; and semi-privatized (because of the huge role played by the litigants'

lawyers). They bring to the table an outsider's perspective on issues that regulators, afflicted with tunnel vision, might botch. Judges are also less mission-oriented than regulators. Being generalists, and coming from diverse professional as well as personal backgrounds, they are less likely to identify with particular policies, and hence bring a more balanced approach to issues than regulators, committed to a particular policy, do. If an agency were established to eradicate drug trafficking, and was given the authority to try violators of the drug laws, it would probably give short shrift to procedural safeguards for accused violators.

Courts: cons. Judges in the Anglo-American judicial systems are among the last generalists in an increasingly specialized government and society, and this is a source of weakness as well as of strength. The judges' lack of specialized knowledge, their limited staffs, limited investigatory resources, cumbersome and to a degree antiquated procedures, commitment to incremental rulemaking, and delay in responding to serious social problems—courts cannot act until a case is brought, which often is long after the practice giving rise to the case began—are impediments to effective regulation, especially of technical subject matter. These problems are aggravated by the heavy use—idiosyncratic by world standards—of juries in civil cases. When technical issues are committed to courts, such as issues concerning medical malpractice, product-design defects, and patents on drugs or software, the results are often unsatisfactory. The costly practice of "defensive medicine," a response to the threat of malpractice liability, is an example of costs resulting from the commitment of technical issues to generalist judges and jurors bound to make many errors. But it is unclear whether the costs of defensive medicine outweigh the benefits of tort liability in creating increased incentives to exercise care in medical treatment.

(4) *Public versus Private Enforcement Mechanisms*

The common law litigation system, as indeed any private-law system, depends on private individuals and firms to activate the system. The award of monetary damages as the standard outcome of a successful private suit provides the incentive for a private party to sue. The common objection that because of the expense of litigation victims of small harms (though they may be great harms when cumulated over all victims) has been overcome or at least diluted by the class-action device, which allows the aggregation of small claims to create a prospective damages award large enough to motivate a suit. Penalties of various sort can also be annexed to compensatory damages in order to increase the private motivation to sue.

Nevertheless litigation is very costly in the United States, and quite slow as well; and the existence of the criminal laws is proof (if any is needed) that damages awards (or injunctions, the other common remedy awarded to a plaintiff who prevails in a private suit) are not always an adequate device for controlling behavior. Hence the rise of regulatory mechanisms that do not require recourse to litigation, although judicial review of the result of the regulatory proceeding is typically available. The rise of the administrative agency, beginning with the creation of the Interstate Commerce Commission in 1887, reflected a desire to increase the role of expert knowledge in regulation and to counteract what was regarded as stubborn judicial resistance to modern social-welfare policies, but also to provide cheaper and more expeditious remedies administered by civil servants, a key element of which was to vest the agency, not a private individual, with the decision whether to initiate proceedings. Gained was a degree of expertise, expedition, and procedural and remedial flexibility; lost was the superior ability of most judges (at least federal judges) to agency administrators, their greater sensitivity to rule of law values, and the energy and initiative of private persons and firms affected by regulation.

III. Litigation and Regulation in Practice

(1) *Pure versus Mixed (Hybrid) Systems (Corner versus Interior Solutions)*

A pure system of regulation would be only administrative regulation or only litigation; a mixed system combines the two modes of control. There are virtually no pure regulatory systems, because most regulatory decisions by administrative agencies are subject to judicial review, though that is a more limited form of judicial intervention than a proceeding that begins in court rather than in an agency.

Nearest to a pure system of administrative regulation is a system in which compliance with a regulatory rule or order precludes subsequent lawsuit (preemption). The financial industry has the closest approach to a pure regulatory regime. Medical malpractice approaches a pure litigation system, except that there is some regulation of hospital and physicial practices and of course there is licensure. Antitrust approaches a pure litigation system too, despite the merger guidelines published by the Department of Justice and the FTC, which provide a basis for advance determinations by the agencies whether to approve proposed mergers; but the guidelines are nonbinding.

Pure: pros. A pure system is cheaper; simpler; operates much more quickly; and provides better guidance. The need for speed, well illustrated by the response of the Federal Reserve and the Treasury Department to the financial collapse in September 2008, can be a compelling reason for a regulatory system in which courts play little or no role.

Pure: cons. But when time is not of the essence, a pure system of administrative regulation has many disadvantages. For one thing it increases incentives for and therefore the likelihood of regulatory capture by interest groups, since an interest group has only to "buy" an agency, and not the courts as well (and the federal courts are very difficult to "buy"), assuming judicial review of agency actions does not disturb the basic orientation of the

regulatory agency. Also, a pure system of either kind (agency or judicial) sacrifices complementarities, because courts and agencies are complements as well as substitutes.

Mixed: pros. A mixed system, as just mentioned, exploits complementarities between agencies and courts. Sentencing judges fine-tune sentencing guidelines; antitrust judges fine-tune Justice Department or FTC merger guidelines; and judges review the rulings of administrative agencies for compliance with statutes and with principles of fair procedure, which are subjects that judges have more experience with than administrators. A mixed system is also less susceptible to capture by interest groups because in a mixed system the interest group has to buy both the agency and the courts. Andn it provides a back-up or fail-safe regulatory capability. In regard to drug safety, for example, when the Food and Drug Administration fails to prevent the sale of an unsafe drug, the tort law of products liability provides an alternative, ex post control over the sale of the drug

Mixed: cons. A mixed system conduces to delay and uncertainty of outcomes and imposes costs of duplication.

(2) *Competitive Regulation*

Often more than one agency regulates the same activity. Both the Justice Department and the FTC enforce the federal antitrust laws, and in addition state attorneys general enforce state antitrust laws modeled on the federal laws and applicable to the same enterprises. Private suits can also be brought to enforce both the federal and the state antitrust laws. To complicate the picture still further, state attorneys general can bring federal antitrust suits on behalf of their states. Regulatory competition increases the likelihood that a violation will be detected and punished, but also increases compliance costs for the firms subject to the dual or multiple regulatory regime.

"Regulatory arbitrage" refers to the unedifying practice of firms' configuring their businesses in such a way as to bring them

within the regulatory jurisdiction of an agency likely to favor the firm, perhaps because the agency is supported by fees of the firms it regulates and therefore, to increase its budget, seeks to entice firms by an implicit promise of light regulation. Thus, a bank might decide to seek a state rather than federal charter because it thought the state banking commissioner would be more tolerant of the bank's loan policies than a federal banking regulator, and the commissioner might welcome the newcomer because of the effect on the commissioner's budget of having a new fee-paying "client."

It is difficult to generalize about the choice between monopoly and competitive regulation. In the case of safety regulation, it is common to allow states to impose stricter safety standards than the federal regulators, although the federal regulation will invariably be deemed to preempt state regulation that contradicts the federal, so that if applicable it would impose inconsistent duties on the regulated firms.

A competitive system should not be confused with a mixed (regulation plus litigation) system. In the mixed system, the different regulators (administrative agency and court) are regarded as complementary rather than competitive, although there may of course be disagreement. In a competitive system, two agencies (or two judiciaries) may find themselves empowered to regulate the same activity, and the hope is that the competitive setting will keep each on its toes.

Employment discrimination is a good example of competitive regulation. The federal state laws are similar, and they are enforced both by federal courts and federal agencies (principally the Equal Employment Opportunity Commission), and by state courts and state agencies. Pollution is handled quite similarly.

Whether the benefits of competitive regulation exceed the costs in all, most, or any settings is unclear. The principal reason for the competitive system is simply the constitutional status of the states. They are not merely bureaucratic subdivisions of the

national government, but instead quasi-sovereignties that make and administer their own laws until Congress or the federal courts intervene to prevent actual conflicts with federal law. Competitive regulation within a state or within the federal government, illustrated by the dual enforcement of the federal antitrust laws by the Department of Justice and the Federal Trade Commission, is rare.

(3) *Comparative Analysis*

There are significant differences in regulatory institutions and procedures across countries and also across states of the United States and between federal and state governments. For example, civil-law courts are much like U.S. regulatory agencies (bureaucratic, rule-bound), and state courts are on average more politicized than federal courts. In general, rules are more important, and standards less important, in civil-law than common law countries. Hence mixed systems in civil-law countries are likely to involve fewer agency-court complementaries than in common law countries.

III. Reforming Existing Regulatory Regimes: Transition Costs

Suppose some new area of activity is sought to be brought under regulation; or there is dissatisfaction with the scope or implementation of an existing regulatory system. The choice then is often between seeking to reform the existing system or creating a new system. In the usual case this comes down to a choice between tinkering with an existing agency (its powers, resources, leadership, or staff) and creating a new agency.

Tinkering with the existing agency: pros. This has the advantage of speed, economy, avoiding turf warfare (the creation of a new agency is likely to step on bureaucratic toes by taking powers from or competing with other agencies), and avoiding increasing the complexity of government. Also, it is easier to rescind changes

in an existing agency, if they prove unsound, than to abolish an entire agency, which will have developed a constituency in Congress or among interest groups.

Tinkering with the existing agency: cons. Agency staff, having civil service tenure protection, may be bold in resisting change and may resist it effectively, with assistance in the form of backing from members of Congress and interest groups. Giving an agency new responsibilities may reduce its ability to perform its old responsibilities and create tension between staff assigned to old responsibilities and staff assigned to the new ones. Seniority considerations may given "old timers" significant positions in administering new programs with which they are unsympathetic.

Creating a new agency: pros. Creating a new agency is a strong signal of a new departure and may attract committed leaders and staff from outside the existing governmental bureaucracy. Exclusively committed to the new programs that gave rise to the new agency, leaders and staff will be judged by the success of the programs and will not be able to bury them in a bureaucracy that has many other programs and constituencies to attend to.

Creating a new agency: cons. These are the converse of the pros of tinkering with an existing agency. Creation of the agency will be time-consuming and involve struggle with existing agencies and their backers in both Congress and industry, will be difficult to reverse, and will increase the complexity of government.

CONCLUSION

The costs and benefits of the different control institutions and techniques have changed over time. The optimal (and actual) mixture has therefore changed. For example, diseconomies of scale in litigation (a court system is pyramid-shaped to maintain uniformity, and if there is too much litigation too many layers of review are required, creating unacceptable delay and confusion) may require the creation of regulatory alternatives to litigation. And the rise of public finance as a consequence of more efficient methods

of taxation has made regulation, which is more costly to the government than litigation (which is largely financed by the litigants themselves), more feasible. Rising information costs because of greater technological complexity may also increase the gain to expertise and hence the comparative advantange of specialized agencies relative to generalist courts.