Chicago Law and Economics

> A second meaning of "justice," and the most
> common I would argue, is simply "efficiency." When
> we describe as "unjust" convicting a person without
> a trial, taking property without just compensation, or
> failing to require a negligent automobile driver to
> answer in damages to the victim of his carelessness,
> we can be interpreted as meaning simply that the
> conduct or practice in question wastes resources.
> *(Posner 1975, p. 777)*

INTRODUCTION

WHILE the roots of Law and Economics go back at least to Adam Smith and
Jeremy Bentham,[1] it came of age as an intellectual discipline within economics
and law in the 1960s and 1970s through the work of such notable figures as
Ronald H. Coase, Guido Calabresi, Henry Manne, Gary Becker, and Richard
A. Posner.[2] The work of these scholars—of whom Posner as professor, scholar,
and judge is perhaps the foremost exponent—forms the core of the Chicago ap-
proach to law and economics, an approach that has attracted a large following
and has come to dominate scholarship within the economic analysis of law. For
the purposes of this discussion, the Chicago school is deemed to encompass
most of the scholarship falling under the umbrella of "mainstream law and
economics," as reflected in the pages of, for example, the *Journal of Law and
Economics*, the *Journal of Legal Studies*, the *American Law and Economics
Review, Research in Law and Economics*, and many of the major law reviews.[3]

This having been said, the history of law and economics at the University of
Chicago begins well before the birth in the 1960s of the "new law and econom-
ics" that is now synonymous with the Chicago school label. The distinction that

---

[1] For Smith, see his *Lectures on Jurisprudence* (1978); on Bentham, see Posner (1998a; 2001b
ch. 1).

[2] Coase, Calabresi, Manne, and Posner were honored as the "four founders" of law and eco-
nomics at the Plenary Session of the American Law and Economics Association on May 24, 1991.
Coase and Becker have been awarded the Nobel Prize in economics.

[3] This is also the approach to law and economics that is reflected in the major texts in the field,
including Posner (2003), Cooter and Ulen (2004), Hirsch (1999), Shavell (2004), and a compre-
hensive casebook by Barnes and Stout (1992).

is often made between the "old" and "new" Chicago law and economics within the Chicago oral tradition is part folklore and part fact,[4] and, to appreciate the nature of this distinction, it is useful to mention a few of the highlights of this earlier era of economics and law. As will become evident, although some of the early history and subsequent success of Chicago law and economics is attributable to scholars in the Department of Economics at Chicago, much of the credit owes to the intellectual edifice provided by the faculty of the law school.[5]

The Law School

As we noted in chapter 1, the Realist-Institutionalist interaction of the 1920s and 1930s did a great deal to bring law and economics together. A similar interaction, but with a distinctly different flavor, commenced at nearly the same time at the University of Chicago. Within the law school, the origins of Chicago law and economics can be traced back to the latter part of the 1930s, when the faculty, under the deanship of Wilber Katz, instituted a four-year interdisciplinary legal studies curriculum that included courses in economics and accounting (Katz 1937). In 1939, the law school appointed its first economist, Henry Simons, to the faculty. Simons had been a lecturer in the Department of Economics and was a former student of Professor Frank Knight, who was in many respects the father of the modern price-theoretic tradition of Chicago economics (Kitch 1983a, p. 167).[6] Simons's appointment to the law school was the culmination of many episodes of political infighting between Frank Knight and Paul Douglas within the Department of Economics. Douglas—a strong proponent of Keynesian thinking and advocate of government intervention in the economy—and others within the economics department fought both Simons's promotion and tenure. As a result, Simons was ultimately appointed to the law faculty to teach a course titled "Economic Analysis of Public Policy" (essentially a course in applied microeconomics for lawyers), and, in 1945, he was promoted to full professor and became the first economist granted tenure by the law school. By the mid-1940s, Simons

[4] The use of the "old" and "new" terminology is evidenced in Posner (1975, p. 759) and Kitch (1983a) and is explored more fully in the following section.

[5] This brief survey of the Chicago school of law and economics prior to 1960 draws on the work of Reder (1982), Duxbury (1995, ch. 5), Kitch (1983a), and Coase (1993); see also Hovenkamp (1995) and Medema (1998, 2007a).

[6] Frank Knight's students included, among others, Milton Friedman, James M. Buchanan, Aaron Director, and George Stigler, all of whom, with the exception of Director, have received the Nobel Prize in Economics. Gary Becker was a student of Jacob Viner, another stalwart in the economics department in the early days of the Chicago school and later a professor at Princeton, where Becker did his undergraduate work. Becker then went on to become a professor at the University of Chicago and was later awarded the Nobel Prize in economics.

had established himself as a contributing, but by no means seminal, figure at the headwaters of the Chicago law and economics movement,[7] and his view that law should be structured so as to promote competitive markets reflected the perspective that has come to be associated with Chicago law and economics. His tenure at the law school helped to lay a solid foundation and provide a receptive environment for the then-emerging field of law and economics.

Simons helped to ensure the continuity and growth of Chicago law and economics through his role (together with Friedrich A. Hayek and financial backing from the Volker Fund) in bringing to the law school the individual most responsible for firmly establishing the law and economics tradition at Chicago, Aaron Director. Director was a student of Frank Knight, had been a member of the economics faculty in the early 1930s, and had authored articles that had a distinct Chicago price-theoretic flavor.[8] After being away for a short time, Director returned to the University of Chicago in 1946. Simons urged the law school to have Director assume the directorship of a university center (affiliated with the law school) dedicated to undertaking "a study of a suitable legal and institutional framework of an effective competitive system" (Coase 1993, p. 246), and he was appointed to that position in 1946. Director also took over responsibility for teaching the course on "Economic Analysis of Public Policy" from Simons, who died in the summer of 1946.

Subsequent to his law school appointment, Director was invited by Edward Levi to collaborate in the teaching of the antitrust course. Director continued teaching the course throughout his time at the law school, and it was during this period that the Chicago approach to antitrust law began to emerge as a distinct body of thought. Antitrust had long been an area of law particularly open to the influence of economic ideas, and it was through his teaching in this course that Director, armed with the tools of price theory, had an especially formidable impact on both Chicago law students and a number of his colleagues. Director's students included Robert H. Bork, John S. McGee, Henry Manne, Edmund Kitch, Bernard H. Seigan, and Wesley J. Liebeler; colleagues who were impacted included Ward Bowman and Lester G. Telser. Director's antitrust students, many of whom went on to very successful careers as economists in their own right, speak of the "conversion" that they experienced in this class.[9] Levi would teach a traditional antitrust course for four days each

[7] In his review of the history of law and economics at Chicago, Coase stated that "[w]hat can, I think, be said with confidence is that Simons . . . played little or no part in the development of the ideas which make up the modern subject of law and economics" (Coase 1993, p. 242). Simons' work on tax law, however, is very worthy of note in this context, and Posner continues to cite it in his discussion of tax law in his *Economic Analysis of Law*. See, for example, Simons (1934, 1938, 1948, 1950) and the discussion by Groves (1974).

[8] See, for example, Director (1933).

[9] This "conversion" is openly discussed throughout "The Intellectual History of Law and Economics" conference as recounted by Kitch (1983a).

week, and then Director would come in on the fifth day and, using the tools of price theory, show that the traditional legal approach could not stand up to rigors of microeconomic analysis. As Posner points out,

> The "kinked demand curve," "workable competition," "cut-throat competition," "leverage," "administered prices," and the other characteristic concepts of the industrial organization of this period had this in common: they were not derived from and were often inconsistent with economic theory, and in particular with the premises of rational profit maximization. They were derived from observation, unsystematic and often superficial, of business behavior. Director's approach was the opposite. He explained tie-ins, resale price maintenance, and other business behavior described in antitrust cases not by studying the practices but by looking for an explanation for them that squared with basic economic theory. (Posner 1978, p. 931)

The result was, as Liebeler put it, "We learned that there was a system of analysis [price theory] that (1) was quite relevant to the stuff we talked about in law school and (2) was much more powerful than anything that the law professors, than anything Ed Levi had to tell us" (quoted in Kitch 1983a, p. 183).[10]

Director taught his students that the appropriate goal for antitrust law was the promotion of efficiency. Reflecting the emphasis within the Chicago tradition on the efficacy of the competitive system, monopoly was viewed as occasional, unstable, and transitory—a potential outcome of the competitive process, but one that would soon be removed (in effect if not in existence) by competitive pressures. Given this, rigorous antitrust enforcement was thought to be unnecessary, and, even when monopolies were shown to generate long-term inefficiencies, the governmental cure was thought to be often worse than the disease, owing to the inefficiencies of government. The past several decades of legal-economic scholarship bear witness to Herbert Hovenkamp's (1986, p. 1020) comment that "the Chicago School has done more for antitrust policy than any other coherent economic theory since the New Deal. No one . . . can escape [its] influence on antitrust analysis." Indeed, and as we will show later in this chapter, the Chicago school has become so firmly established within this area that it can claim a partial victory in what Eleanor Fox (1987a) has called the "battle for the soul of antitrust."[11]

Director's impact at the law school, however, went far beyond antitrust: he was the prime mover in the early professionalization of Chicago law and

---

[10] The transcript of the conference indicates that, over the course of time, Professor Levi also became converted (Kitch 1983a). Indeed, in 1951 Director and Levi (1951, p. 282) wrote, "We believe the conclusions of economics do not justify the application of the antitrust laws in many situations in which the laws are now being applied."

[11] See also Fox (1981).

economics. Director formally established the nation's first law and economics program (derivative of the school's antitrust project), which maintained visiting fellowships for law and economics scholars, set in place the workshop in the law school (an arena of vigorous debate over issues of current research, already an established tradition within the economics department), and, in 1958, founded the *Journal of Law and Economics*. Throughout his tenure at Chicago, he imparted a persuasive message to the students—that regulation was the proper function of markets, not government. This message was one that often resulted in legal reasoning losing out to economic analysis and was, to the students of law, "a message which was at once both unfamiliar and yet quite understandable" (Duxbury 1995, p. 344).

Kitch, a Chicago-trained lawyer (1961–64) and later the director of the Law and Economics Program at Chicago, describes the thrust of law and economics at the Chicago law school during this era in terms that make it seem like a perfectly natural activity for law school professors to pursue—a view, in fact, very consistent with the rationale behind the Legal Realists' blending of economics and law:

> The interest . . . in economics did not come out of any anti-interventionist thinking. It essentially came out of the idea that the legal system is going to be doing this now [i.e., impacting the economic system] and that means we need to learn how to do it right and maybe economists know something about how to do it right. . . . There is a great legitimacy given to the idea that government is going to be doing these things and we in the law schools should try to help the government do it right. (Kitch 1983a, pp. 175–76)

Indeed, this was the very perspective motivating much of the work of one of the other major founders of Chicago law and economics in this early period, Ronald H. Coase.

Coase was educated at the London School of Economics and served on the faculty there from 1935–51, when he emigrated to the United States. He held appointments at the Universities of Buffalo and Virginia before moving on to the University of Chicago Law School in 1964. Although Coase is most closely associated with the Chicago school, his seminal article, "The Problem of Social Cost," which is the cornerstone of Chicago law and economics literature (see pp. 107–19), was actually written several years before he arrived at Chicago. One of Coase's major reasons for going to the law school was to assume the editorship (with Director) of the *Journal of Law and Economics*, and he describes his motivation in going to Chicago as follows: "I don't think that I would ever have come to the University of Chicago had it not been for the existence of the *Journal of Law and Economics*. That's what I wanted to do. I wanted to get what Aaron had started going for the whole profession—and when I say the profession, I mean the economics profession; I have no interest in lawyers or legal education" (Coase, quoted in Kitch

1983a, p. 192). That someone so avowedly indifferent to legal education became such a major figure in the evolution of the economic analysis of law is itself an interesting episode in the history of ideas (Medema 1994, 1998)—but then, Coase could not have anticipated that the influence of his insights in "The Problem of Social Cost" would be at least as great on legal scholarship as on economics.

The Economics Department

Although the events in the law school laid the foundation for the development of law and economics at Chicago, a full understanding of what transpired also necessitates an appreciation of the scholarship generated by the faculty of the Department of Economics, both before World War II and thereafter. Especially with respect to economic method, there were really two Chicago schools of thought, roughly divided in time by the war (Duxbury 1995, p. 367; Reder 1982; Emmett, forthcoming).[12] The perspective of the prewar Chicago school is evidenced in the scholarship of Frank Knight, Jacob Viner, Paul Douglas, and Henry Schultz, who were themselves by no means of a homogeneous perspective. In simple terms, proponents of the early Chicago approach generally accepted the propositions that had been at the heart of economics since the publication of Adam Smith's *Wealth of Nations* (1776)—within a liberal democracy, the rational pursuit of economic self-interest by economic actors was taken as given, competition was seen as inherent within and intrinsic to economic life, and market-generated outcomes were said to be generally superior to those resulting from government interference with the market mechanism. Although these propositions (the latter two in particular) were being increasingly called into question within the profession at large during the 1930s, their continuity within the Chicago tradition served (and continues to serve) to set the Chicago perspective apart from much of the rest of the economics profession.

From this group, it was Knight who had the most impact on what has come to be known as Chicago law and economics. While his writings were a significant force, his greatest influence came through the perspective that he imparted to his students—most importantly, for present purposes, Milton Friedman, George Stigler, and Director. Knight's interest and strength did not lie in the use of formal mathematical and quantitative tools, but rather in the economic way of thinking and in applying this way of thinking to the development of

---

[12] For wide-ranging discussions of the Chicago school of economics, see Bronfenbrenner (1962), Coats (1963), Miller (1962), Samuels (1976), Reder (1982) and Emmett (forthcoming). It should be noted that, even today, the Chicago school is not nearly as homogeneous as some would make it out to be.

economic ideas and to the tearing down of what he saw as false theories that were becoming increasingly fashionable within economics (Reder 1982, pp. 4, 6).

The next generation of Chicago economists kept one foot in this camp but also moved beyond it. As Duxbury (1995, p. 368) has pointed out, whereas the earlier generation of Chicago economists "had grasped and applied certain of the basic insights of Adam Smith . . . [p]ost-war Chicagoans were more intent on elaborating and extending these insights." In line with this, the new generation undertook to demonstrate, in formal terms, the detailed nexus between competitive markets and efficient outcomes. The nature of these price-theoretic undertakings was necessarily abstract and typically ahistorical, largely relying on positive, empirical research and mathematical analysis, all very much in keeping with broader movements within neoclassical economics at the time, and toward which Knight, ironically, had been rather hostile.[13] In this sense, then, as Reder (1982, p. 6) has noted, "Knight contributed to the formation of their minds but did not influence the direction of their research." Following the lead of Friedman and Stigler, postwar Chicago economists, buttressed by their empirical research, emphasized the efficacy of the competitive market system, arguing for less government intervention, fewer wealth redistribution policies, reliance on voluntary exchange with a concomitant reliance on the common law for mediating conflicts, and an across-the-board promotion of more private enterprise—which, based on the evidence provided by their empirical research, would facilitate a more efficient allocation of resources.

Exploding the Boundaries: Becker and Posner

There is no doubt that the economic analysis of law—Chicago law and economics—has been the most successful of economists' imperialistic forays into other disciplines, though their move into political science in the form of public choice theory, too, has had major influence on the "invaded" field (see chapter 3 of this book). Coase's "The Problem of Social Cost" (1960) and Calabresi's "Some Thoughts on Risk Distribution and the Law of Torts"

---

[13] James M. Buchanan, a Chicago-trained economist and also a Nobel Prize winner, has not been reluctant to castigate the methodological thrust of "economic science" as practiced in the 1980s and exemplified by many of the positivists working within the Chicago tradition. His reference to "economic science" is purely pejorative, and he argues that the modern economists' focus on tools and mathematical prowess has rendered them illiterate as to fundamental ideas and principles of economics and, further, has turned them into "ideological eunuchs." Most academic programs in economics, he says, are "now controlled by rent-recipients who simply try to ape the mainstream work of their peers in the discipline . . . and they seem to get their kicks from the discovery of proofs of propositions relevant only for their own fantasy lands" (Buchanan 1986, pp. 14–15).

(1961) raised many issues for both economists and lawyers; in particular, they revealed both the economic nature of many of the questions of legal analysis—that legal rules and decisions across many traditional fields of law beget both benefits and costs, and thus are amenable to analysis in efficiency terms—and the potential for the application of economic analysis to the law.[14] The intellectual direction and concerns of both the law school and the economics department transformed the academic landscape at Chicago, as the institutionalization of Chicago law and economics and the growing imperialism of the science of economics moved forward. This idea began to stimulate economic analyses of legal questions in the areas of property, contract, and tort law. Then, in 1968, came Becker's pathbreaking economic analysis of criminal behavior and criminal law. By the early 1970s the new law and economics had emerged as a recognized field of inquiry, best marked, perhaps, by the publication of Posner's *Economic Analysis of Law* (1972).

Becker's contribution to the Chicago school during this era is of particular import for the development of the economic analysis of law. Seizing upon the idea that economics is the science of choice, Becker employed the Chicago price-theoretic framework to explain choices across a broad range of nonmarket behavior, including racial discrimination in labor markets (Becker 1957); criminal behavior and law enforcement (Becker 1968); the organization of the family (including marriage and divorce), the decision to have children, and the division of labor within the household (Becker 1976); altruism (Becker and Barro 1988; Becker and Murphy 1988a); and addictive behavior, including drug use (Becker and Murphy 1988b; Becker, Grossman, and Murphy 1991). These works illustrate Becker's distinct approach to the economic analysis of law and typifies, perhaps better than any other scholarship, what has come to be known as the "economics imperialism."[15]

Posner received his LL.B. from Harvard Law School in 1962, and the period following his graduation was spent in Washington, first clerking for Supreme Court Justice William J. Brennan, Jr., and then working in the Kennedy and Johnson administrations. Posner was appointed associate professor of law at Stanford in 1968, and it was there that he came into contact with Aaron Director, who exposed him to the economic approach to analyzing legal rules. Posner moved on to the University of Chicago Law School in 1969. Since 1981, he has served as a judge on the U.S. Court of Appeals for the Seventh Circuit, including as Chief Judge from 1993 until 2000. During his tenure on the court, Posner has continued both to teach regularly at Chicago and to publish at a prolific rate.

---

[14] See Grembi (2003) on the distinction between the respective contributions of Coase and Calabresi; see also Posner (2005) and Hylton (2005) for discussions of Calabresi's role in the development of the economic analysis of law.

[15] On economic imperialism, see Brenner (1980), Posner (1993a), and Lazear (2000).

It would surely not be an overstatement to rank Posner among the foremost legal scholars of the second half of the twentieth century. For if, as both its advocates and critics acknowledge, the law and economics movement ranks as the most significant development in jurisprudential analysis during this period, Posner, as the leading presence in this movement in scholarship and on the bench, deserves much of the credit.[16] His *Economic Analysis of Law*, now in its sixth edition, served both to develop the field well beyond the classical applications to property, contract, tort, and criminal law, and to present the subject matter in a way that facilitated its integration into the law school curriculum. He has taken on issues as diverse as sex, aging and euthanasia, the AIDS epidemic, law and literature, sexually transmitted diseases, the Clinton impeachment proceedings, the Supreme Court's role in the 2000 Bush-Gore election, cloning, homosexuality, surrogate parenting, religious freedom, and adoptions, as well as a vast array of more traditional topics spanning virtually every area of law and jurisprudence—all these analyses being infused, in some cases to a greater extent and in other cases to a lesser extent, with an underlying economic flavor. If Becker opened the floodgates to an economic analysis that touches on all areas of life, it was Posner who took this approach and ran with it to the far corners of the legal arena.

FUNDAMENTAL BUILDING BLOCKS OF THE CHICAGO APPROACH

The defining characteristic of the Chicago approach is the straightforward application of microeconomic (or price-theoretic) analysis to the law.[17] As such, this approach is grounded in the following premises: (i) individuals are rational maximizers of their satisfactions in their nonmarket as well as their market behavior; (ii) individuals respond to price incentives in nonmarket as well as market behavior; and (iii) legal rules and legal outcomes can be assessed on the basis of their efficiency properties. Along with (iii) comes a two pronged normative prescription; first, that legal decision-making should promote efficiency, and, second, that in formulating public policy, decision-makers should rely heavily on markets. Each of these premises will be discussed in turn.

The Rational Maximization of Satisfaction

The assumption that economic agents are rational maximizers—that is, they make purposeful choices so as to pursue consistent ends using efficient means[18]—stands as a cornerstone of modern economic theory. Under this view,

---

[16] For a further discussion of Posner's contributions see Medema (2007b).

[17] See Ulen (1989).

[18] Note that this definition of rationality does not imply the necessity of conscious deliberation.

individuals are assumed to have a set of preferences that are complete, reflexive, transitive, and continuous. Given these conditions, it can be shown that consumer preferences can be represented by an ordinal utility function,[19] based on which, assuming that individuals are perfectly able to process all relevant information about the alternatives available to them, they can rank all possible outcomes according to their relative desirability. Consumers will then choose to consume the bundle of goods/activities that maximizes utility. Similarly, firms are viewed as profit maximizers, with the result that the level of output produced, the price charged, the composition of inputs and of payments thereto, the contracting practices of the firm, and so on are those that will maximize the firm's profits.

The rational maximization assumption has straightforward implications for the individual choice process. As was made clear in appendix A to chapter 1, individuals will engage in additional units of an activity as long as the additional benefit derived from another unit of that activity is greater than or equal to the additional cost—that is, as long as marginal benefit is greater than or equal to marginal cost—whether that activity is the consumption of goods, production, or the supply of labor. Thus, for example, the decision as to how many apples to consume comes down to an evaluation of whether the additional benefit from each additional apple consumed is greater than or equal to the cost, in terms of the best foregone alternative to consuming that apple. In the present context, decisions regarding unlawful activity (e.g., to breach a contract, to take precaution against breach, to engage in potentially tortious conduct, or to engage in criminal behavior) become a matter of comparing marginal benefits with marginal costs. From this perspective, those who break the law are not essentially different from the rest of the population; they simply have different preferences, opportunity costs, and constraints and engage in "illegal" activities because these are the activities that maximize their net benefit. Some commentators have challenged this idea on the grounds that individuals do not have full and complete information—but information, too, is a good, additional units of which are consumed only as long as marginal benefit is greater than or equal to marginal cost. Thus, as the price of information rises relative to the associated marginal benefit, so too does the optimal level of ignorance.[20] This, in fact, is part of the economic rationale for the failure to specify fully all contingencies in a contract, or the failure to fully abate pollution or eliminate all crime.

The "rational" individual of economics contrasts, of course, with the "reasonable" individual of traditional legal theory—an individual who is socialized into the norms and conventions of a community, and whose behavior corresponds to these norms. The law is said to reflect these norms and conventions, and thus is obeyed by reasonable individuals. Those who engage in

[19] See Kreps (1990, ch. 2) and Varian (1992).
[20] See Stigler (1961).

"illegal" activities are seen as unreasonable in that they have violated these norms and conventions. In contrast, the economic approach says that behavior can be (and usually is) rational, even when it conflicts with these social norms (Cooter and Ulen 1988, pp. 11–12).[21]

Legal Rules as Prices

The idea that individuals are rational maximizers implies that they respond to price incentives: consumers will consume less of a good as its price rises and producers will produce more of a good as its price rises (all of this, ceteris paribus).[22] Within the legal arena, legal rules establish prices, such as fines, community service, and incarceration, for engaging in various types of illegal behavior. The rational maximizer, then, will compare the benefits of each additional unit of illegal activity with the costs, where the costs are weighted by the probability of detection and conviction.

The adjustment of the level of illegal activity, be it tortious acts, breach of contract, or criminal behavior, thus becomes a matter of adjusting the prices reflected in the legal rules. To reduce the amount of such activities, one simply raises their price through the imposition of higher fines or greater jail time by an amount sufficient to induce the desired degree of behavioral change. As Posner (1983, p. 75) has said, "The basic function of law in an economic or wealth maximizing perspective is to alter incentives." This is the core of what was termed "legal centralism" in the discussion of the "logic of law and economics" in chapter 1. An increase in the price of engaging in an illegal activity will induce individuals to reduce or even eliminate their involvement in such activity, and what illegal activity remains will be that for which the marginal benefits to these individuals continue to exceed even the higher marginal cost.

The predictions derived from this are very straightforward. The imposition of liability on polluting firms will raise the price of pollution and induce the firm to reduce the level of pollution as long as the marginal benefits (e.g., foregone damage payments) from doing so exceed the marginal cost of pollution abatement. Similarly, the institution of higher damage payments for negligence in tort (e.g., punitive damages, pain and suffering, etc.) will induce potential tortfeasors to take additional precaution to prevent the occurrence of a tort. In addition, higher fines and longer jail terms will reduce the amount of crime.

---

[21] The influence of social norms will be explored in chapter 7.

[22] As Becker (1976, pp. 151–58) has shown, however, we may expect an increase in price to lead to a reduction in quantity demanded even when people are not rational.

Efficiency

The third defining characteristic of the Chicago approach to law and economics is that legal decision-making and the evaluation of legal rules should be grounded in economic efficiency.[23] As Posner (1990, p. 382) has put it, "The economic task from the perspective of wealth maximization is to influence [individuals] so as to maximize [their] output." One criterion employed is Pareto efficiency—that a course of action is efficiency-enhancing if at least one person can be made better off without making anyone else worse off. The Pareto criterion is generally recognized to be quite limited as a guide to legal decision-making because of the ubiquity of losses due to legal change. The impossibility and/or prohibitive cost of compensating all of these losses make it virtually impossible to conceive of alterations in legal rules that *would* satisfy the Pareto criterion. As such, the use of the Pareto criterion would forever perpetuate the status quo. The standard definition of efficiency employed in Chicago law and economics is Kaldor-Hicks efficiency, or wealth maximization: A legal change is efficiency-enhancing if the gains to the winners exceed the losses to the losers[24] or, alternatively stated, if the wealth of society (as measured by willingness to pay) is increased.

An example will easily demonstrate this efficiency concept in action. Suppose that a firm dumps chemicals into a stream, the effect of which is to reduce the property values of downstream landowners by a total of $1 million. Suppose further that the downstream landowners file suit seeking a permanent injunction against this dumping. If the downstream landowners have no way of preventing the damage, but the firm could eliminate the damage by installing a filtering device at a cost of $600,000, efficiency would dictate that the firm install the filter, since the cost of abatement is less than the damage from the chemical discharge. Thus, the injunction should be granted, which in turn would induce the firm to install the filter, increasing societal wealth by $400,000. Suppose, however, that due to the nature of the damage and the technology available, the downstream landowners could, in fact, eliminate the damage themselves for a cost of $300,000. In this case, efficiency would dictate a denial of the injunction, as the landowners could eliminate the harm at a lower cost than could the firm, and would in fact do so if the injunction were denied, since the gain from doing so ($1 million) exceeds their cost ($300,000).

The Chicago school's reliance on efficiency is in part due to the belief that it is a legitimate and important goal for legal-economic policy, but it also reflects their concern as to what can "come in" under the name of "justice" or "fair-

---

[23] See appendices to chapter 1.

[24] Thus the winners could, hypothetically, compensate the losers for their losses and still be better off, creating a hypothetical (or potential) Pareto improvement.

ness," neither of which has a unique definition that commands universal acceptance. As a consequence, Chicagoans are skeptical of the motives of those who invoke its use, and it is not unusual to read characterizations of the concepts of justice or fairness such as "A suitcase full of bottled ethics from which one freely chooses to blend his own type of justice" (Stigler 1972, p. 4), or, "Fairness is a vagrant claim applied to any value that one happens to favor" (Areeda 1978, p. 21). The purported objectivity of the efficiency criterion makes it the preferred alternative to the ambiguities inherent in using justice or fairness, in the eyes of its proponents. In like manner, markets—readily capable of generating efficient outcomes—are viewed as the preferred system of social control.

The view of efficiency as justice is what many of the critics of the Chicago approach to law and economics find so troubling.[25] Posner (1992, p. 27), however, has argued both that efficiency is "perhaps the most common" meaning of justice and that "a moral system founded on economic principles is congruent with, and can give structure to, our everyday moral intuitions" (Posner 1983, p. 84).[26] At one point, Posner (1983, p. 89) attempted to ground the ethical basis for a rule of wealth maximization in the principle of consent, which he describes as "an ethical criterion congenial to the Kantian emphasis on treating people as ends rather than as means, in a word, on autonomy." The notion of consent employed by Posner here is based on ex ante compensation. The connection between consent and ex ante compensation lies in the idea that individuals would consent to wealth maximization as a criterion for establishing common law rules of adjudication as long as there is a sufficient probability that they will benefit (i.e., be net winners) from the application of such rules in the long run, even though they may be losers from the application of a particular rule. It thus is not necessary to compensate those who lose from the application of a particular wealth-maximizing decision rule, because these individuals have garnered ex ante compensation in the form of the greater wealth (lower costs) that accompanies the adoption of these wealth-maximizing rules.[27]

Posner has also suggested that wealth maximization, which he sees as blending certain elements of the utilitarian tradition and Kantian tradition (the latter with its emphasis on human respect and autonomy) is superior as an ethical concept to both utilitarianism and Kantianism. First, he says, "the pursuit of wealth, based as it is on the model of the voluntary market transaction, in-

---

[25] See, for example, the *Hofstra Law Review*, "Symposium on Efficiency as a Legal Concern" (1980), especially Coleman (1980); see also Dworkin (1980), Kronman (1980), and Michelman (1978).

[26] Posner does, however, rule certain wealth-maximizing ideas out of bounds; see Posner (1983, chs. 3 and 4) and Posner (1990, pp. 374–87).

[27] For an elaboration of these ideas, see Posner (1983, pp. 88–115). See also Mercuro and Ryan (1984, p. 125).

volves greater respect for individual choice than in classical utilitarianism" (Posner 1983, p. 66). Second, economic liberty "can be grounded more firmly in wealth maximization than in utilitarianism" (p. 67). Third, "the wealth-maximization principle encourages and rewards the traditional 'Calvinist' or 'Protestant' virtues and capacities associated with economic progress" (p. 68). Finally, wealth maximization "provides a firmer foundation for a theory of distributive and corrective justice," along with a firmer commitment to the principle of rights than is evident in utilitarian and Kantian thinking (p. 69).

Posner has since overtly moved away from the attempt to ground wealth maximization in moral philosophy. The strongest argument for wealth maximization, he says now, is pragmatic: "We look around the world and see that in general people who live in societies in which markets are allowed to function more or less freely not only are wealthier than people in other societies but have more political rights, more liberty and dignity, are more content . . . so that wealth maximization may be the most direct route to a variety of moral ends" (Posner 1990, p. 382). Posner's perspective here is consistent with the larger Chicago school aversion to concepts, such as justice and fairness, that are grounded in moral philosophy.[28] That is, Posner's position here is not the result of any sort of weakened commitment to efficiency but rather to any defense of efficiency based in moral philosophy.

The three features of the Chicago approach to law and economics that we have just laid out here—rational maximization, legal rules as prices, and the efficiency criterion—are all reflected in what some would call *the* cornerstone of the economic analysis of law: the Coase theorem and its corollary, which we shall call "the Coase lesson."

The Coase Theorem

We noted earlier in this chapter that the Chicago approach to the analysis of common law rules received its impetus from Coase's contribution in "The Problem of Social Cost" (1960). The traditional legal and economic approaches to harmful effects are similar in that each seeks to label one party (A) the cause of the harm and the other (B) the victim. From a legal perspective, then, the question is whether A should be liable for the harm it causes to B, with various common law rules available to resolve the liability issue under different conditions. Coase took the discussion in a different direction by pointing out that, while it is traditional to think of externality issues in terms of these conventional notions of causation—for example, the polluting factory *causes* damage to the surrounding neighborhood—the problem, in reality, is reciprocal in nature. That is, while it is true that there would be no harm to B

---

[28] See also Posner (2003) and Medema (2007b).

(the members of the neighborhood) absent A's pollution, it is equally the case that there would be no harm to B if B were not located in the vicinity of A. Looking at matters from a cost perspective, A's pollution imposes costs on B, but for B to be free from pollution (through A's abatement) imposes costs on A (the costs associated with abatement). As Coase points out, the real legal *and* economic question is, "[S]hould A be allowed to harm B or should B be allowed to harm A?" "The problem," he says, "is to avoid the more serious harm" (Coase 1960, p. 2). This approach contrasts starkly with the unidirectional notions of causation and the attendant decision rule implication, as reflected in, for example, the past doctrines of "first-in-time" or "coming to the nuisance" as heretofore enunciated in traditional tort and property law.[29] It is also in marked contract to the Pigovian approach to externalities in economics, where taxes, subsidies, or regulations are used to alter the behavior (i.e., abate the activities) of the party "causing" the externality.[30]

Coase then went on to suggest that, beyond being reciprocal in nature, externalities are also amenable to efficient resolution without government tax or regulatory measures being imposed—at least if markets operate without frictions. This requires that three conditions be met. First, rights over the resources in question must be fully specified. That is, some party must have legal control over those resources; if there are no rights over said resources or those rights are incompletely defined or unassigned, market-oriented solutions are all but precluded. Second, legal rights must be alienable: if rights cannot be exchanged, the processes envisioned here are rendered inoperative. The final assumption is that transaction costs, including the costs of acquiring information, are zero.[31] When these conditions are met—that is, when the exchange or market process is frictionless—all that is necessary is that the government decide liability one way or the other, in effect granting one party the right to act or the other party the right to be free from the action of the first party. Once these rights are defined and assigned, the parties are then free to trade the rights, and will do so if it is in their self-interest—ultimately, to an efficient solution. A simple example will illustrate this idea at work.

Suppose that a rancher (R) and a farmer (F) occupy adjoining parcels of property. If the rancher puts more than ten head of cattle on his property, the cattle tend to wander onto the farmer's property and do damage to the crops. Suppose further that the rancher's profits and the value of damages to the farmer's crops associated with various quantities of cattle, are as given in

---

[29] For example, Epstein (1973) challenges the notion of reciprocity.

[30] The term "Pigovian" derives from the influence of Cambridge economist A. C. Pigou (1920) in developing the neoclassical approach to the analysis of externalities—an approach that was elaborated and extended by a host of others in the middle third of the twentieth century.

[31] These will later be explored more fully. The implication of this is that all agents have full and complete information regarding benefits, costs, etc.

TABLE 2-1

Q_R (head of cattle)	Π_R (total profit to rancher)	$M\Pi_R$ (marginal profit to rancher)	TD_F (total damage to farmer)	MD_F (marginal damage to farmer)	NSB (net social benefit)
10	100	—	0	—	100
11	106	6_{11}	1	1_{11}	105
12	111	5_{12}	3	2_{12}	108
13	115	4_{13}	6	3_{13}	109
14	118	3_{14}	10	4_{14}	108
15	120	2_{15}	15	5_{15}	105
16	121	1_{16}	21	6_{16}	100
17	120	-1_{17}	28	7_{17}	92

table 2-1. If the government (legislature or the court) determines that the ranchers are *not liable* for the damages caused by their wandering cattle, this rancher will raise sixteen head of cattle, as that is the quantity that gives him the highest profit, $121. In deciding how much to produce, the rancher has no incentive to take into account the damages that his cattle cause to the farmer. Note that net social benefit (of $109) is maximized at $Q_R = 13$.

Observe that the sixteenth head adds only $1 to the rancher's profit while simultaneously causing $6 in damage to the farmer's crops. The farmer will be willing to offer (bribe) the rancher up to $6 to have fifteen cows rather than sixteen.[32] Of course, the rancher will be willing to reduce herd size to fifteen for any payment in excess of $1. Any payment from the farmer to the rancher between $6 and $1 will make both parties better off. Additional bribes by the farmer to the rancher will continue to make both parties better off. For example, for the fifteenth cow, the additional profit to the rancher is $2 and the marginal damage is $5; and for the fourteenth cow, the marginal damage to the farmer still exceeds marginal profit to the rancher. A reduction in herd size from thirteen to twelve, however, reduces damage to the farmer by $3 while reducing the rancher's profits by $4. Given this, the farmer will not be willing to offer the rancher an amount of money sufficient to get him to reduce his herd size to twelve. Thus, when the rancher is not liable, the farmer will be willing to offer the rancher an amount sufficient to induce him to limit his herd size to thirteen. This is exactly the herd size where net social benefit is at its highest level, $109.

Now suppose instead that the government decides to make the ranchers liable for the damages caused by their wandering cattle—that is, our farmer has the right to be free from harm. This means that the rancher cannot expand his

---

[32] The use of the would "bribe" in discussing these payment is entirely nonpejorative.

herd size beyond ten cows unless the farmer agrees to allow him to do so. Obviously, $Q_R = 10$ is below the efficient herd size, $Q_R = 13$. Notice however, that the addition of the eleventh cow increases the rancher's profit by \$6 while causing only \$1 in damage to the farmer. The rancher will thus be willing to pay up to \$6 to increase herd size to 11, while the farmer will accept any payment in excess of \$1 to agree to this. Any payment from the rancher to the farmer between \$6 and \$1 will make both parties better off. Likewise, the twelfth and thirteenth cows increase the rancher's profits by more than they increase damage to the farmer's crops, so the rancher will be able to bribe the farmer to allow him to increase herd size to thirteen. The fourteenth cow, however, adds only \$3 to the rancher's profits while increasing damage to the farmer by \$4; hence, the rancher will not be willing to offer the farmer an amount sufficient to convince him to allow herd size to expand to fourteen. This time, when the rancher is liable, we see that the herd size is thirteen, just as it is when the rancher is *not* liable.

Thus, whether the government, through its definition and assignment of rights (i.e., liability), decides that the rancher has the right to allow his cattle to roam freely or the farmer has the right to be free from damage caused by roaming cattle, we end up, through the bargaining process, with the same solution—a herd size of thirteen, \$6 worth of crops destroyed, and the highest net social benefit attainable = \$109. This result has come to be known as the Coase theorem. The theorem has been stated in numerous ways over the years, but a statement that closely follows the spirit of Coase's own discussion is,

> *If rights are fully specified and transaction costs are zero, parties to a dispute will bargain to an efficient and invariant outcome regardless of the initial specification of rights.*[33]

The Coase theorem guarantees not only the attainment of the efficient outcome but also that the efficient outcome will be reached in the most efficient fashion. Suppose, for example, that an alternative to reducing herd size is to install a fence, and that a fence that eliminates damage can be installed by the farmer at a cost of \$6, or by the rancher at a cost of \$8. Of course, if a fence is to be installed, both the rancher and farmer would prefer that the other party install and pay for it. If the rancher is not liable for damage caused by his

---

[33] See Medema and Zerbe (2000) for a litany of statements of the Coase theorem, as well as a fairly comprehensive survey of the vast literature on the theorem. Other surveys include Zerbe (1980), Cooter (1982a), Medema (1994, chapter 4; 1995), and Farber (1997). One of the most controversial aspects of the Coase theorem is what is known as the invariance proposition—that the *same* efficient outcome will be reached regardless of the initial assignment of rights. It is widely accepted that the presence of income effects is sufficient to invalidate the invariance claim; that is, *an* efficient outcome will be reached regardless of the assignment of rights, but these efficient outcomes will not give *identical* allocations of resources. This issue is dealt with in detail in the surveys cited in this note.

cattle, then the farmer would choose to install the fence for $6, resulting in zero damage to his crops. He would prefer to incur the $6 cost as it is cheaper than either bribing the rancher to reduce herd size down to zero, which would have cost him at least $21 ($\Sigma = \$6 + 5 + 4 + 3 + 2 + 1$), or bribing the rancher at least $8 to install the fence. Once the fence is in place, the rancher will then maintain a herd size of sixteen.

If instead the rancher is liable and the farmer has the right to be free from damage, then the rancher will bribe the farmer to install the fence, because the $6 cost is cheaper than either installing the fence himself for $8 or the bribe that he would need to pay the farmer in order to expand his output. Indeed, for the same $6 the rancher would be able to bribe the farmer to allow a herd size only equal to thirteen ($\Sigma = \$1 + 2 + 3$). With the availability of this cheaper alternative, once the rancher bribes the farmer $6 to install the fence, the rancher will once again be able to maintain a herd size of sixteen. Thus, no matter to whom the right is initially assigned, the farmer's land gets fenced and the rancher has a herd size of sixteen. Recalling that damage associated with any herd size is zero when the fence is installed, the net benefit to society of any given herd size is rancher profit minus the $6 fencing cost. As such, net social benefits are maximized where profits are maximized, so efficient herd size here is sixteen—the exact result reached by the parties themselves, regardless of the initial assignment of rights.

The Coase theorem can also be depicted graphically, as reflected in figure 2-1.[34] Here, $M\pi_R$ shows the marginal profit to the rancher of additional head of cattle, while MD_F shows the associated marginal damage to the farmer. Assume that the three conditions outlined earlier—all rights fully specified, alienable rights, and the zero transaction costs—are satisfied. If the government determines that the rancher is *not liable* for the damages caused by his wandering cattle, the rancher will produce a level of output Q_3, where his marginal benefits are exhausted and his profits maximized. Because $MD_F > M\pi_R$ at Q_3, there exists scope for a mutually beneficial bargain between the rancher and the farmer such that the rancher will agree to reduce the size of his herd. That is, the farmer is willing to pay the rancher an amount up to MD_F to reduce the herd size, and the rancher is willing to accept any payment greater than $M\pi_R$ to do so. Given that $MD_F > M\pi_R$ between Q_3 and Q^*, the farmer will bribe the rancher to reduce his herd size to Q^*. That is, because it is in the interest of both parties to bargain, and because all parties have full and perfect information and face zero costs of transacting, they will bargain until the gains from exchange are exhausted—here, at Q^*.

If the government instead makes the rancher liable and the farmer is entitled to be free from the harm caused by the rancher's activity, the rancher will be

---

[34] The figures in this section are designed for a more general discussion and are not intended to reflect the numbers in the previous table.
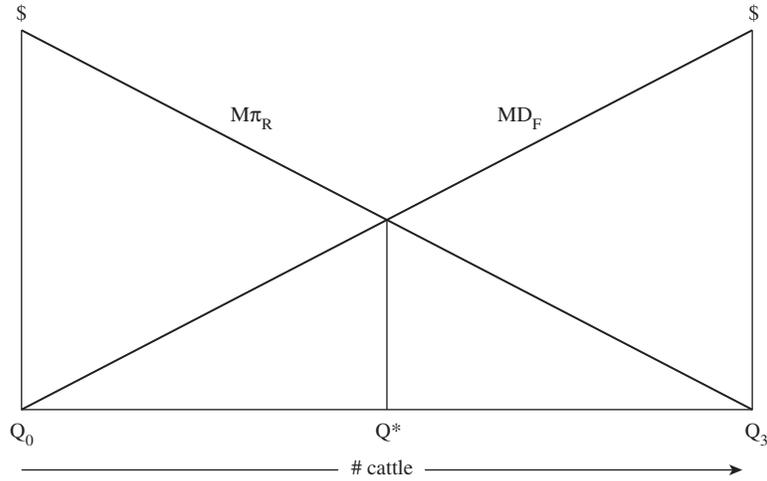
FIGURE 2-1. Cattle Rancher and Farmer

forced to maintain Q_0 cattle unless he can induce the farmer to agree to let him undertake some non-zero amount if this activity. Since $M\pi_R > MD_F$ at Q_0, there is scope for a mutually beneficial bargain whereby the rancher compensates the farmer for damage done. Furthermore, because $M\pi_R > MD_F$ between Q_0 and Q^*, the rancher will bribe the farmer to allow him to increase his herd size to Q^*, at which point the gains from trade are exhausted. Thus we see, diagrammatically, the thrust of the Coase theorem: it does not matter to which party the rights are initially assigned—that is, whether the rancher is or is not liable for the damage caused by his straying cattle. As long as the government defines and assigns the right, one way or the other, the parties will bargain to an efficient and invariant level of output, Q^*.

   The Coase theorem is one of the most debated, most powerful, and most misunderstood concepts in the entire legal-economic literature. Literally hundreds of articles have been written on the theorem, attempting to prove, disprove, test (experimentally or empirically), praise, or denigrate it.[35] To understand the theorem and its implications properly, there are three simple facts to keep in mind.

   First, the Coase theorem is correct. Period. *All* supposed disproofs and refutations of the theorem involve violations of one or more of the theorem's underlying assumptions—usually that of zero transaction costs.[36] Second, the world that we live in bears little correspondence to the world assumed by the

[35] Again, much of this literature is discussed in Medema and Zerbe (2000).
[36] See as well as Medema and Zerbe (2000).

Coase theorem.[37] The world of zero transaction costs is one in which there are no frictions, no imperfections—where, in particular, all parties have full and complete information. It is a world in which, as Coase (1988, p. 15) has remarked, "[E]ternity can be experienced in a split second." Of course, Coase concludes from this that, among other things, analyzing a world of zero transaction costs is akin to "divining the future by the minute inspection of the entrails of a goose" (1981, p. 187). The Coase theorem simply tells us that, under certain conditions, the form of legal rules does not affect the allocation of resources. It is a starting point for analysis—a useful fiction, if you will. A look at the world around us tells us not only how unrealistic those conditions and assumptions really are, but also that much insight into economics and the law can be gained by looking at the consequences of relaxing those assumptions—the thrust of the "Coase lesson," which we explore in the next section. Third, the Coase theorem undergirds the strong philosophical argument that can be made for efficiency as the criterion for legal rule-making. The logic here is both simple and persuasive, and it is related to Posner's consent-based argument presented earlier. The Coase theorem tells us that parties will bargain to the efficient outcome regardless of how rights are initially assigned if transaction costs do not preclude said bargaining. This, then, is the outcome that agents would choose voluntarily if they were able to do so. Given this, why should the judge not assist them in doing so by imposing that outcome—the efficient outcome—when transaction costs preclude bargaining to it? So, you see, the Coase theorem—a purely positive depiction of life in a world without transaction costs—provides the underpinnings for a normative law and economics based upon the efficiency criterion. Against those who would argue for the moral/ethical/philosophical bankruptcy of Chicago law and economics, the theorem, with its roots in voluntary agreements among affected parties, provides a strong rebuttal.[38]

The Coase Lesson

We have already emphasized that the Coase theorem's results hinge crucially on the assumptions that all rights are fully specified and alienable, and that the costs of transacting are zero. The assumption that transaction costs are zero is far more restrictive, and certainly more controversial, than the others. The most simple and obvious notion of transaction costs encompasses the costs of negotiating, monitoring, acquiring information, and enforcing contractual

---

[37] In fact, Coase pointed to the unrealistic nature of the zero transaction cost assumption already in "The Problem of Social Cost" (1960).

[38] There are certain important commonalities between this idea and the catallactic approach that undergirds public choice theory and constitutional economics (see ch. 3 and Buchanan [1983]), as well as with the analysis by Rawls in *A Theory of Justice* (1971).

agreements.[39] The most important subset of transaction costs, in terms of the implications for the allocation of resources, is information costs. Costly information and the resulting informational asymmetries give rise to the specter of strategic behavior, which, as many commentators have shown, negates the operation of the theorem's mechanisms.[40] In fact, however, all of these costs are significant. Thus, a more accurate definition is, "Transaction costs are the costs of establishing and maintaining property rights" (Allen 1998, p. 108).

A moment's thought makes clear how violation of the zero-transaction-costs assumption impedes bargaining to the efficient solution. First, sufficiently large transaction costs will preclude bargaining altogether. That is, if the costs associated with the bargaining process, gathering information, and so forth are greater than the expected gains from bargaining, no one has any incentive to bargain over rights once the court has rendered its decision on liability (or rights in general). Second, even if transaction costs are not so high so as to preclude bargaining, the marginal cost of additional negotiations may exceed the marginal benefit beyond *some* point, with the result that the final allocation of resources *will*, in fact, depend on to whom the rights are initially assigned. We can see the result in figure 2-2. The point to be emphasized here is that, given the presence of transaction costs, parties will bargain over rights so long as the expected gains from the bargain exceed the expected costs, which include the costs of transacting.

Continuing with our example of the rancher and the farmer, $M\pi_R$ and MD_F remain as defined above, and we will assume initially, for convenience, that marginal transaction costs (MTC) are constant and equal for both parties—that is, $MTC = MTC_R = MTC_F$—and that both parties will simultaneously incur these costs regardless of to whom the government assigns liability. The fact that both parties will incur transaction costs requires us to adjust $M\pi_R$ and MD_F by the amount of the transaction costs, which provides us with the two parties' respective bargaining curves, BC_R and BC_F.[41] The diagrammatic impact of the simultaneous imposition of transaction costs depends on whether or not the cattle rancher is liable for the damage caused by his straying cattle.

Suppose that the government were first to determine that the rancher was *not* liable for the damages caused by his wandering cattle. In the absence of some agreement between the rancher and the farmer, the rancher would maintain Q_3 cattle. We can see in figure 2-2 that when the rancher is not liable, the presence

[39] For useful treatments of the concept of transaction costs, see Dahlman (1979), Allen (1991, 1998), and Medema and Zerbe (2000).

[40] For example, the rancher may take advantage of informational asymmetries by attempting to overrepresent his net marginal benefit and the farmer by overstating his marginal damage—in both cases toward the end of inducing the other party to pay a larger bribe.

[41] In the discussion up to this point, marginal damage and marginal profits illustrate the parties' willingness to pay or the minimum payment that they are willing to accept to bring about a change in herd size. The presence of transaction costs affects willingness to pay and willingness to accept within the bargaining process, causing them—and thus the parties' bargaining curves—to differ from those in the zero-transaction-costs case by an amount equal to the size of the transaction costs incurred.
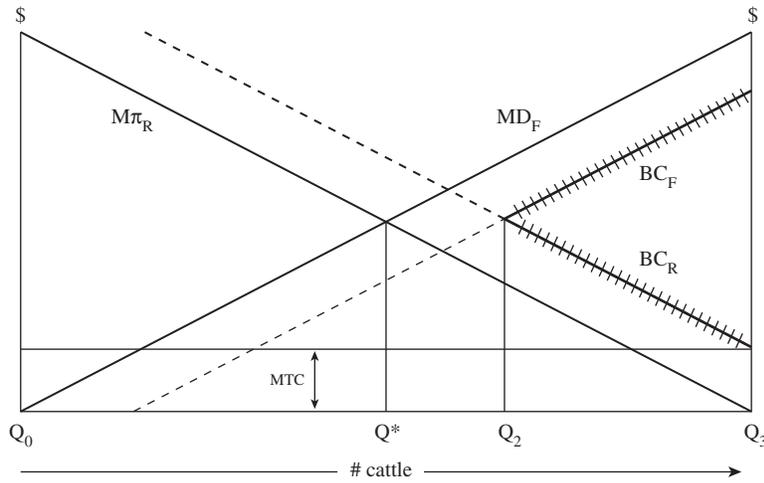
FIGURE 2-2. Rancher Not Liable

of transaction costs causes the rancher's bargaining curve, BC_R, to lie above the $M\pi_R$ and the farmer's bargaining curve, BC_F, to lie below MD_F. Thus, starting at Q_3, the farmer will have to initiate the bribe to get the rancher to reduce herd size, but the maximum bribe that the farmer is willing to pay will now be reduced owing to the presence of transaction costs the farmer must now incur. The farmer bargains down along the hatched segment of the BC_F curve. At the same time, the rancher will agree to reduce the herd size only if the bribe from the farmer (hatched segment of BC_F) covers both the reduction in his marginal profits and his costs of transacting. Thus, starting at Q_3, the rancher bargains up along the hatched segment of the BC_R curve, and the resulting bargaining solution is at herd size Q_2. That is, when transaction costs are positive, symmetric, and simultaneously incurred by both parties, the bargaining process will result in a herd size of $Q_2 > Q^*$ if the government determines that the rancher is *not* liable.[42]

Alternatively, when the rancher *is* liable, the rancher will have to initiate the bribe to induce the farmer to move away from Q_0. As we see in figure 2-3, when the rancher is liable the presence of the transaction costs causes the rancher's bargaining curve, BC_R, to lie below $M\pi_R$ and the farmer's bargaining curve, BC_F, to lie above MD_F. The maximum bribe that the rancher is willing to pay is reduced by the presence of the transaction costs that he must now incur. Thus, starting at Q_0, the rancher will bargain down along the hatched segment of the BC_R curve. At the same time, the farmer will agree to

[42] Of course, as previously described, if transaction costs are sufficiently high, no alteration of output will take place because the expected costs of bargaining will exceed the expected gains.
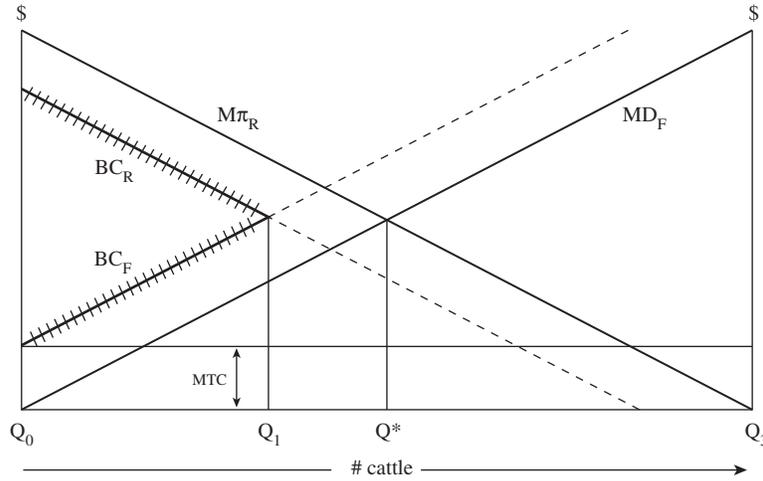
FIGURE 2-3. Rancher Liable

allow the rancher to increase the size of his herd only if the bribe from the rancher (hatched segment of BC_R) covers both the marginal damages of the wandering cattle to the farmer and the farmer's costs of transacting. Thus, the farmer bargains up along the hatched segment of the BC_F curve. The resulting equilibrium herd size is Q_1.[43] The results (Q_1, Q_2) that diverge from the optimum (Q^*) are efficient in the presence of transaction costs, since the costs of movement away from such points exceed the benefits.[44] The upshot is that, in the presence of any of the various types of costs associated with the transacting process, there is no unique optimal outcome. Instead, the outcome ultimately reached will be a function of the initial assignment of rights and the relative magnitude of transaction costs.[45] We call this the "Coase lesson":

[43] Note that if marginal transaction costs are increasing, the divergence between Q_1 and Q^* or between Q_2 and Q^* will be even greater.

[44] These positive transaction-costs-bargaining results are efficient in the Paretian sense—that is, all gains from trade have been exhausted, given the level of transaction costs, but the result is not efficient in the globally wealth-maximizing sense. See Buchanan (1983).

[45] Of course, income effects, endowment effects, and so on—implicitly assumed to be zero throughout this discussion—will have a similar influence on equilibrium outcomes, pushing them away from Q^* and, potentially, exacerbating the divergence already caused by the presence of transaction costs. On income effects, see, for example, Mishan (1967, 1971). Endowment effects exist when having or not having a right influences one's valuation of it—that is, causes a divergence between the amount that one is willing to pay to acquire a given right and the amount that one is willing to accept to give up that right. See, for example, Kahneman, Knetsch, and Thaler (1990). For a survey of the literature dealing with divergences between willingness to pay and willingness to accept, see Hoffman and Spitzer (1993). Many of these issues are discussed in Medema and Zerbe (2000), and they are also part of the subject matter of the new "behavioral law and economics."

The assignment of legal rights to one party or another impacts the allocation of resources and thus has clear efficiency implications. Moreover, depending on the assignment of rights (e.g., deciding that the cattle rancher is or is not liable for damages caused), the outcome may be in direct contradiction with the outcomes generated by policies reflecting uni-directional notions of causation prevalent in the past doctrines within traditional legal theory of property and tort law and the Pigovian approach to externalities in economics.

Now let's consider the economic impact of asymmetric transaction costs, which has further implications for the Coase lesson. Here we make two assumptions. First, we assume that the transaction costs incurred by the rancher are greater than those incurred by the farmer ($MTC_R > MTC_F$). Second, for expositional simplicity, we are assuming that once the government assigns the right to one party, say the farmer, then the other party, the rancher, would have to bear the entire burden of the rancher-specific transaction costs; hence, under this rights structure, the farmer would not incur any transaction costs.[46]

We begin by considering the case where the rancher is not liable for the damages caused by his straying cattle, as shown in figure 2-4. Thus, the farmer must incur the transaction costs, resulting in a bargaining curve BC_F, where BC_F is the farmer's marginal damage from additional head of cattle adjusted for transaction costs. In this case, the bargaining process will begin at Q_3. Through incremental bribes paid by the farmer (down along the hatched segment along BC_F) to the rancher in excess of the rancher's marginal benefits (up along the hatched segment along $M\pi_R$), the farmer will be able to bribe the rancher to get the rancher to reduce herd size. The bargaining will proceed to the point where $M\pi_R = BC_F$, yielding a herd size of Q_2. Here then we see that the existence of transaction costs precludes the bargaining process from yielding the wealth-maximizing solution Q^*.

Alternatively, if the farmer is granted the right to be free from trespass damage (i.e., the rancher is liable), then the rancher incurs all of the transaction costs. In figure 2-5, BC_R depicts the marginal benefits of additional units of herd-size expansion to the rancher adjusted for transaction costs. The rancher's willingness to pay is thus reduced by the amount of the transaction costs incurred when the farmer is granted the right by the government. This time bargaining will begin at Q_0. Through incremental bribes by the rancher (down along the hatched segment of BC_R) that are in excess of the farmer's marginal damages (up along the hatched segment of MD_F), the rancher will bribe the farmer to allow him to enhance the size of his herd. The bargaining will

---

[46] We could generalize the example by assuming that both parties simultaneously incur transaction costs, but that the transaction costs facing one party are larger than those facing the other. However, making what is quite clearly an unrealistic assumption here—that only one party at a time bears transaction costs—makes for a greatly simplified graphical treatment, as seen below.
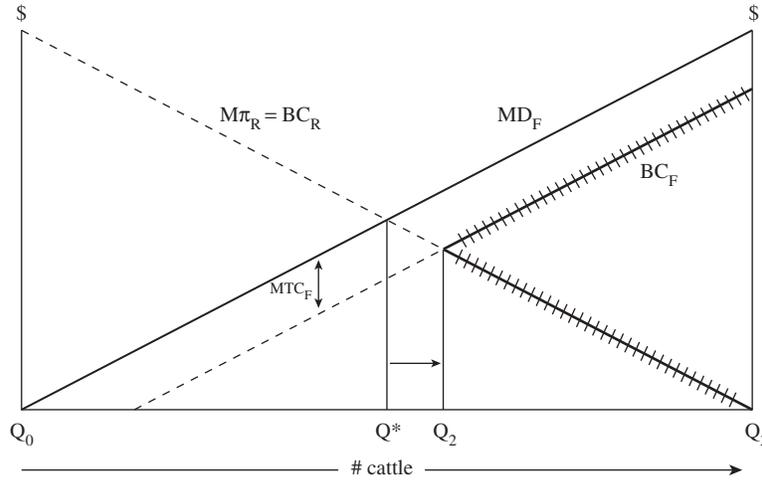
FIGURE 2-4. Rancher Not Liable

proceed to the point where $BC_R = MD_F$ yielding a herd size of Q_1. Once again it is clear that the existence of transaction costs does not allow the bargaining process to yield the wealth-maximizing, Q^*. Inasmuch as Q_2 is closer to the efficient solution Q^* than is Q_1, however, the normative implication for government policy is that the efficiency-enhancing rights structure associated with Q_2 is said to be preferred to the alternative rights structure associated with Q_1.

As noted earlier, the assumption that only one party actually incurs costs within the transacting process is unrealistic and made here only for expositional convenience. It is more likely that both parties will simultaneously incur costs, and probably not of the same magnitude. Moreover, it may well be that the sum of transaction costs across parties varies greatly with the assignment of rights. One party may be more difficult to locate or deal with, particularly if given the rights in question, and the transacting process may thus be smoother—that is, less costly—if the other party is assigned these rights instead. Ceteris paribus, if, in making legal decisions, we are going to use up some of society's scarce resources by incurring transaction costs, then the "Coase lesson" suggests that the courts should structure rights in a way that minimizes the amount of resources used up so that society can use the "saved" resources for other, more highly valued activities.

It follows that if efficiency is considered an important value in determining the law, in general, courts should attempt to assign rights in such a way as to minimize the sum of all costs associated with the externality, including transaction costs. It should be clear that this may very well generate remedies opposed to those that would be reached if courts subscribed to the more traditional uni-directional notions of causation prevalent in either the "first-in-time"
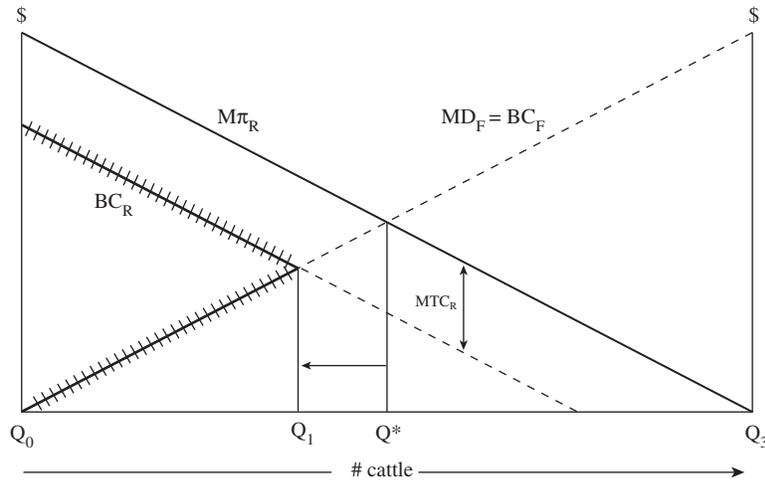
FIGURE 2-5. Rancher Liable

Figure labels: $\$$, $M\pi_R$, $MD_F = BC_F$, BC_R, MTC_R, Q_0, Q_1, Q^*, Q_3, # cattle

and "coming to the nuisance" rules of the traditional legal theory of property and tort law, or the Pigovian approach to externalities in economics—all of which would act to protect the victim-farmer and restrain the harm-causing rancher. Here, in contrast, the rancher is instead given the right.

The economic logic that underpins the Coase lesson is that a wide range of benefits and costs can, and usually do, attend alternative assignments of rights. That is, no matter what one's view of the normative import of efficiency, law does have important efficiency implications, and these include the costs associated with transacting over rights. The logic of this reasoning is inherent in Coase's approach where he states,

> The problem which we face in dealing with actions which have harmful effects is not simply one of restraining those responsible for them. What has to be decided is whether the gain from preventing the harm is greater than the loss which would be suffered elsewhere as a result of stopping the action which produces the harm. In a world in which there are costs of rearranging the rights established by the legal system, the courts, in cases relating to nuisance, are in effect, making a decision on the economic problem and determining how resources are to be employed. (Coase 1960, pp. 27–28)

Coase also argued that, as he put it, "the courts are conscious of this" and that judges "often make, although not always in a very explicit fashion, a comparison between what would be gained and what would be lost by preventing actions which have harmful effects" (Coase 1960, p. 28). This, in turn, stimulated legal-economic scholars to examine the extent to which legal outcomes of the past may reflect efficiency concerns.

The Efficiency of the Common Law

While many within law believe that the common law is a grouping of largely separate fields of analysis, each with its own set of judge-made legal rules and doctrines, the Chicago approach to law and economics suggests that the common law as a whole is unified by an underlying economic logic.[47] Perhaps somewhat ironically, the literature on the efficiency of the common law bears more than a passing resemblance to doctrinalism and, in certain respects, functions as an attempt by the Chicago school to return to formalism and to autonomous legal thought—here, organized around the doctrine of efficiency. Posner makes this very point:

> Many economic analysts of law, such as myself, are very interested in cases, and when we use economics to reconcile and to distinguish cases, we are carrying on the tradition of the doctrinal analysts. But insofar as we are trying to use modern economic concepts in this task, we fall into the category of positive analysts who use social science. . . . The positive analysts such as myself resemble traditional doctrinal analysts in believing that there really are rules of law—that the law is not wholly a matter of judicial discretion, as the more extreme Legal Realists believed. We use economics to inquire to what extent the common law is a coherent system of rules concerned with promoting efficiency. (Posner 1981, p. 1115, 1120)

As noted before, the idea that judges' decisions in common law cases seem to evidence an underlying economic logic was first raised within the Chicago tradition by Coase in "The Problem of Social Cost" (1960, p. 8–15, 19). The efficiency hypothesis has subsequently been probed in numerous studies by Posner and others who purport to describe the common law's economic logic.[48] As Posner (1987b, p. 5) defines it, the economic theory of the common law (broadly conceived to include all judge-made law) "is that the common law is best understood not merely as a pricing mechanism but as a pricing mechanism designed to bring about an efficient allocation of resources, in the Kaldor-Hicks sense."

There are three threads to the argument that the common law tends toward efficiency. The first is that the institutions of the common law have been designed to promote efficiency by fostering market transactions through contract. This approach uses the expected utility maximization model to provide insight into the issue of pretrial settlement versus formal adjudication and has been analyzed in the contexts of civil, criminal, labor, and antitrust cases. The

---

[47] See Cooter (1985).

[48] The first Chicago study that advanced the efficiency hypothesis in-depth was Posner (1972b). See also Rubin (1977), Priest (1977), and Goodman (1978). An overview of the efficiency theory of the common law can be found in Cooter and Ulen (2004, ch. 10) and Posner (1992, chs. 8, 19–21). Landes and Posner (1987) present an extensive analysis of the efficiency theory of the common law as applied to the area of torts.

issue facing each party is that of settlement versus adjudication, and each party's decision is assumed to turn on which course of action gives them the higher expected utility, given their resource constraints and their subjective probability estimates of prevailing should the case go to trial.

Let's look at the working of this process in the context of a civil case. Assume that both parties agree that the judgment, if the defendant is found liable, will be $100,000 and that the probability of the plaintiff prevailing is 0.40.[49] Suppose further that the plaintiff's trial costs are $20,000 and the defendant's trial costs are $10,000. The plaintiff's expected gain from going to trial is thus $20,000 ($100,000 × 0.40 − $20,000) and the defendant's expected loss is $50,000 ($100,000 × 0.40 + $10,000). Given that the plaintiff's expected gain from a trial is $20,000, she will be willing to settle out of court for any amount greater than this. Similarly, the defendant, with expected costs of $50,000 associated with going to a trial, will be willing to settle for any amount less than $50,000. There is thus a $30,000 range within which a settlement can occur, and if the transaction costs associated with the bargaining process are relatively low, it can be expected that the parties will reach a mutually agreeable settlement within this range.[50]

This same type of analysis can be extended to criminal law and issues surrounding settlements in criminal cases. In a criminal case, the defendant weighs his expected sentence (including the probability of conviction) against the sentence offered in the plea bargain. The prosecutor, whose utility may be defined to include the number of convictions, weighted by the severity of the sentences, will consider the expected sentence (again, weighted by the probability of conviction) in determining what, if any, terms to offer in a plea bargain. If there exist expected gains to both parties from a settlement, then there will be a range of sentences within which a settlement may be reached.[51]

Numerous factors affect the issue of settlement versus trial, and many of these have been introduced into the economic models describing this process. These include differential bargaining costs across disputants, differences in their respective attitudes toward the risks associated with the adjudicatory outcome, and different estimates of the probability of prevailing at trial. Moreover, the various methods by which the legal costs of going to trial are allocated between the litigants (e.g., the standard U.S. method versus the English Rule) have been analyzed in an attempt to explain the trial settlement pattern under each method.[52]

---

[49] Of course, this analysis generalizes as to the differing expectations of judgment value and differing subjective probabilities of prevailing at trial between plaintiff and defendant. The assumed symmetry within this example serves only to simplify the discussion.

[50] The reader should be able to see the parallels between this idea and the Coase theorem in the allocation of mutual gains from trade through the bargaining process.

[51] See, for example, Landes (1971).

[52] See, for example, Gould (1973) and Shavell (1982).

The second theory offered in support of the efficiency hypothesis is that common law *judgments* will tend to mimic the efficient result of a free market, because inefficient legal rules are likely to be more frequently and more intensively (in the sense of resources expended) challenged in court than are efficient ones. The reason for this is that inefficient rules generate higher expected judgments for the challenging parties, and thus individuals will be willing to spend more time, money, and effort in challenging inefficient rules than they will in challenging efficient ones. This is apparent when one recognizes that the overturning of inefficient rules generates both a greater quantity of total wealth and a redistribution of wealth, whereas the overturning of efficient rules *decreases* total wealth at the same time that it redistributes it (Cooter and Ulen 2004, pp. 436–43; Posner 1990, p. 360). Alternatively, parties may seek some form of alternative dispute resolution, such as arbitration, if judicial decisions continue to promote inefficient allocations (Posner 1992, p. 535). As a result, we can expect to see inefficient rules overturned more frequently, thus increasing the stock of efficient rules over time.

When this idea is combined with the logic of the trial-settlement model, the underlying rationale for the economic theory of the evolution of the common law, as evidenced, for example, in the work of Paul H. Rubin (1977) and George L. Priest (1977), becomes quite clear. Rubin describes it as follows:

> The presumed efficiency of the common law and the decision to use the courts to settle a dispute are related. In particular, this relationship will occur because resorting to court settlement is more likely in cases where the legal rules relevant to the dispute are inefficient, and less likely where rules are efficient. Thus, efficient rules may evolve from in-court settlement, thereby reducing the incentive for future litigation and increasing the probability that efficient rules will persist. In short, the efficient rule situation noted by Posner is due to an evolutionary mechanism whose direction proceeds from the utility maximizing decisions of disputants rather than from the wisdom of judges. (1977, p. 51)

Rubin goes on to point out,

> If rules are inefficient, parties will use the courts until the rules are changed; conversely, if rules are efficient, the courts will not be used and the efficient rule will remain in force. An outside observer coming upon this legal rule would observe that this rule is efficient; but this efficiency occurs because of an evolutionary process, not because of any particular wisdom on the part of judges. (1977, p. 55)

That is, efficiency will tend to obtain regardless of what motivates the judges' decisions. That having been said, Rubin (1977, p. 55) also notes that "intelligent judges" may cause the path to efficiency to be a bit shorter than it would be otherwise.

Priest's model builds on the notion that inefficient rules impose higher costs on parties than efficient rules. As a result, the value of overturning an

inefficient rule tends to be higher to the parties than the value of overturning an efficient one, and thus inefficient rules will, on average, lead to more litigation than efficient ones. As a result, the common law has a tendency toward efficiency because the lower probability of relitigation means that efficient rules are more likely to survive than inefficient rules. Priest goes further, asserting that (i) even if judges were actively hostile to the concept of efficiency in their decision-making and/or (ii) (going beyond Rubin) even if both parties do not have a continuing interest in precedent, the ongoing process will result in a tendency toward the efficiency of the common law.

In summary, for both Rubin and Priest, the efficient rule situation noted by Posner is due to an evolutionary mechanism whose direction proceeds from the utility-maximizing litigation decisions of the disputants rather than from the wisdom of judges (Rubin 1977, p. 51).[53]

John Goodman's "differential investment" approach takes the analysis a step further than Rubin and Priest, attempting to set out specific reasons *why* inefficient rules will be overturned. Goodman assumes that judges are amenable to persuasion by the litigants appearing before the court and are completely unbiased with respect to efficiency, and that any increase in legal expenses by either party will increase that party's probability of securing a favorable judgment. Since the economic stakes are higher under inefficient rules, the party on whom liability initially rests has a greater incentive to spend a larger amount on litigation expense here than under efficient rules. The additional expenditures on litigating inefficient rules thus increase the likelihood that such rules will be replaced by efficient ones. Goodman (1978, pp. 394–95) describes this as follows:

> A model of an adversary proceeding is proposed in which the probability that a particular litigant will win a favorable decision depends upon the efforts of both litigants to influence the court and upon the weight of judicial bias. Since parties before the court have an obvious interest in the decision, they have incentives, not necessarily equal, to affect that decision through efforts that incur legal costs—expenses for legal research, factual investigation, forensic talent, and so forth. The fundamental assumption made throughout is that any increment in legal expenses . . . will induce an increment, however small, in the probability . . . of winning a favorable decision. . . . Even if the weight of past precedents favors inefficient solutions, the side with the greater economic stake in the issue will still have a higher probability of winning any succeeding case so long as the ratio of his economic stake to his opponent's exceeds . . . the ratio of legal expenses by the two litigants that must be maintained in order to insure that they both have the same probability of winning.

[53] More recently, Rubin (2005a) advanced what he has termed the "macro" argument (inspired by Hayek) for explaining the efficiency of the common law to complement the previously advanced "micro" (evolutionary) models.

The third, most strongly held, rationale for the efficiency of the common law lies in the view that judges—implicitly or explicitly—select legal rules that generate efficient outcomes. Simply stated, the hypothesis is that the development of the common law, especially the law of torts, can be explained *as if* the goal was to maximize allocative efficiency—that is, *as if* the judges who created the law through decisions operating as precedents were trying to promote efficient resource allocation. As Posner puts it, "The hypothesis is not that judges can or do duplicate the results of competitive markets, but that within the limits set by the costs of administering the legal system (costs that must be taken into account in any effort to promote efficiency through legal rules), common law adjudication brings the economic system closer to the results that would be produced by effective competition—a free market operating without significant externality, monopoly, or information problems" (1983, pp. 4–5).[54] "It is," says Posner (1990, p. 356; emphasis in original), "*as if* the judges wanted to adopt the rules, procedures, and case outcomes that would maximize society's wealth."

Posner (1992, p. 252) suggests that economic logic pervades the common law and that, in general, when transaction costs are low the common law gives incentives for individuals to "channel their transactions through the market," whereas when transaction costs are high, making market allocation infeasible, "the common law prices behavior in such a way as to mimic the market." The instances where law and economics scholars have found common law rules comporting with the dictates of efficiency are far too numerous to detail here, but a few examples will suffice to illustrate the point. The law of property structures property rights in such a way as to promote value-maximizing exchange. Tort law, through the application of the Learned Hand formula, promotes the taking of cost-justified precautions. The doctrine of impossibility in contract law places liability on the party who could most easily anticipate or insure against the unforeseen contingency. Other examples are found in admiralty, expectation damages, assumption of risk, and the application and nonapplication of punitive damages.[55]

Chicago law and economics does not have a clearly delineated model of judicial behavior or motivation in decision-making. Yet, there is a more-or-less general theme that judges have in mind the overall well-being of society—a concept incorporating numerous social values—in resolving cases. Utility maximization is sometimes rejected as a motivating force for judges on the grounds that judges almost never have any personal stake in the case at hand, and that the judicial system has been designed to insulate judges from significant economic incentives. Posner (1993b) has questioned this view, arguing

---

[54] The Posnerian distinction between the efficiency properties of common law and those of statutory law has been challenged by Rubin (1982). See also the essays in Hirshleifer (1982).

[55] See, for example, Posner (1992, part II), Posner (1990, ch. 12), Cooter and Ulen (2004), and Landes and Posner (1987).

that judges, like everyone else, are rational utility-maximizers, where the utility functions of judges are primarily a function of income, leisure, and judicial voting. In fleshing out the implications of this utility function in the context of the institutional structure within which judges make decisions, Posner is able both to explain certain aspects of judicial behavior and to offer testable predictions that arise from the model.

While the actual motivation underlying judicial decision-making is open to debate, it remains the case that judges are, at times, called upon to "legislate" (Posner 1993b, p. 40). Whereas societal well-being entails both efficiency and distributional considerations, Posner (1990, p. 359) suggests that "prosperity" (i.e., wealth maximization) is a goal "that judges are especially well equipped to promote," while judges can do little, if anything, to promote the redistribution of wealth.[56] Moreover, says Posner (1990, p. 359), judges wish to avoid controversy, and wealth maximization "is a relatively uncontroversial policy." Because the cases at hand are unrelated to judicial self-interest, and/or because it is difficult to rationalize the judge's personal considerations as to "deservingness" within the context of a judicial opinion, the judge is "[a]lmost by default . . . compelled to view the parties as representatives of activities," and "[i]n these circumstances it is natural that he should ask which of the competing activities is more valuable in an economic sense" (Posner 1992, pp. 523–24).

Critics have often maintained that the common law reflects no overt economic logic or phraseology; yet, the proponents of the efficiency theory suggest that the underlying economic logic is clear, even if judges usually do not speak the language—that judges employ language other than efficiency to explain efficient arrangements (Cooter and Ulen 2004, p. 440). Posner (1992, pp. 254–55) even goes so far as to suggest that these efficient doctrines simply reflect common sense, and that whereas the articulation of these doctrines in formal economic terminology lies beyond the capacity of most judges, the commonsensical intuition does not. Against those who suggest that efficiency is not a value that would enter into the judge's decision-making calculus, Posner (1992, p. 255) responds that justice is often "a version of efficiency" and maintains that efficiency "has always been an important social value," especially during the laissez-faire period of the nineteenth century, when the common law received much of its modern shape.[57]

It must be underscored that acceptance of this view—in any or all of its manifestations—that the common law is efficient serves as an ideological barrier to the general promotion of statutory law. That is, it serves to denigrate the

---

[56] See also Posner (1992, p. 255).

[57] Cooter and Kornhauser (1980) examined the evolution of the common law through a Markov chain analysis. They demonstrated that regardless of (i) differential litigation rates, (ii) whether initial liability rules are efficient of inefficient, or (iii) how judges choose rules in making decisions, there is a positive probability that any given rule, efficient or inefficient, will persist in the long run. That is, we can expect to see *some* inefficient rules persist over time.

role of the legislative process—the means by which statutes are passed—on the grounds that statutes have no corresponding mechanism to ensure efficient outcomes (this will be explored in chapter 3).[58] Thus, the logic underlying the Chicago efficiency theory of the common law generally follows the following contour: Whenever the market falls short of providing an efficient allocation of resources due to externalities or some other form of market failure, one can rely on the common law and damage measures, which, the theory's proponents contend, have been demonstrated to be comprised of rules and doctrines that produce efficient results, to give the market a gentle nudge in the direction of maximum social welfare. In general, given the existence of some form of market failure, society need not rely on the legislative branch to adopt regulatory statutes or bureaucratic mechanisms to remedy these problems; all one needs to do is rely on the common law to generate the efficient outcome.[59]

CHICAGO LAW AND ECONOMICS AT WORK

We have seen that the "positive" side of the Chicago approach to law and economics is concerned with assessing the degree to which common law and other legal doctrines comport with the dictates of economic efficiency. When law is seen to depart from the dictates of efficiency, or when new legal issues present themselves, the concern is with fashioning and adopting efficient legal rules to guide judicial decision-making—the *normative* side of law and economics. The goal of this section is to describe how the Chicago school uses the elements of Chicago price theory to make normative legal choices to decide "what the law should be."

The Formulation of Efficient Legal Rules—A General Discussion

The starting point for analyzing instances where the common law departs from the dictates of economic efficiency is with the assertion by Coase (1960, p. 2) that situations of harm are reciprocal in nature. Examples are present throughout the fields of property, contract, and tort law, and so the concept

---

[58] Stearns (1998, p. 715) describes the evolving role of the Chicago school over time: "While early Chicago School scholars employed economic analysis to *explain* then-existing institutions and rules, a new generation of law and economics scholars employed the same tools to *challenge* the alleged inefficiencies of the expanding regulatory state."

[59] Halper (1993, pp. 230–31) describes the same notion as follows: "The Chicagoans argue that outcomes legitimate the superiority of the market as a form of collective decision-making—the market is the most efficient means to allocate resources. . . . [S]tate intervention in the market is warranted only to correct market failure, . . . the common law, effectively the outcome of the day-to-day interactions of market participants, orders the market more efficiently than statutory law, the artificial creature of the legislative state."

applies generally. Under the reciprocal view of harm, the problem of causation becomes much more open-ended than under the traditional legal or Pigovian approaches.[60]

One can see an excellent example of the reciprocity concept in action in Landes and Posner's (1983, p. 110) description of the economics of causation in tort:

> If the basic purpose of tort law is to promote economic efficiency, a defendant's conduct will be deemed the cause of an injury when making him liable for the consequences of the injury would promote an efficient allocation of safety and care; and when it would not promote efficiency for the defendant to have behaved differently, then the cause of the accident will be ascribed to "an act of God" or some other force on which liability cannot rest. In this view, the injurer "causes" the injury when he is the cheaper cost avoider; not otherwise.

The contrast between the economic approach and the traditional measures of "cause-in-fact" (where, in order for a plaintiff to sue, a defendant must have caused the harm) and "proximate cause" (where remote actions may have been the cause of harm and thus causation becomes a matter of degree) is apparent, but the most interesting characteristic of the economic approach here is that "causation becomes a result rather than a premise" within the economic analysis of torts (Landes and Posner 1983, p. 110).

We have already seen that, in a world of zero transaction costs, rights will end up in their highest-valued uses because all mutually beneficial bargains can and will be costlessly struck by rational maximizing individuals.[61] Given such a situation, the judge will not need to be concerned about efficient rights assignment, as the same efficient outcome will be reached regardless of the judge's decision.[62] Indeed, attempts by judges to engage in social engineering will inevitably be fruitless because rights, regardless of how they are initially assigned, will always end up in their highest-valued use. Consider the earlier example of the polluted stream, where the pollution discharge causes $1 million in damage to downstream landowners. The polluter can prevent the damage by installing a filtering device at a cost of $600,000, whereas downstream landowners could eliminate the damage at a cost of $300,000. Efficiency clearly dictates that the pollution be eliminated, since the damage is greater than the cost of abatement, and that the optimal way of abating the pollution is for the downstream landowners to undertake the abatement. The Coase theorem tells us that, no matter how rights are initially assigned, this is exactly the

---

[60] It also constitutes an excellent example of the effects of the "naming and framing" process discussed in chapter 1.

[61] This is analogous to movements from positions off contract curves to positions on contract curves in the Edgeworth Box analysis of microeconomic theory. See appendix B to chapter 1 and Varian (1993, pp. 484–507).

[62] Again, this is subject to the qualifications dealt with in note 42.

solution that will obtain, with the actual distribution of the gains from this exchange depending on the relative bargaining power of the two sides.

While the pollution example provides a useful context for understanding the legal-economic resolution of a combined property and tort disputes, the approach is also applicable to contract law. Suppose that A signs a contract with B which stipulates that A will sell her house to B for $200,000. After this contract has been signed, C enters the picture, offering to pay A $210,000 if A will sell to him. A then breaches her contract with B and sells the house to C for $210,000. Here, it is clearly efficient for A to breach the contract with B and sell to C instead, since C values the house more highly than does B. Suppose next that B files suit against A for breach of contract, asking that the contract be enforced according to its terms. If the court rules against B, then C retains the house—the efficient result obtains. If instead the court finds in favor of B, the house is transferred from C to B, resulting in an inefficient solution. Just as in the pollution example, however, this is not the end of the story. If C values the house at $210,000 and B only at $200,000, then C will be willing to offer B any amount up to $210,000 for the house. B, who values the house at $200,000, will be willing to accept any amount greater than that to give up the house. Thus, a mutually beneficial bargain will be struck, and the efficient result—C owning the house—will obtain. In a world of zero transaction costs, we thus get the efficient result regardless of the initial assignment of rights— that is, whether or not the court rules on behalf of A or on behalf of B. Coase (1960, p. 10) recognized this, observing that "[w]ith costless market transactions, the decision of the courts concerning liability for damage would be without effect on the allocation of resources." In recounting a judicial ruling in a land-use case, he then almost went on to chide the judges by stating,

> It was of course the view of the judges that they were affecting the workings of the economic system—and in a desirable direction. Any other decision would have had "a prejudicial effect upon the development of land for residential purposes." . . . The judges' view that they were settling how the land was to be used would be true only in the case in which the costs of carrying out the necessary market transactions exceeded the gain which might be achieved by any rearrangement of rights. (p. 10)

Otherwise, the parties would bargain to an efficient use of the land. "But of this" he says, "the judges seem to have been unaware" (p. 10).

The problem, of course, is that the real world is, at least for the most part, a world of positive transaction costs. As discussed earlier, these costs may at times may be sufficiently high to preclude bargaining all together,[63] resulting in what may be called a legal "fly-paper effect"—the right sticks where it hits. Here, the court's decision on behalf of one party or the other will directly

---

[63] More specifically, where transaction costs exceed the economic surplus from a bargain.

determine the final resting place of the right and thus will impact the efficiency of the end state. The goal of efficiency in Chicago law and economics manifests itself in the normative prescription—the Coase lesson—that rights should be assigned in a way that maximizes the wealth of society.[64]

Acknowledgment of the fly-paper effect requires an approach that constitutes what has been called "mimicking the market" and involves assuming a hypothetical zero-transaction-costs world, as in the Coase theorem. The question then becomes one of ascertaining where, in a zero-transaction-cost world, the ultimate resting place of the right will be. By engaging in this type of inferential analysis, courts can attempt to discern the efficient outcome that would have resulted from the market-like machinations of the Coase theorem. By assigning rights in such a way as to achieve this result, the courts can accomplish the goal of wealth maximization. Thus, in a dispute over terms or provisions that are missing from a contract because positive transaction costs made it uneconomical, ex ante, to specify fully the contract over all states of the world, the task of the judge is to attempt to infer which party would have accepted liability for the contingency at issue in a world of costless bargaining. Since it can be inferred that the party accepting liability would have been the one who could have done so at least cost, the efficient decision would be to place ex post liability on that party. The same logic explains why, if the gains from a breach of contract exceed the costs of breach, then the court should allow the contract to be breached.

Law of Torts

The evidence for the efficiency of the common law is thought to be particularly prevalent in the field of tort law, although, as we have noted, much of the literature purports to offer evidence supporting the view that common law doctrines of property and contract also promote efficiency. In the Chicago analysis of tort law, the subtle transition from positive description of efficient common law evolution to normative remedy prescription is easy to discern: if

---

[64] Thus, in terms of the pollution example, since the least-cost method of reducing pollution is for the downstream landowners to abate that pollution, the efficiency criterion would dictate that the polluter be given the right to pollute. As we saw, this would induce the landowners to undertake abatement since the alternative is sustaining $1 million in damage. Of course, the optimal amount of pollution may be some intermediate amount, as opposed to all or nothing. In this case, a property rule will not give the optimal result when transaction costs preclude bargaining. As Calabresi and Melamed (1972) have pointed out, efficiency would dictate the use of a liability rule in such a situation, since the polluter would then pollute only as long as its benefits from additional pollution exceed its costs (as measured by compensatory damages), thereby engendering the optimal outcome. Similarly, referring back to the example from contract law, the court should, from an efficiency perspective, allow A to breach her contract with B and sell to C, since, in doing so, the house is placed in a higher-valued use.
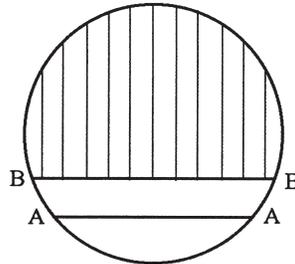
FIGURE 2-6.

society has an array of liability rules from which to pick, rules that may, for example, lead to a reduction in accidents or to the abatement of pollution, then the economic logic suggests choosing the least-cost alternative.

This line of reasoning was originally presented by Calabresi in "Some Thoughts on Risk Distribution and the Law of Torts" (1961), coincidentally at about the same time Coase authored "The Problem of Social Cost."[65] The analysis of efficient liability rules is based upon three suppositions: (i) that all losses can be expressed in monetary terms, (ii) that the quantity of undesirable activities can be reduced by devoting more of society's scarce resources to precautionary or preventive activities, and (iii) that those individuals who are potentially involved in harmful activities are sensitive to cost pressures (Burrows and Veljanovski 1981, p. 11). The aim of efficiency-based tort law is to use ex post damage awards to replace what may be termed "ex ante unfeasible agreements" that would have occurred had market transactions been possible (that is, if transaction costs were sufficiently low).

The economic logic of this approach can be expressed in three simple steps. First, let figure 2-6 represent all the costs associated with, say, automobile accidents (including the actual damages suffered by victims, litigation expenses, legal administrative costs, enforcement costs, etc.) before the government attempts to reduce society's costs of accidents. Assume that the government first installs traffic signals and stop signs, with the net savings to society being the area beneath line AA. Now assume, additionally, that the government erects costly median barriers, again with some net savings, this time equal to the area ABBA. The area above line BB represents the residual costs of accidents, or, as they are commonly referred to, the *interaction damage costs.* It is these costs that are the focus of the literature on economics of liability rules.

The economic approach to the analysis and selection of liability rules can be viewed as one method by which risks in society are determined to be either

---

[65] We will explore Calabresi's approach more fully in chapter 6 in the section on the New Haven school.
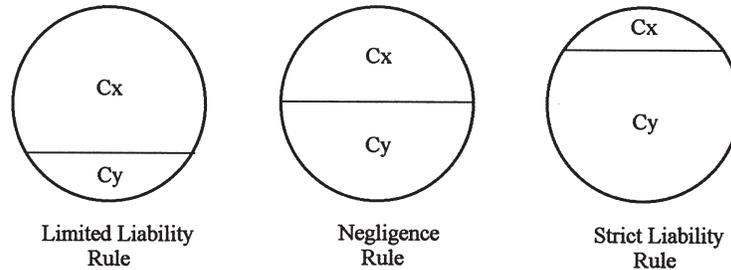
FIGURE 2-7.

background risks, with the costs borne by the victim, or, alternatively, *compensated risks*, which arise from tortious acts and whose costs are borne by the injurer. The primary focus of the economic approach to liability rule formulation is to develop a standard of liability that minimizes the sum of interaction damage costs with full recognition that the selection of a specific liability rule will determine which actions will be considered background risks (and go uncompensated) and which will be deemed compensated risks. Focusing only on the interaction damage costs—the area above the line BB in figure 2-6, now replicated three times and equal to each of the three circles in figure 2-7—we can easily describe the two facets of the problem.

First, in each panel of figure 2-7 let C_x represent all costs borne by victims, because the liability rule chosen considers the risks associated with these interaction damage costs to be background risks. In a like manner, in each panel of figure 2-7 let C_y represent all compensable costs borne by the injurers, because the liability rule chosen considers the risks associated with these interaction damage costs as tortious and not part of society's background risks. It is evident that the liability rule which is chosen will determine the extent of noncompensable background risk vis-à-vis that of compensable risk. Consider, for example, three hypothetical liability rules. With a limited liability rule, most risks are considered background risks, and most injurers are not liable for the damages caused—that is, the victims bear most of the costs. At the other extreme, where injurers are held strictly liable, most risks are considered compensated risks and thus the injurers are liable for most of the damages caused. Perhaps somewhere between the two lies a third liability rule, a negligence rule, that establishes a system under which injurers are liable for the damages only if they are negligent (i.e., at fault). The first fundamental point illustrated here is that the legislative or judicial choice of one liability rule over another results in a different distribution of interaction damage costs for society. It is only for expositional purposes that the three circles have the same area (i.e., represent the same interaction damage costs). Indeed as stated earlier, the thrust of the economic approach to liability rule formulation is to develop a standard of liability that minimizes the sum total of interaction damage costs.

Cx = $200

Cy = $800

Cx = $500

Cy = $300

Cx = $360

Cy = $240

Limited Liability
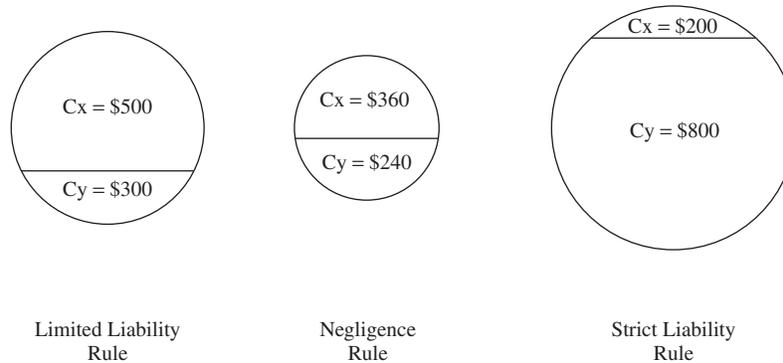Rule

Negligence
Rule

Strict Liability
Rule

FIGURE 2-8.

Figure 2-8 illustrates the efficiency consequences of the choice of one liability rule over another. The sizes of each of the three circles—indicating the sum of interaction damages costs, arbitrarily sized, for each rule—are purposely different to convey the notion that there would be no reason to believe that all rules would result in identical interaction damage costs. For example, the third circle, which illustrates a hypothetical strict liability rule, depicts both the distribution of costs—80 percent of which are borne by the injurers, and only 20 percent of which are uncompensated and thus borne by the victims—and the total size of the interaction damage costs, here equal to $1,000. The second circle, by way of contrast, illustrates the situation for the rule of negligence, which yields a different distribution of costs (40 percent being borne by the injurers and 60 percent of the costs going uncompensated and thus being borne by the victims), with the total interaction damage costs equal to $600. In comparing these two rules, the economic approach would suggest adopting the negligence liability rule on the grounds that it allocates risks, alters incentives, and thereby affects behavior so as to minimize the sum cost to society ($600, clearly less than $1,000). Similarly, and also on efficiency grounds, the negligence rule is preferred to the limited liability rule (the first circle, where most costs go uncompensated and thus borne by the victims as part of background risks), since the limited liability rule results in total interaction damage costs of $800, as compared to $600 under the negligence rule. It should be noted that the explicit adoption of the least-cost solution carries with it the implicit adoption of the specific distribution of risk that goes along with the chosen rule. This is an important point of departure between the Chicago school and the traditional justice-fairness approach inherent in conventional tort law. Whereas the justice-fairness approach is concerned with the distribution of the costs associated with risk (i.e., the question as to where *within* the circles one draws the line separating compensable risk from background risk),

the Chicago approach to law and economics is concerned with minimizing the total costs associated with risk (that is, choosing the smallest circle), the distribution of the costs being irrelevant except as it affects the total cost.

Before exploring one of the most notable examples of these ideas entering case law, it is important to understand the nature of the concept of negligence as used in common law. Consider the following quote by Chief Justice Rosenberry in *Osborne v. Montgomery* (234 NW 2nd 372, 378–380 [1931]), which provides a general outline as to what constitutes a negligent action:

> Every person is negligent when, without intending to do any wrong, he does such an act or omits to take such precaution that under the circumstances present he, as an ordinarily prudent person, ought reasonably to foresee that he will thereby expose the interest of another to an unreasonable risk of harm. In determining whether his conduct will subject the interests of another to an unreasonable risk of harm, a person is required to take into account such of the surrounding circumstances as would be taken into account by a reasonably prudent person and possess such knowledge as is possessed by an ordinarily reasonable person and to use such judgement and discretion as is exercised by persons of reasonable intelligence and judgement under the same or similar circumstances.

What the court determines is or defines as "reasonable" is the driving force in separating background risks from compensated risks in society. While one can argue that this determination should be made on the basis of moral and ethical considerations, the Chicago approach to law and economics suggests the use of the efficiency criterion—that society select the rule that minimizes costs and live with the accompanying risk distribution.

Interestingly, however, the injection of economic thinking into the legal concept of negligence did not originate with the Chicago school. Rather, it was introduced by Judge Learned Hand in his decision in *United States v. Carroll Towing Co.*[66] In his ruling, he gave explicit formulation to the standard that the courts had applied throughout the period that the negligence standard flourished (as in *Osborne* v. *Montgomery*)—and he did so in economic terms. This formula has come to be known as the Hand negligence formula—the injurer is liable for accident damages if the expected accident costs (the probability of an accident occurring times the victim's accident costs) are greater than the costs that the injurer must bear to avoid the accident. Otherwise, the injurer is not negligent and the victim bears the damages. Hand stated that the matter of liability turns on the minimization of expected costs:

> Possibly it serves to bring this notion into relief to state it in algebraic terms: if the probability be called P; the injury, L; and the burden, B; liability depends

---

[66] *United States v. Carroll Towing Co.*, 159 F.2d 169 (1947). Not all agree as to the jurisprudential significance of the Hand rule; for instance, see Wright (2003).

upon whether B is less than L multiplied by P: i.e., whether B is less than PL. (*United States v. Carroll Towing Co*.,1947, at 173)

A simple numerical example will help explicate the workings of the Hand rule and the incentives it creates for potential injurers. Following Hand's notation, let L be the costs of accidents borne by the victims and let B be the potential injurer's costs of avoiding accidents. Let P be the probability that the accident will occur if no avoidance is taken and P′ be the probability that the accident will occur if some avoidance activities are undertaken. Let L = $1000, B = $500, P = .8, and P′ = .2.[67] Again, if P × L > B, then the injurer would be ruled negligent and thus liable for the damages caused. In this case, the Hand formula asserts that the injurer should be liable because expected accident costs, $800 (.8 × $1000), are greater than the costs of avoidance, $500. Since he would be liable for injuries caused in these circumstances, the potential injurer will engage in accident avoidance measures and thereby reduce the probability of accidents from .8 to .2.

The economic rationale for the Hand rule is that it minimizes expected costs (i.e., the sum of accident and avoidance costs). That is, by undertaking avoidance activities, the potential injurer reduces the probability of accidents to .2, resulting in expected accident costs equal to $200, which when added to the avoidance costs of $500 sums to a total of $700. If the Hand rule was not in force and injurers were therefore ruled not liable in these situations, then our prospective injurer would not engage in avoidance activities, leaving expected accident costs at $800 and, given his lack of undertaking any avoidance measures, the avoidance costs would be zero, for a sum total of $800. Inasmuch as the $700 expected cost associated with undertaking avoidance activities is less than the $800 expected costs in the absence of avoidance, adherence to the Hand formula dictates that the injurer should be liable, thereby forcing avoidance measures and minimizing the expected costs associated with accidents.

Consider an alternative case where the simple Hand formula would indicate that the injurer should not be held liable. As before, let L = $1000, P = .8 and P′ = .2, but now assume that B = $900. Since the expected accident costs of $800 (.8 × $1000) are less than the costs of avoidance, $900, the injurer should not be held liable. Here, then, the Hand rule would make the victim bear the costs. The efficiency consequences of this application of the Hand formula parallel those in the previous example. Since injurers are not liable here when avoidance is not cost-justified, potential injurers will refrain from any accident avoidance activities in those circumstances. In the present example, expected accident costs are $800 (.8 × $1000) and avoidance costs are equal to zero for a total expected cost of $800. Note that if the courts had not followed the simple Hand formula and injurers were ruled liable, then, in

---

[67] Note that P and P′ need not sum to 1.

these circumstances, the potential injurers still would not engage in avoidance activities. The reason is undertaking avoidance results in expected accident costs of $200 (.2 × $1000) plus avoidance costs of $900, for a total expected cost of $1100. The potential injurer will choose the option with the lower expected cost, which here means foregoing avoidance activities.

The Hand formula has been seized upon by those opposed to suggestions that the determination of negligence should be based on moral and ethical considerations. As Posner wrote, "Characterization of the negligence standard as moral or moralistic does not advance analysis. . . . Negligence is an objective standard. . . . To characterize the negligence concept as a moral one is only to push inquiry back a step" (Posner 1972a, pp. 31, 32). The more general contrast between the Chicago law and economics and the justice-fairness perspectives is seen in the following statements by Richard Posner and Richard Epstein—the latter of whom was a long-time critic of Chicago law and economics[68]—in their debate over the rule of strict liability. Speaking from the law and economics perspective, Posner (1973, p. 221) argued that "[s]ince the efficient use of resources is an important although not always paramount social value, the burden, I suggest, is on the authors to present reasons why a standard that appears to impose avoidable cost on society should nonetheless be adopted. They have not carried this burden." Epstein (1973, p. 152), in contrast, contends,

> Once it is admitted that there are questions of fairness as between the parties that are not answerable in economic terms, the exact role of economic argument in the solution of legal questions becomes impossible to determine. It may well be that an acceptable theory of fairness can be reconciled with the dictates of economic theory in a manner that leaves ample room for the use of economic thought. But that judgment presupposes that some theory of fairness has been spelled out, which, once completed, may leave no room for economic considerations of any sort.

In essence, Posner is claiming that the goal of legal decision-making should be efficiency, and that the burden rests with the advocates of the justice-fairness approach to demonstrate why fairness considerations should dominate efficiency. Epstein argues from the opposite pole, setting up fairness as the goal and stating, in a sense, that the burden rests with the advocates of the economic approach to show why efficiency considerations should dominate fairness.

TORT REFORM

To a large extent, this is the very logic that underlies the ongoing impetus for tort reform in the United States. To place the issue in perspective, a milestone was reached in 1990: for the first time, the number of new cases filed in state

---

[68] He is now a fellow traveler.

courts topped 100 million. Advocates of tort reform, like the former Vice President Dan Quayle, have argued that "the costs of the civil justice system amount to a self-inflicted competitive disadvantage" (Jost 1992, p. 435). By the year 2000, the President's Council of Economic Advisers figured that "excessive tort costs" amounted to a $650 "litigation tax" on the economy for every person in the United States. With the Republicans winning both the U.S. House and Senate in 2002, they promised to follow up on the tort reform initiatives that they tried to push through in the 1990s.[69] The rhetoric has heated up even further. In his 2003 State of the Union address, President Bush said, "No one has ever been healed by a frivolous lawsuit"—to which the Trail Lawyers Association responded, "True, but neither has anyone been healed by tort reform" (Karp 2003, p. 1). The result is that for the past decade we have witnessed tort-reform advocates demonizing trial lawyers, while trial lawyers simultaneously argue that the tort reformers are turning their backs on those who have suffered damages and deserve to be made whole once again.[70]

An understanding of the theory reflected in the three different-sized circles enables one to gain some insight into the positions of those who argue for the efficient reform of tort law on the one hand and those who want to focus on using the rules based on justice and fairness on the other. It also brings into sharper relief much of the political rhetoric between the conservative Republicans, who advocate tort reform along lines consistent with Chicago-school thinking, and the liberal Democrats, who echo the fairness claims of the trial lawyers, whose interests liberal Democrats tend to represent. J. M. Balkin explains why the Chicago approach to law and economics has found such a warm reception among conservative academics (1987, pp.1454–55). In his article, he describes the close nexus between American conservatism and the Chicago approach to law and economics, and their corresponding "relatively individualistic" positions in regard to legal-economic policy in general and tort law in particular. It is this "relatively individualist" approach that is, he says, at the root of the advocacy of tort reform by the Chicago school—an approach that tends to limit the defendant's liability, or deemphasizes the defendant's responsibility for an injury to the plaintiff. On the other side of the argument over tort reform are those who adopt what Balkin terms the "relatively communalist" position, an alternative approach that tends to expand the defendant's liability, emphasizes the defendant's responsibility to others, and oftentimes serves a redistributive function. The relevant point here is that, for the

---

[69] This is documented by Coyle (2002).

[70] While the American Tort Reform Association (ATRA) was formed in 1986 to represent hundreds of U.S. and foreign corporations in their bid to overhaul civil liability laws at the state and national levels, much of the early efforts in tort reform came from Vice President Dan Quayle, as head of President Bush's Council on Competitiveness, established in 1989. The Council also assigned then-Solicitor General Kenneth Starr the task of developing a plan to overhaul the country's civil liability laws.

Chicago school, the higher levels of duty in tort law become superfluous because they waste societal resources on administrative costs, whereas the relatively individualist rules—those that lower standards of liability—are preferred based on efficiency considerations. Chicago advocates of the individualist approach take the additional step of denying that liability rules are an efficient method for redistributing income and make the further point that their effects on this front are often unpredictable.

Recall from our hypothetical example that each circle (in figure 2-8) represents all the costs associated with, say, "generic" accidents, including the actual damages suffered by victims, litigation expenses, legal administrative costs, enforcement costs, and so forth, after the government had taken certain obvious actions to reduce some of the most obvious costs. Also, keep in mind that the costs associated with generic accidents are, on the reverse side of the ledger, an income entry for those who collect the fees: the bigger the circle, the greater the costs to one group, hence the larger the income earned by the other. Once the two major parties to the dispute over tort reform are identified—the Chamber of Commerce on one side and the Association of Trial Lawyers of America on the other—their positions emerge. Generally, the argument goes something like this: The trial lawyers want to have the smallest possible share of damages treated as part of life's background risks; hence, they argue that we should rely on the principles of justice and fairness and advocate a rule of law more along the lines or in the direction of strict liability. This is clearly the very rule that increases costs and hence incomes—including incomes of the trial lawyers.

The position of the Chicago-efficiency tort reformers and, in particular, the Chamber of Commerce, is that there are a host of efficiency-enhancing reforms that the legislature should set in place. These would reduce the sum of the costs borne by society owing to generic accidents (and with it, of no concern to them, the incomes of trial lawyers). As a consequence, they advocate a package of reforms that they believe would, when taken together, yield the smallest circle of costs, as reflected in the more limited liability rule—all this, of course, independent of the concepts of justice and fairness advocated by the trial lawyers.[71]

In actuality, tort reforms have been difficult to set in place, in part because

---

[71] The ultimate validity of each of the parties' positions, is partly an empirical question related to the actual differences in the costs under each liability rule and what portion of costs get translated into trial lawyer incomes. In mid-2003, a report on the economics of the litigation of tort claims was issued by the Center for Legal Policy—Manhattan Institute (2003) titled "Trial Lawyers Inc.: A Report on the Lawsuit Industry in America 2003." The report found that in the United States, total costs for tort litigation now exceed $200 billion annually (approximately 2 percent of the nation's GDP), with trial lawyers getting about $40 billion a year. In addition, the study found that the trial lawyers' revenue has grown at about 9 percent per year over the past three decades.

they are seen as a national problem but one with little federal jurisdiction. The federal courts handle only a small portion of the nation's caseload; it is the individual states that carry the bulk of the cases. Thus, federal initiatives for tort reform typically confront states' rights barriers and force the reformers to go state by state to win approval of their reforms. Since 1991, tort reform advocates have set up dozens of tax-exempt groups in at least eighteen states to get out their "lawsuit abuse" message in the media in an effort to influence legislation, the judiciary, and jurors. As of May 2003, tort reform packages including some or all of the provisions described here have been introduced in seventeen states and tend to contain some or all of the following provisions:

> Abolishing joint and several liability—the longstanding principle, which requires losing defendants in tort suits to pay all the damages regardless of their level of fault. Restrictions on venue to keep plaintiffs from seeking the friendliest courts. Legislation to allow trial court evidence on so-called collateral sources of compensation, such as medical insurance. An outright ban on punitive damages, though in practice caps are often imposed as a compromise. Restrictions on noneconomic damages such as pain and suffering. Caps on appeal bonds. And the allowing of appellate challenges to class action certification. (Ballard 2003, p. 1)

Thus far, judges have not clamped down on plaintiffs' lawyers as much as the Chamber of Commerce would like, and Congress has failed to deliver meaningful tort reform. Individual states sputter along taking incremental steps—some succeeding in obtaining liability caps, eliminating various causes of action, and implementing procedural rules making it much harder to sue wrongdoers. Many times, however, their reforms are challenged as unconstitutional by members of the Association of Trial Lawyers of America, with varying results.

The Law of Contracts: Efficient Breach

The economic analysis of contract law provides a unique perspective on the rules, distinctions, categories, and boundaries of the evolving legal doctrine of contract law.[72] From an economic perspective, the major thrust of contract law is to facilitate the accomplishment of the goals of the contracting parties—and to do so efficiently (Goldberg 1989, p. x). Contract law concerns the structuring of contracts, the negotiation of settlements among the parties of interest in a dispute over their asserted legal rights, rules governing arbitration, and, in a lawsuit, the defenses of breach of a contract. If the contracting parties took the

---

[72] Comprehensive treatments of the economics of contract law include Craswell and Schwartz (1994), Cooter and Ulen (2004, chs. 6, 7), Goldberg (1989), Kornhauser (1986), Kronman and Posner (1979), and Posner (1992, ch. 4).

time and incurred the costs to incorporate explicitly into the contract terms that deal with every type of nonperformance under every possible circumstance, there would be no occasion for the courts to prescribe remedies. Obviously, however, from an economic standpoint, it is too costly to negotiate and draft contracts that provide for every conceivable contingency, and this makes disputes unavoidable. In Coasean terms, the problem is one of transaction costs, and thus, as A. Mitchell Polinsky (1989, p. 27) describes it, "Contract law can be viewed as filling in these 'gaps' in the contract—attempting to reproduce what the parties would have agreed to if they could have costlessly planned for the event initially." Thus, contract law has a clear economic justification.

Obligations that arise consequent to a contract are influenced by the answers given to two questions. The first has to do with the circumstances under which an otherwise binding promise would be excused, and the second concerns the issue of what legal consequences flow from breaching a contract. For our purposes here, legal rules of contract can be divided into two distinct classes. In one class are the so-called *immutable rules*, those that govern even if the parties try to contract around them. The other class consists of the so-called *default rules*, which parties can contract around by prior agreement; these are the rules that fill the gaps in incomplete contracts (Craswell and Schwartz 1994; Ayres and Gertner 1989). The assessment of default rules and the legal consequences that flow from breach constitute a major line of literature of the economics of contract law and are the focus of the present discussion.

It may seem odd at first glance to think of a breach of contract as a "good" thing. Yet as Robert Cooter and Thomas Ulen (1988, p. 290) have observed, "One of the most enlightening insights of law and economics is the recognition that there are circumstances where breach of contract is more efficient than performance." They go on to define efficient breach as follows: "the breach of a contract is more efficient than performance of the contract when the costs of performance exceed the benefits to all parties" (Cooter and Ulen 1988, p. 290). The major thrust of the literature, then, is to analyze and assess rules according to the extent to which they promote efficient breach. As with much of Chicago law and economics, doing so requires a return to Pareto efficiency and therefore to Coase. Markets and contracting are not costless processes; thus, following the Coase lesson, contracts should be structured to approximate the efficient outcome by providing incentives for value-maximizing conduct. The ubiquitous reliance on Coasean thinking throughout the economics of contract law is evidenced in Craswell and Schwartz's (1994, p. 54) assertion that "[w]henever a remedial rule requires conduct that seems inefficient, you should consider whether the party who would lose the most from that inefficiency will be able to pay the other party to abandon his or her insistence on the inefficient remedy."

If a party breaches a contract and the court determines that (i) the contract has been validly formed, and (ii) performance will not be excused, then a remedy for the breach must be determined. For our purposes here we will consider

only three options, two of which—legal relief and equitable relief—are court-determined and the third of which—liquidated damages—is not court-determined, but structured by the contracting parties themselves.

1. *Legal relief*: This remedy consists of court-imposed money damages and is the remedy most often imposed by the court. Legal relief can take one of three forms:

Expectation damages: award the victim of breach an amount of money that leaves him in a position identical to the one he would have been in if the contract had been completed according to its terms.

Reliance damages: award the victim of breach an amount of money that puts him in the same position he would have been in if he had not entered into the contract in the first place.

Restitution damages: award the victim of breach an amount of money equal to any benefits (e.g., payments) he has provided to the breaching party.

2. *Equitable relief (specific performance)*: Although not used often, equitable relief is the second court-determined remedy that can be imposed, and is termed "specific performance." Specific performance protects the potential victim's right to performance by ordering the breaching party to perform the contractual promise. That is, the court essentially protects the victim's right with the equivalent of a property rule.

3. *Liquidated damages*: Unlike the first two remedies—legal relief and equitable relief—that are imposed upon the parties by the court, the third remedy—liquidated damages—is imposed by the parties themselves. The liquidated damage remedy awards the breached-against party an amount of money agreed to by the parties at the time of contract formation (Polinsky 1989, pp. 64–65).

The economic analysis of breach of contract asks, at the most basic level, whether and under what conditions the expectation, reliance, restitution, specific performance, and liquidated damages remedies for breach of contract are efficient alternatives to a fully specified (ergo, Pareto-efficient) contract. Putting it slightly differently, the goal, from an efficiency perspective, is to employ the legal remedy that at once prevents inefficient breach and at the same time induces breach when breach is efficient. Each of the remedies mentioned can be examined with this in mind, and we shall do so in the context of a simple example.

Suppose that Acme Autos has a car for sale at a price of $12,000. Fred is willing to pay up to $15,000 for the car, signs a contract with Acme to purchase the car at the $12,000 price, gives Acme a down payment of $5,000, and agrees to return in three days to make the remaining payment and take possession of the car. Shortly thereafter, Fred goes to Parts Unlimited—an auto parts dealer, where he purchases new wheels and tires for this car during their going-out-of business sale, spending a nonrefundable $800 in the process. Not

long after Fred leaves Acme, Jim arrives and offers Acme $13,000 for the car—an amount equal to the value that *he* places on owning it. Setting to one side for the moment the issue of damages, Acme will wish to breach its contract with Fred in order to sell the car to Jim, since it can realize an extra $1,000 from the transaction. Since Fred, however, values the car at $15,000, whereas Jim only values it at $13,000, such a breach would be inefficient, because it would place the vehicle in a lower-valued use. With this as background, we can now examine the effect that the various remedies described earlier would have on Acme's potential breach of the contract with Fred.

EXPECTATION DAMAGES

An expectation damages remedy makes the victim as well off as if the contract had been performed. Thus, if Acme went ahead and sold the vehicle to Jim, an expectation damages remedy would require Acme not only to refund Fred his $5,000 down payment, but also pay Fred $3,000 (the surplus from the transaction that Fred lost due to the breach). Thus, under the expectation damages remedy, Acme would decide not to breach the contract since the gain from selling to Jim ($1000) is less than the damages Acme would have to pay to Fred ($3,000). If, on the other hand, Jim had offered Acme $20,000 for the car, a decision by Acme to breach would be efficient, since Jim values the car more highly than Fred. That is, Acme would breach its contract with Fred and sell to Jim, since the gain from selling to Jim ($8,000) exceeds the damages that they would have to pay to Fred ($3,000). More generally, under the expectation damages remedy, Acme will refrain from breaching its contract with Fred for any amount less than $15,000 and will breach its contract for any amount greater than $15,000.

From the vantage point of economics, the expectation damages remedy encourages efficient breach and discourages inefficient breach. The economic reasoning here is straightforward and can be generalized. The expectation remedy requires the breaching party to pay damages equal to the value of the good or resource to the victim of the breach. If some other buyer enters the arena and values the good or resource more than the initial buyer, then it is efficient for that third party to acquire the good—that is, efficiency dictates that goods and resources should gravitate to their highest-valued uses. Consequently, if the parties anticipate that the court will impose an expectation damages remedy, the seller will have an incentive to breach in order to obtain the higher offer, pay the victim his expectation damages, keep the surplus, and thereby end up better off—with no one made worse off. On the other hand, if a prospective buyer who did not value the good as much as the initial buyer entered the arena, the expectation remedy would appropriately and efficiently discourage the breaching of the contract.[73]

[73] On the limits of expectation damages, see Goetz and Scott (1983).

RELIANCE DAMAGES

Reliance damages place Fred in the same position as he was in prior to signing the contract, and as such compensate him for expenditures undertaken in reliance on the contract. Recall, in this case, Fred had made a nonrefundable $800 expenditure on wheels and tires, and thus, under a reliance damages remedy, Acme would be forced to compensate Fred in the amount of $800, plus refund his $5,000 down payment.

It should be apparent that the reliance damages remedy provides an incentive for inefficient breach. With Jim's offer of $13,000, Acme can gain an additional $1,000 by selling the car to Jim, and, since Acme must compensate Fred only an amount of $800, this provides Acme with an incentive to breach the contract. Since Fred, however, places a higher value on the car ($15,000) than does Jim ($13,000), this breach is inefficient—resources would not be gravitating to their highest-valued use. More generally, reliance damages do not internalize to the potential breacher the full cost of the breach to the victim (i.e., the "price" of breaching is inefficiently low) and, as a result, can lead to breaches where the social costs are in excess of the social benefits.

RESTITUTION

Under the restitution damages remedy, the breacher need only return to the victim any monies received from the him—here, Fred's $5,000 down payment. Given this, Acme will breach its contract with Fred and sell to Jim, and, indeed, would do so at any price in excess of $12,000. As with reliance damages, restitution damages do not internalize all costs of breach to the potential breacher, and, as a result, give rise to the potential for inefficient breach.

Thus, among the monetary remedies, only expectation damages, by internalizing the full cost of his actions to the potential breacher, act simultaneously to discourage inefficient breach and to promote efficient breach.[74]

SPECIFIC PERFORMANCE

Specific performance involves an order by the court that the contract be completed according to its terms. Thus, in the context of our example, the court would order Acme to sell the car to Fred for $12,000. As such, specific performance eliminates the possibility of inefficient breach here. It might be argued, however, that specific performance will have the effect of preventing efficient breaches. For example, if Jim had offered $20,000 for the car, then the court's

---

[74] There are two additional efficiency issues that arise in the discussion of the appropriate form of damages. These relate to the effect of the damage remedy employed on (i) the potential breacher's incentive to undertake precaution against breach and (ii) the potential victim's incentive to undertake expenditures in reliance on the contract. For a discussion of these issues, see Cooter and Ulen (1988, ch. 7).

requirement of specific performance, by awarding the car to Fred, would generate an inefficient outcome. Yet, in such a situation there exists a subsequent exchange that would exhaust the gains from trade. Jim would presumably also be willing to pay Fred $20,000 for the car, and Fred, valuing the car at $15,000, would be willing to part with it at the $20,000 price. Thus, unless transaction costs are prohibitive, we would expect that the specific performance remedy would, through subsequent transactions, generate the efficient outcome, and that the difference between specific performance and expectation damages (previously shown to be efficient) will be purely distributional. That is, under specific performance, Fred will capture the gains from Jim's greater willingness to pay, whereas, under expectation damages, Acme will capture those gains.

Based on this, it has been argued that courts should rely more heavily on specific performance as a remedy for breach.[75] This is particularly true if the good is one for which the value of performance to the victim is difficult for the court to discern with certainty. If, for example, the victim places a very high subjective value on the good in question (a value much higher than others would place on it), the court, in questioning whether the victim is accurately revealing his preferences, may well award insufficient damages. Under a specific performance remedy, however, the breacher will have every opportunity to negotiate with the victim and pay him his full value if he is to consent to the breach. More generally, since valuations are more likely to be accurately reflected through the bargaining process than through court-stipulated damages, specific performance offers a greater potential for an efficient resolution of the dispute than do monetary damages as long as transaction costs are low; conversely, monetary damages are to be preferred when transaction costs are high.[76]

LIQUIDATED DAMAGES

Liquidated damages clauses are inserted into the contract by the parties themselves during the contract-formation process and specify the damages that one party must pay to the other in the event that the contract is breached. Nonetheless, courts have traditionally been unwilling to enforce such clauses when the specified damages appear to be overly burdensome or punitive (that is, appear to exceed compensation for the harm caused to the victim). The economic approach, however, places these damages in a somewhat different relief. Since rational maximizing agents will agree only to a contract that both believe to be in their interests, these clauses would seem to be efficient; otherwise, the parties would not have agreed to their inclusion in the first place. There are two important reasons why maximizing agents might agree to such terms. First, the potential victim might place a very high subjective value on performance

---

[75] See, for example. Cooter and Ulen (2004, pp. 259–60) and, more generally, Schwartz (1979).

[76] This is parallel to the argument made by Calabresi and Melamed (1972), discussed earlier.

of the contract—much, much higher than the average person. The high level of liquidated damages specified in the contract may thus reflect the high subjective value placed on performance by the victim, and, in a sense, insurance against nonperformance. Alternatively, the potential breacher may be willing to insert such a clause as a quality signal to attract customers. In either case, the clause is mutually beneficial ex ante and thus, from an efficiency perspective, should be enforced.[77]

We hope this brief review makes it clear that Chicago law and economics provides novel and interesting theoretical insights into the field of contract law. It must be noted, however, that the substantial impact that the Chicago approach to law and economics has had in the areas of torts and antitrust (see discussion later in this chapter) has not carried over to the realm of contracts. In an attempt to discern the influence of economic analysis on contract law, Eric Posner (2002, p. 35) recently reviewed all state and federal court opinions in contract disputes that cited an economics article appearing in either a major law review or a faculty-edited journal since 1980. He found that only twenty-three such opinions were issued, four of which discussed the article and the rest of which only cited it. He also found that the notes to the Restatement (2nd) of Contracts contained only a handful of references to economic ideas.[78] It seems safe to conclude that, while the economic analysis of contract law remains a major feature in the law and economics academy, its impact on the law remains fairly negligible.

Antitrust Policy

It has been said that, of all the fields of law, Chicago law and economics has had its greatest impact on antitrust.[79] Antitrust policy had gone through a remarkable evolution since its inception in the late 1800s.[80] The earliest cases (in the late 1800s to early 1900s) relied primarily on common law rules. Throughout most of the first half of the twentieth century, antitrust was a blend

---

[77] On the economics of liquidated damages, see, for example, Goetz and Scott (1977), Cooter and Ulen (2004, pp. 251–54), and Polinsky (1989, pp. 63–65).

[78] Beecher-Monas (2002) argues that a number of the influential judges who espouse the tenets of law and economics and rational choice theory in contract law are wrong-headed, and he suggests that their analysis of contract doctrine needs to be amended in light of the findings of behavioral economics and evolutionary game theory.

[79] Thus, it is not all unusual to read comments such as, "Since the 1980s, the Chicago School model of antitrust has reigned as the predominant approach of both the courts and the agencies" (Jacobs 1995, p. 219). Duxbury (1995, p. 349) observed that "[T]here exists very little in the way of contemporary antitrust theory which has not been inspired to some degree by Chicago economic analysis." Lande contends, "[T]he dominant paradigm today is that the only goal of the existing antitrust laws is to increase economic efficiency" (1990, p. 258).

[80] This brief history is retold in many places; we have relied on Hovenkamp (1985).

of statutory law and the judge-made law of the U.S. Supreme Court and Appellate Courts. In the early part of the twentieth century, the courts adopted the "rule of reason" approach; later, in the mid-1900s, Edward Chamberlin and Joan Robinson's economic theory of monopolistic competition served as a partial basis for deciding cases. This was followed by J. M. Clark's theory of workable competition, which was later endorsed by the Attorney General's National Committee to Study the Antitrust Laws. In the 1950s and 1960s, antitrust policy was driven by the so-called liberal school, as reflected in the decisions of the Warren Court. While some suggest that the inclusion of economic thinking into antitrust policy is a relatively recent phenomenon, the very economic nature of the problem means that the formulation of antitrust policy has, to various degrees, always incorporated some form of economic thinking. Indeed, as Hovenkamp (1985, p. 217) has pointed out, "Only an extreme form of historical myopia will admit . . . [that] economic theory had nothing useful to say about antitrust policy until the 1970s."

THE SETTING FOR CHICAGO ANTITRUST

To appreciate fully the contribution of the Chicago school to antitrust analysis, it is important to understand the economic and legal milieu from which it arose. In economics, antitrust has long been a part of the field of industrial organization, while in law it is the legacy of the New Deal–liberal school, where the prevention of industrial concentration was pursued for political and social reasons.

Early industrial organization theory argued that the way to understand industrial policy was to utilize the *structure-conduct-performance paradigm*. In the simplest of terms, this approach stated that the structure of an industry (number of firms, industrial concentration, technology, etc.) would affect the behavior of market participants (that is, the conduct of individual firms with respect to pricing decisions, levels of advertising, investment in research and development, etc.), and consequently, market performance (taking into consideration both the levels of profits and social welfare).[81] Some who utilized this paradigm argued that to attain the desired economic performance, antitrust policy should concern itself with the "structure" of the industry under question. These were the so-called "structuralists."[82] The structuralists tended to view large firms and

---

[81] A typical presentation of the structure-conduct-performance approach can be found in Scherer (1970, pp.1–7). For a concise review of this paradigm and its relation to antitrust law, see Carstensen (1983). Some proponents of the Chicago school often suggest this approach is of little value, for example: "Today it is hard to find an economist who believes the old structure-conduct-performance paradigm" (Easterbrook 1986, p. 698). An interesting recent application of the structure-conduct-performance paradigm, however, is contained in Michael R. Baye's *Managerial Economics and Business Strategy* (2002), now in its fourth edition.

[82] Baye (1997 [2nd ed.], p. 262) terms this approach "the causal view," where it is taken to mean that the market structure "causes" firms to behave in a certain way. The distinction between the structuralists and the behaviorists is described by Scherer (1970, pp. 6–7).

mergers with suspicion and explained a variety of business practices as being motivated by monopolistic intent. The antitrust remedies they offered typically involved changing the structure of the industry through legislative, judicial, and/or administrative actions—actions that would deconcentrate the industry.[83] It is with this approach that we typically associate the term "trust-busters."

There were others who countered that emphasis on the structure of an industry was misleading, and that the concern should be with the "behavior" of market participants, regardless of market concentration. These were the so-called "behaviorists."[84] They deemphasized structure and instead probed structure-conduct and conduct-performance associations, thereby emphasizing the intermediate behavioral links. In doing so, they were able to set out an antitrust policy agenda that drew attention away from market structure and, by implication, from firm size and concentration, to focus instead on pricing behavior and economic rationales for horizontal or vertical integrations and merger activity. In so doing, they helped to create a perspective on antitrust policy that proved conducive to the Chicago approach.

The distinction between the structuralists and the behaviorists is not purely academic. One of the arguments for aggressive enforcement of the antitrust laws—pre-Chicago school—was partly based on the emphasis on structure and the associated relationship said to exist between high concentration in market sectors and resultant lessening of competition. Merge this economic thinking with the New Deal–liberal school concerns regarding (i) the political ramifications of bigness of firms and of fewness of competitors and (ii) the socioeconomic impacts of high concentration on the small businessperson and the consumer, and one can then come to appreciate more fully the legal-economic environment surrounding the rise of the Chicago school.[85] This economic structure–socio-political approach to antitrust policy came to be known as the New Coalition[86] and served as the basis for a variety of legislative and

[83] A sampling of the scholarship underlying this general approach includes the work of Mason (1937, 1957), Kaysen and Turner (1959), and Bain (1956).

[84] Baye (1997, pp. 262–63) characterizes the behaviorists as part of the so-called "feedback critique."

[85] Hovenkamp (1985, pp. 218, 219) correctly observed the following:

Despite all that has been said about the lack of sophistication or even the hostility toward economics manifested by Warren Court and Eisenhower administration antitrust policy, that policy was in fact very much informed by academic economists. The price theory and industrial organization that dominated the academic study of economics in the 1960s *were simply quite different* from the dominant economic ideology of the 1980s. . . . [Antitrust policy reflected] the academic thinking of the 1960s, in which product differentiation, industrial concentration, barriers to entry, and large firm dominance *rather than* tacit collusion were the principal areas of economic concern for the competitive process. (emphasis added)

[86] This term is employed by Fox (1987a, p. 917). White (1992, pp. 1055–57) labeled them the "Modern Populist School." In fact, there is no single broadly accepted moniker for this approach—that said, we stay with the "New Coalition."

administrative initiatives to reduce concentration levels in selected industries. Eleanor M. Fox and Lawrence A. Sullivan's characterization of the "two central concerns" of antitrust in this period makes all of this very clear:

> The first was political—distrust of bigness and of fewness of competitors as well as a policy preference for diversity and opportunity for the unestablished. The second was socioeconomic, especially as seen from the vantage point of the small businessperson and the consumer. Antitrust set fair rules for the competitive game. What mattered was getting a fair shot as an entrepreneur, and having choice and receiving a fair deal as a consumer. Antitrust was not a tool for increasing aggregate national wealth (sometimes called or equated with allocative efficiency). While a more efficient allocation of resources would probably result from competition as compared with more direct government intervention or blatant laissez-faire, improved resource allocation was never a norm for antitrust, nor a condition precedent to antitrust enforcement. In order to get to the jury, neither the government nor a private plaintiff was expected to show that a particular enforcement action would achieve efficiency. What they had to show was that the competitive process was being harmed. (Fox and Sullivan 1987b p. 944)

This approach to antitrust wove its way into policy and ultimately into the courts; in fact, the U.S. Supreme Court incorporated these very ideas into their case opinions. For example, we read from *United States v. Von's Grocery Co.* (384 U.S. 270, 274–275 [1966], at note 7), quoting Judge Hand, who was reviewing the policy of the antitrust laws and other laws designed to foster small business in *United States v. Aluminum Co. of America* (148 F.2d 416 (1945)):

> Throughout the history of these statutes it has been constantly assumed that one of their purposes was to perpetuate and preserve, for its own sake and in spite of possible cost, an organization of industry in small units which can effectively compete with each other.

More to the point, from *Northern Pacific Railway Co. v. United States* (356 U.S. 1, 4 (1958)), the court wrote:

> The Sherman Act was designed to be a comprehensive charter of economic liberty aimed at preserving free and unfettered competition as the rule of trade. It rests on the premise that the unrestrained interaction of competitive forces will yield the best allocation of our economic resources, the lowest prices, the highest quality and the greatest material progress, while at the same time providing an environment conducive to the preservation of our democratic political and social institutions.

While this is representative of much of the pre-Chicago thinking about antitrust policy in the U.S. Justice Department, the Federal Trade Commission, and the U.S. courts, a different approach had begun to evolve at the University

of Chicago beginning in the 1940s. This early, pre-Chicago work exploring the interaction between economics and the law has been labeled "old law and economics"—a label given to those fields of law that were, by their very nature, inherently concerned with both economics and the law (for example, public utility regulation, corporate law, federal taxation, labor law, and, of course, antitrust).[87] Thus, the interest of Director and others at Chicago in the area of antitrust, described in chapter 1, is not novel; the novelty, rather, lies in the challenge that this new "Chicago view" posed to the traditional structure-conduct-performance paradigm and the policies derived therefrom. It is safe to say that, as a result of the efforts of Director and his disciples, by the 1980s, antitrust academia, the antitrust bar, and the federal judiciary were becoming populated with people who had made serious efforts to learn about price theory and industrial organization (Hovenkamp 1985, p. 216).

THE CORE THEORY OF CHICAGO ANTITRUST

The Chicago approach to antitrust is predicated on four propositions.[88] First, neoclassical price theory (microeconomics) is the most useful social science to explain and understand antitrust issues, and hence should be the science used in formulating antitrust policy. The idea is that linking antitrust policy to the Chicago price-theoretic approach would largely remove antitrust policy from politics and set it on a scientific course.[89] Proponents of the Chicago approach argue that the concept of efficiency explains many of the phenomena that the New Coalition–liberal school described as anticompetitive or exclusionary.

Second, efficiency should be the sole goal of antitrust policy; wealth distribution is not an economic concern.[90] The use of efficiency provides the enforcement authorities with simple quantitative tests by defining anticompetitive behavior solely in terms of price and output. Relying exclusively on the efficiency criterion frees courts from the need to hazard qualitative judgments about the political ramifications and the socioeconomic impacts of high

---

[87] Whereas the "new law and economics"—the application of economic theory (primarily microeconomics and the basic concepts of welfare economics) to examine the formation, structure, processes, and the economic impact of law and legal institutions—begins in 1960 with the work of Coase and Calabresi.

[88] In describing the contours of the core theory of the Chicago school of antitrust, we have, owing to their clear exposition, relied heavily on two sources—Hovenkamp (1985) and Jacobs (1995); our description is a synthesis of portions of these two works. Three of the more significant books outlining the Chicago approach to antitrust are Bowman (1973), Posner (2001a), and Bork (1993); see also Gerhart (1982).

[89] The argument that economics is indeed a science and one that can (and should) be used as the basis for antitrust policy is set forth in Bork (1993, p. 8).

[90] See Hovenkamp (1985, p. 229).

concentration on the small businessperson and the consumer, judgments typically invoked by advocates of the New Coalition–liberal school.[91]

Third, the legislative history of the antitrust laws allows the judiciary to decide antitrust cases in a manner consistent with the goal of efficiency. The Chicago school also asserts, however, that judges are ill-equipped to identify and understand the true sources of a broad array of market imperfections. Consequently, they argue that judicial enforcement should proceed cautiously and with greater reliance on the market (as opposed to the courts) as the regulator; that is, their contention is that the market punishes inefficiency faster than the machinery of the law.[92]

Finally, the simplicity and clarity of the Chicago school approach to antitrust would enable firms to predict more accurately the legal consequences of important business practices, and would thereby promote capital investment and generally facilitate private ordering.[93]

With this as its foundation, the Chicago school constructed a relatively simple analytical framework from which to fashion antitrust policy. This policy holds, contrary to the New Coalition–liberal school, that antitrust enforcement should punish only inefficient conduct. The contention is that

> the focus of the antitrust laws . . . should instead be on: (1) cartels and (2) horizontal mergers large enough either to create monopoly directly, as in the classic trust cases, or to facilitate cartelization by drastically reducing the number of significant sellers in the market. . . . [T]his implied a breathtaking contraction in the scope of antitrust policy. (Posner 1978, p. 928)

More specifically the Chicago approach contends that markets are competitive and tend toward efficiency, even if they contain a relatively small number of sellers. Market imperfections are typically thought to be transitory. Thus, if a monopoly does exist, the problem tends to be self-correcting in that the higher profits of the monopolist would attract new firms into the market and thereby erode its monopoly position. Closely associated with their position on monopoly is the belief that barriers to entry (apart from those erected by government) tend to be more imagined that real and that this allows resources to flow into any market where there is an above-normal rate of return. They also argue that economies of scale are far more prevalent than suggested by the New Coalition–liberal school, as a result of which many more industries than commonly thought should be allowed to operate at fairly high concentration levels.

---

[91] As Fox (1987b, p. 945) described it, "In [the Chicago school's] intellectual universe, antitrust is embodied in a reductionist paradigm: antitrust concerns the functioning of markets; microeconomics is the study of the functioning of markets; therefore, antitrust is microeconomics. The potential and desired effect of markets is the efficient allocation of resources; therefore, the sole purpose of antitrust is to prevent inefficient allocation of resources."

[92] This is one of five "beliefs" of the Chicago school as described by Fox (1987b, p. 957).

[93] On this point, see Jacobs (1995, p. 231) and Easterbrook (1984a, p. 14).

For the Chicago school, given that producers strive to maximize profits, almost all competitively ambiguous business behavior is explained as a drive for greater efficiency. So-called exclusionary practices are generally understood by firms not to be in their long-run interests, and thus, if such practices do exist in particular market settings, they must be more efficient than alternative legal arrangements. As Jacobs (1995, p. 229) observed, "This framework . . . permits courts to interpret commercial behavior coherently, efficiently, and free of political bias. Since price theory views firms as profit-maximizers, it regards conspiracies as inherently unstable, monopoly markets as self-correcting, and entry barriers—except for those imposed by government regulation—as inadequate to prevent the flow of capital to profitable markets."

In all this, the Chicago school of antitrust evidences a strong preference for a market as free as possible from state interference in all of its manifestations. Unless a powerful reason exists for interfering, the state should avoid doing so inasmuch as each intervention may cause other unintended economic distortions and redistribute wealth, which was not part of its intention. On this point the Chicago school is unambiguous—it rejects wealth redistribution through antitrust enforcement as an appropriate role for the state and attempts to restore the state to the position of neutral umpire, which it held in the classical model.

The Chicago school's ascendancy to its dominant position in antitrust in the 1980s was greatly aided by two political factors. First, there was a shift to the right on the U.S. Supreme Court in the 1970s, as the retirements of Chief Justice Warren, Justices Black, Harlan, Fortas, and Douglas led to the appointment, by Presidents Nixon and Ford, of Harry A. Blackmun, Warren E. Burger, William H. Rehnquist, Lewis F. Powell, Jr., and John Paul Stevens.[94] Then, upon his election in 1980, President Ronald Regan appointed William Baxter as his first Assistant Attorney General to take charge of the Antitrust Division, and James C. Miller III as head of the Federal Trade Commission.[95] These two appointments were fundamental in assisting President Reagan in implementing his promise to reduce government's size, scope, and role in business.

The judicial, political, and administrative environments could not have been more accommodating to Chicago school ideas. Baxter was unequivocal in setting forth the standards that he would apply as the Justice Department's chief antitrust enforcer; the same can be said for Miller at the FTC. As characterized by Stoll and Goldfein (1999, p. 3), under Baxter,

> Criminal enforcement proceedings would be limited to cases presenting conduct proscribed by the rule of per se illegality, the use of which would be appropriate

[94] Of these, Blackman and Stevens are the two least Chicago-oriented.

[95] Miller is associated with "the Virginia school" of public choice analysis, which shares certain philosophical commonalities with the Chicago school and is discussed in chapter 3.

only in those instances involving cartel-type behavior. Consistent with his consumer welfare standard, civil actions challenging horizontal and vertical restraints would be limited to situations where there was a potential for substantial output restrictions. In this vein, he turned his back on enforcement of the Robinson-Patman Act because there was no threat of higher consumer prices [nor were there] reductions in output. Regarding enforcement of §2 of the Sherman Act, Baxter's view was that as long as a firm was pursuing some legitimate competitive goal the firm's conduct should not be challenged, even if it held a dominant market position.

This became known as the "Baxterization" of the Antitrust Division.[96]

Furthermore, by the late 1980s, several opinions of the United States Supreme Court strongly suggested that antitrust authorities and the courts had come a long way in adopting the ideas being advocated by members of the Chicago school.[97] The Chicago school also had an immediate impact with its influence on the revisions to the Department of Justice Merger Guidelines, to which we now turn.

ANTITRUST MERGER GUIDELINES INCORPORATE THE CHICAGO APPROACH

Shortly after becoming assistant attorney general in charge of the Antitrust Division, Baxter announced that the division would undertake to review and revise its merger guidelines. The new guidelines were to replace an earlier set that had been issued by the Justice Department in 1968 under then Assistant Attorney General Donald Turner, the first Ph.D. economist to lead the antitrust division and architect of the 1968 guidelines. Early indications were that the division believed that some loosening of the 1968 guidelines was appropriate. On June 14, 1982, the Antitrust Division issued its long-awaited Merger Guidelines and the Federal Trade Commission (FTC) issued its own "Statement Concerning Horizontal Mergers." The merger guidelines described the criteria and process by which the Antitrust Division would analyze the competitive impact of proposed mergers and, consequently, the basis upon which it would decide whether to challenge them under the antitrust laws. The FTC simultaneously issued a statement on horizontal mergers, announcing that it would give "considerable weight" to the Justice department's guidelines in making its own enforcement decisions.

The 1982 Merger Guidelines were considered a "revolutionary leap forward." They were drafted in the light of the then-recent Supreme Court merger decisions, as well as economic studies of antitrust issues—many of which emanated from the Chicago school. The new guidelines were intended to evidence a much greater economic logic, relying primarily on the Chicago notions of

---

[96] This moniker is largely attributed to Victor H. Kramer (1981).

[97] Two U.S. Supreme Court cases that illustrate the impact of the Chicago school on judicial thinking are *Continental T. V., Inc. v. GTE Sylvania Inc.* 433 U.S. 36 (1977) and *Matsushita Electronic Industrial Co., Ltd. v. Zenith Radio Corp*. 475 U.S. 574 (1986).

economic price theory, largely to the exclusion of legislative concerns that re-
sulted in the passage of the antitrust laws (Stoll and Goldfein 2002, p. 3).[98] The
guidelines began with an explicit statement of their underlying theme:

> Mergers should not be permitted to create or enhance "market power" or to facil-
> itate its exercise. A sole seller (or "monopolist") of a product with no good substi-
> tutes can maintain a selling price that is above the level that would prevail if the
> market were competitive. Where only a few firms account for most of the sales of
> a product, those firms can in some circumstances coordinate, explicitly or implic-
> itly, their actions in order to approximate those of the monopolist. . . . Although
> they sometimes harm competition, mergers generally play an important role in a
> free enterprise economy. . . . While challenging competitively harmful mergers,
> the Department seeks to avoid unnecessary interference with that larger universe
> of mergers that are either competitively beneficial or neutral. (Blumenthal 2000,
> p. 11)

The stated theme explains the guidelines' emphasis on horizontal mergers and
the abandonment of approaches to vertical and conglomerate mergers es-
poused by the 1968 guidelines.

The new guidelines consisted of five basic elements. First, markets, each
consisting of a group of products and a geographic area, would be defined for
each product of each of the merging firms by analyzing the ability of buyers
to substitute to other products and/or to substitute to the same product pro-
duced in other areas. That is, a market would consist of a group of products
and an associated geographic area such that (in the absence of new entry) a
hypothetical, unregulated firm that made all the sales of those products in that

---

[98] It should be noted here that some questioned the basis upon which Baxter set forth the 1982
(and thereafter, the 1984) Department of Justice Merger Guidelines. As background, Baxter and
Bork were members of the Neal Committee of 1968, the committee that came up with recommen-
dations for legislation to break up so-called concentrated industries (quite consistent with the pre-
Chicago, New Coalition thinking). Bork consistently dissented from this effort, describing it as ille-
gitimate from its inception and specious from the outset (see Rowe—in panel discussion—1985,
p. 31). Baxter, on the other hand, went along with what can only be termed a policy of large-scale
industrial deconcentration, based on the premise of the report that rested on oligopoly theory. For
Bork and other members of the Chicago school who rejected the traditional theory of oligopoly of
the day, the merger guidelines promulgated in 1968, 1982, and then 1984 were deemed to be with-
out foundation, because their explicit premise rested on the economic oligopoly model—a model
that provided the only coherent rationale undergirding the notion of high concentration and market
share levels and one they thought to be legally suspect. Baxter's response to this was simply that
the position he took "*does* rest on an oligopoly theory, but it is not *that* oligopoly theory. It is a the-
ory that says firms are always tempted to fix prices, to raise prices if they can. Usually, it's very,
very difficult; they may try but they can't coordinate, it falls apart. But, if a certain set of conditions
is met, it becomes more nearly feasible; if products tend to be homogeneous and the number of
firms tend to be small. The criteria that show up in the Merger Guidelines represent a sort of pro-
phylactic approach to conspiratorial collusion" (Baxter—in panel discussion—1985, p. 34).

area could increase its profits through a small but significant and nontransitory increase in price above prevailing or likely future price levels. Under the new guidelines, a market would be considered an analytical construct consisting of a product and an area, not a group of sellers or buyers or both. Thus, the market definition issue would turn on the question, "If the merging firms perfectly coordinated their actions, could all present and potential sellers of the product in the area profitably raise prices?" The premise here is that if, absent regulation, the sole seller of a product in an area could not raise prices profitably above prevailing or likely future levels, then the merger of two sellers of that product in that area could not enhance market power meaningfully.

Second, the guidelines include as "competitors" in a market not only present sellers of the relevant product in the relevant area but also sellers of other products who could quickly and economically, in response to a small but significant and nontransitory increase in price, begin to sell the relevant product in the relevant area using existing facilities.

Third, market shares would be assigned to each of the present competitors in the market according to particular geographic area based on consumption in the area as well as some consideration for production or capacity to produce in that particular geographic area.

Fourth, while the size distribution of firms is to be examined, the primary analytical procedure set out in the 1982 guidelines is somewhat different than the previous procedure, although the two basic ingredients—concentration and market shares—remain unchanged. Concentration would be measured not by reference to traditional four-firm concentration ratios, but instead on the Herfindahl-Hirschman Index (HHI). Under the new guidelines, there would be three sets of rules for gauging whether the Justice Department would challenge a particular merger, and these rules would be based on different levels of market concentration as determined by the HHI, which would result after the proposed transaction is completed. In essence, the shares of the merging firms were to be examined to determine whether the merger would significantly increase the likelihood of collusion. The guidelines also list many other factors that affect the ease and profitability of collusion, including the nature of the product and its manner of trade, the availability of information about competitors, and the existence of substitutes not included in the market.

Fifth, entry conditions are examined to determine whether collusion by present competitors would be profitable. It was clear that the premise of the guidelines was that high concentration poses little threat to competitive market performance if entry is very easy. The guidelines indicate that the department is unlikely to challenge mergers in markets into which entry is so easy that incumbents would be unable profitably to raise price to a significant extent and for a significant period of time.

The significance of these revisions to the Merger Guidelines is profound, as reflected in Stoll and Goldfein's (2002, p. 3) assertion:

> No policy document issued by the antitrust agencies has been more enduring or far-reaching. . . . Its analysis is as dynamic and vital as the economy to which it is applied and it must be acknowledged that anything less may have unnecessarily impeded the efficiency of the economy, restricted the efforts of American business to compete internationally and, thus reduced the well-being of American consumers.

The new guidelines established the presumption that the "larger universe of mergers [is] either competitively beneficial or neutral" (Stoll and Goldfein 2002, p.3).

The impact on merger activity during the Reagan administration was significant. In 1980, the reported value of mergers was $33 billion; in 1988, $266 billion, an increase of over 700 percent—it turned out to be one of the greatest merger movements in our history. One cannot say definitively that there was a direct causal connection between Chicago-school thinking and the merger wave of the late 1970s through the mid-1980s. A large part of it was conglomerate in nature and, truth be told, no school had a good theory to deal with such mergers. Nonetheless, it is safe to say that the pervasive influence of the Chicago school—with its view that mergers are typically efficiency-enhancing and beneficial to consumers—certainly provided an atmosphere conducive to, if not a "green light" for, those seeking mergers. In the final analysis, the concepts underlying the 1982/1984 guidelines are now so well accepted that even non-Chicago enforcers of antitrust use them. As written and enforced, they do not curtail corporations from engaging in mergers on a scale unthought of thirty years ago.

In summary, during the late 1970s and early 1980s, largely due to what was transpiring at Chicago, economic theory began to play an ever-increasing role in antitrust thinking generally and in merger enforcement in particular. The Chicago law and economics movement posited that the antitrust laws were essentially economic in their underpinnings and that their primary objective should be to increase efficiency and thereby reduce consumer welfare loss. This perception of the objectives of the antitrust laws caused a reformulation of the antitrust enforcement policy, and the use of economic justification has largely supplanted all other considerations as the basis of antitrust enforcement in the merger area. The concepts of increased concentration of resources and trends toward such increases, and their purported impact on economic performance, which for so long acted as the catalyst for merger enforcement, no longer trigger government challenge (Stoll and Goldfein 2002, p. 3).

That having been said, the evolution of antitrust is by no means complete. Stringent opposition to the Chicago school still exists and has had some influence on recent litigation. Yes, it is now generally accepted that economics is

essential to a proper fashioning of antitrust policy, but the Chicago approach is not the final word, according to many. It is commonly recognized that monopolization is not something that can be ignored; that exclusive dealing and predatory pricing can still be a problem; that the tying problem persists; and that we do need to worry about exclusionary practices (in part because of information asymmetries). Finally, serious concerns over policies affecting innovation and mergers still present challenges to antitrust policymakers. In all this, it is clear that the Chicago school has been at work in the field of antitrust law, in and through both the courts and the Antitrust Division. Perhaps Hovenkamp (1985, p. 283) sums it up best when he states, "The Chicago School of antitrust analysis has made an important and lasting contribution to antitrust policy. The School has placed an emphasis on economic analysis in antitrust that will likely never disappear."