ECONOMIC ANALYSIS OF LAW

SIXTH EDITION

RICHARD A. POSNER
Judge, U.S. Court of Appeals for the Seventh Circuit;
Senior Lecturer, University of Chicago Law School

ASPEN
PUBLISHERS

1185 Avenue of the Americas, New York, NY 10036
www.aspenpublishers.com

CHAPTER 1

THE NATURE OF ECONOMIC REASONING

This book is written in the conviction that economics is a powerful tool for analyzing a vast range of legal questions but that most lawyers and law students—even very bright ones—have difficulty connecting economic principles to concrete legal problems. A student takes a course in price theory and learns what happens to the price of wheat when the price of corn falls and to the price of grazing land when the price of beef rises but does not understand what these things have to do with free speech or accidents or crime or the Rule Against Perpetuities or corporate indentures. This book's design is to anchor discussion of economic theory in concrete, numerous, and varied legal questions; the discussion of economic theory in the abstract is confined to this chapter.

§1.1 Fundamental Concepts

Many lawyers still think that economics is the study of inflation, unemployment, business cycles, and other mysterious macroeconomic phenomena remote from the day-to-day concerns of the legal system. Actually the domain of economics is much broader. As conceived in this book, economics is the science of rational choice in a world—our world—in which resources are limited in relation to human wants.[1] The task of economics, so defined, is to explore the implications of assuming that man[2] is a rational maximizer of his ends in life, his satisfactions—what we shall call his "self-interest." Rational maximization should not be confused with conscious calculation. Economics is not a theory about consciousness. Behavior is rational when it conforms to the model of rational choice, whatever the state of mind of the chooser (see §1.3 *infra* on the meaning of "rational" in economics). And self-interest should not be confused with selfishness; the happiness (or for that matter the misery) of other people may be a part of one's satisfactions. To avoid this confusion, economists prefer to speak of "utility" (discussed further in the next section of this chapter) rather than of self-interest.

§1.1 1. See Gary S. Becker, The Economic Approach to Human Behavior (1976), and for criticism of so broad a definition of economics, Ronald H. Coase, Economics and Contiguous Disciplines, 7 J. Leg. Stud. 201 (1978).

2. And woman too, of course. Throughout this book, the "masculine" pronouns are used in a generic rather than a gendered sense. The book devotes more space to issues of particular concern to women (see, e.g., Chapter 5) than is typical in economic analyses of law.

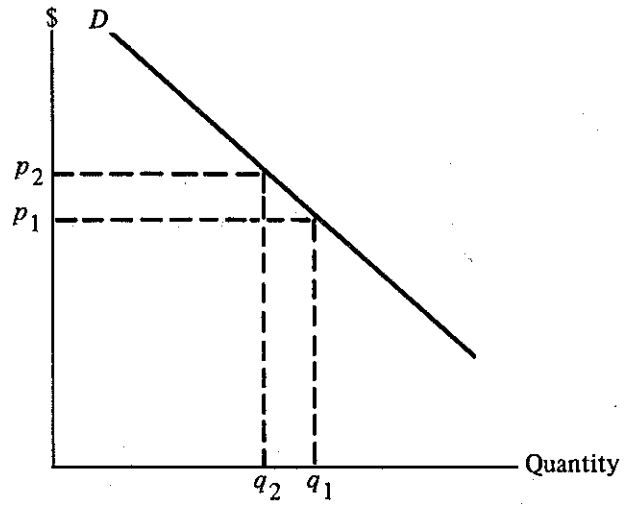

Figure 1.1

Central to this book is the further assumption that man is a rational utility maximizer in *all* areas of life, not just in his "economic" affairs, that is, not only when engaged in buying and selling in explicit markets. This is an idea that goes back to Jeremy Bentham in the eighteenth century, but it received little attention from economists until the work of Gary Becker in the 1950s and 1960s.[3]

The concept of man as a rational maximizer of his self-interest implies that people respond to incentives—that if a person's surroundings change in such a way that he could increase his satisfactions by altering his behavior, he will do so. From this proposition derive the three fundamental principles of economics.

The first is the inverse relation between price charged and quantity demanded (the Law of Demand). If the price of steak rises by 10¢ a pound, and if other prices remain unchanged, a steak will now cost the consumer more, relatively, than it did before. Being rational and self-interested, the consumer will react by investigating the possibility of substituting goods that he preferred less when steak was at its old price but that are more attractive now because they are cheaper relative to steak. Many consumers will continue to buy as much steak as before; for them, other goods are poor substitutes even at somewhat lower relative prices. But some purchasers will reduce their purchases of steak and substitute other meats (or other foods, or different products altogether), with the result that the total quantity demanded by purchasers, and hence the amount produced, will decline. This is shown in Figure 1.1. Dollars are plotted on the vertical axis, units of output on the horizontal. A rise in price from p_1 to p_2 results in a reduction in the quantity demanded from q_1 to q_2. Equally, we could imagine quantity supplied falling from q_1 to q_2 and observe that the effect was to raise the price of the good from p_1 to p_2. Can you see why the causality runs in both directions?

This analysis assumes that the only change occurring in the system is the change in relative[4] price or in quantity. Yet if, for example, demand were increasing at the

3. See Becker, note 1 *supra*, ch. 1 (1976); Richard A. Posner, Frontiers of Legal Theory 54-57 (2001).

4. If the price *level* is rising for all goods (i.e., if there is inflation), there is no quantity effect (why not?).

same time that price was rising, the quantity demanded and supplied might not fall; it might even rise. (Can you graph an increase in demand? If not, see Figure 9.5 in Chapter 9.)

The analysis also assumes away the possible impact of a change in relative price on incomes. Such a change might have a feedback effect on the quantity demanded. Suppose that a reduction in a person's income will cause him to buy more of some good.[5] Then an increase in the price of that good will have two immediate effects on consumers of the good: (1) Substitutes will become more attractive; (2) consumers' wealth will be reduced because the same income now buys fewer goods. The first effect reduces demand for the good, but the second (under the assumption that it is an inferior good) increases the demand for it and might conceivably, though improbably, outweigh the first.[6] The wealth effects of a change in the price of a single good are unlikely to be so great as to have more than a negligible feedback effect on demand; in other words, the substitution effects of a price change ordinarily exceed the income or wealth effects. So the latter can usually be ignored.

The Law of Demand doesn't operate just on goods with explicit prices. Unpopular teachers sometimes try to increase class enrollment by raising the average grade of the students in their courses, thereby reducing the price of the course to the student. The convicted criminal who has served his sentence is said to have "paid his debt to society," and an economist would find the metaphor apt. Punishment is, at least from the criminal's standpoint (why not from society's, unless the punishment is in the form of a fine?), the price that society charges for a criminal offense. The economist predicts that an increase in either the severity of the punishment or the likelihood of its imposition will raise the price of crime and therefore reduce its incidence. The criminal will be encouraged to substitute other activity. Economists call nonpecuniary prices "shadow prices."

The consumers in our steak example—and the criminal—were assumed to be trying to maximize their utility (happiness, pleasure, satisfactions).[7] The same is presumably true of the producers of beef, though in the case of sellers one usually speaks of profit maximization rather than of utility maximization. Sellers seek to maximize the difference between their costs and their sales revenues, but for the moment we are interested only in the lowest price that a rational self-interested seller would charge. That minimum is the price that the resources consumed in making (and selling) the seller's product would command in their next best use—the alternative price. It is what the economist means by the cost of a good, and suggests why (subject to some exceptions that need not trouble us here) a rational seller would not sell below cost. For example, the cost of making a lawn mower is the price the manufacturer must pay for the capital, labor, materials, and other resources consumed in making it. That price must exceed the price at which the resources could have been sold to the next highest bidder for them, for if the

5. This would be what economists call an "inferior" good. Technically, a good is inferior if a reduction in the consumer's income will not have a proportionately negative effect on his purchase of the good. A consumer is apt to change the composition of his diet in favor of potatoes and against caviar if his income falls, but, especially if his income falls a lot, he may not actually buy more potatoes than he did before. A "normal" good is one the demand for which is proportional to income; and a good is "superior" if a fall (rise) in income will cause a proportionately greater fall (rise) in the consumption of the good.

6. This is the Giffen paradox, but no well-authenticated real-world example of a "Giffen good" has been found.

7. We shall examine the concept of utility more critically in the next section.

manufacturer had not been willing to beat that price he would not have been the high bidder and would not have obtained the resources. We postpone the complication that is introduced when the sellers of a resource price it higher than its alternative price.

A corollary of the notion of cost as alternative price is that a cost is incurred only when someone is denied the use of a resource. Since I can breathe as much air as I want without depriving anyone of any of the air he wants, no one will pay me to relinquish my air to him; therefore air is costless.[8] So is a good with only one use. (Can you see why?) Cost to the economist is "opportunity cost"—the benefit forgone by employing a resource in a way that denies its use to someone else. Here are two more examples of opportunity cost: (1) The major cost of higher education is the forgone earnings that the student would have if he were working rather than attending school; this cost exceeds the tuition cost. (2) Suppose the labor, capital, and materials costs of a barrel of oil total is only $2, but because low-cost oil is being rapidly depleted, a barrel of oil is expected to cost $20 to produce in 10 years. The producer who holds on to his oil for that long will be able to sell it for $20 then. That $20 is an opportunity cost of selling the oil now—although not a net opportunity cost, because if the producer waits to sell his oil, he will lose the interest he would have earned by selling now and investing the proceeds. Suppose, however, that the current price of oil is only $4 a barrel, so that if the producer sells now, he will have a profit of only $2. If he invests the $2, it is unlikely to grow to $20 (minus the then cost of production) 10 years hence. So he is better off leaving the oil in the ground. The scarcer that oil is expected to be in the future, the higher the future price will be, and therefore the likelier the oil is to be left in the ground, which will have the effect of alleviating a future scarcity.

This discussion of cost may help dispel one of the most tenacious fallacies about economics—that it is about money. On the contrary, it is about resource use, money being merely a claim on resources.[9] The economist distinguishes between transactions that affect the use of resources, whether or not money changes hands, and purely pecuniary transactions—transfer payments. Housework is an economic activity, even if the houseworker is a spouse who does not receive pecuniary compensation; it involves cost—primarily the opportunity cost of the houseworker's time. Sex is an economic activity too. The search for a sexual partner (as well as the sex act itself) takes time and thus imposes a cost measured by the value of that time in its next-best use. The risk of a sexually transmitted disease or of an unwanted pregnancy is also a cost of sex—a real, though not primarily a pecuniary, cost. In contrast, the transfer by taxation of $1,000 from me to a poor (or to a rich) person would be costless *in itself*, that is, regardless of its secondary effects on his and my incentives, the (other) costs of implementing it, or any possible differences in the value of a dollar to us. It would not diminish the stock of resources. It would diminish my purchasing power, but it would increase the recipient's by the same amount. Put differently, it would be a private cost but not a social one. A social cost diminishes the wealth of society; a private cost rearranges that wealth.

8. That is not to say that *clean* air is costless, cf. §3.7 *infra*.

9. Noneconomists attach more significance to money than economists do. One of Adam Smith's great achievements in *The Wealth of Nations* was to demonstrate that mercantilism, the policy of trying to maximize a country's gold reserves, would impoverish rather than enrich the country that followed it. Other common misconceptions about economics that this book will try to dispel is that it is primarily about business or explicit markets, that it is pro-business, that it is heartless, that it slights nonquantifiable costs and benefits, and that it is inherently conservative.

Competition is a rich source of "pecuniary" as distinct from "technological" externalities—that is, of wealth transfers from, as distinct from cost impositions on, unconsenting parties. Suppose A opens a gas station opposite B's gas station and as a result siphons revenues from B. Since B's loss is A's gain, there is no diminution in overall wealth and hence no social cost, even though B is harmed by A's competition and thus incurs a private cost.

The distinction between opportunity costs and transfer payments, or in other words between economic and accounting costs, helps show that cost to an economist is a forward-looking concept. "Sunk" (incurred) costs do not affect a rational actor's decisions on price and quantity. Suppose that a life-sized porcelain white elephant cost $1,000 to build ($1,000 being the alternative price of the inputs that went into making it) but that the most anyone will pay for it now that it is built is $10. The fact that $1,000 was sunk in making it will not affect the price at which it is sold, provided the seller is rational. For if he takes the position that he must not sell it for less than it cost him to make it, the only result will be that instead of losing $990 he will lose $1,000.

This discussion of sunk costs should help explain the emphasis that economists place on the ex ante (before the fact) rather than ex post (after the fact) perspective. Rational people base their decisions on expectations of the future rather than on regrets about the past. They treat bygones as bygones. If regret is allowed to undo decisions, the ability of people to shape their destinies is impaired.[10] If a party for whom a contract to which he freely agreed turns out badly is allowed to revise the terms of the contract ex post, few contracts will be made.

The most celebrated application of the concept of opportunity cost in the economic analysis of law is the Coase Theorem.[11] The theorem, slightly oversimplified—for necessary qualifications see §3.6 *infra*—is that if transactions are costless, the initial assignment of a property right will not affect the ultimate use of the property. Suppose a farmer owns his land and ownership entitles him to prevent the destruction of the crop that he grows on the land by sparks from an adjacent railroad's locomotives. The crop is worth $100 to him. The value to the railroad of unimpeded use of its right-of-way is much higher, but at a cost of $110 it can install spark arresters that will eliminate the fire hazard and then it can run as many trains as it wants without injuring the farmer's crop. On these assumptions, the real value of the crop to the farmer is not $100 but somewhere between $100 and $110, since at any price below $110 the railroad would prefer to buy the farmer's property right than to install the spark arresters. The farmer can realize the higher value of the crop only by selling his property right to the railroad; he will do this; and as a result his land will be shifted into some fire-insensitive use, just as if the railroad had owned it. Similarly, if the railroad initially has the right to the unimpeded use of its right-of-way, but the farmer's growing a crop produces more value than the spark arresters cost, the farmer will buy the right to use his land free of spark damage and so again the land will be put to its most productive use regardless of the initial assignment of rights.

The forces of competition tend to make opportunity cost the maximum as well as minimum price. (Can you see why our farmer-railroad example is an exception

10. It is not the *emotion* of regret that is irrational, but acting on the emotion rather than letting bygones be bygones. Regret is a form of self-evaluation and is valuable in improving future conduct ("I won't do this again because I know I would regret it").

11. Ronald H. Coase, The Problem of Social Cost, 3 J. Law & Econ. 1 (1960).

to this generalization?) A price above opportunity cost is a magnet drawing resources into the production of the good until the increase in output drives price, by the Law of Demand, down to the level of cost (why will competition not drive price below opportunity cost?). This process is illustrated in Figure 1.2. *D* represents the demand schedule for the good in question and *S* the opportunity cost of supplying a unit of output at various levels of output. Another name for *S* is the industry marginal-cost curve. Marginal cost is the change in total costs brought about by a one-unit change in output; in other words, it is the cost that would be avoided by producing one unit less. (Marginal cost is explored further in Chapters 9 and 10.) This definition should help you see why the intersection of *D* and *S* is the equilibrium price and output under competition. "Equilibrium" means a stable point, that is, a point at which, unless the conditions of demand or supply change, there is no incentive for sellers to alter price or output. Why would any point to either the left or the right of the intersection represent an unstable, disequilibrium price-output level?

Even in long-run competitive equilibrium, there is no assurance that all sales will take place at prices equal to the opportunity costs of the goods sold. This is implicit in the upward slope of the supply curve in Figure 1.2. The fact that the cost of producing the good rises with the quantity produced implies that its production requires some resource that is inherently very scarce in relation to the demand, such as fertile or well-located land. Suppose, for example, that the very best corn land can produce corn at a cost of $1 a bushel, with the cost consisting both of the direct costs of producing corn (labor, fertilizer, etc.) and the value of the land in its next best use, and that the market price of the corn produced on such land would be $10 a bushel were no other corn produced. Clearly there are incentives to expand production, and since the good land cannot be expanded, inferior land will be shifted into corn production—land that requires greater inputs of labor, fertilizer, and so forth to produce the same quantity of corn. This process of reallocation will continue until price and marginal cost are equalized, as in Figure 1.2.

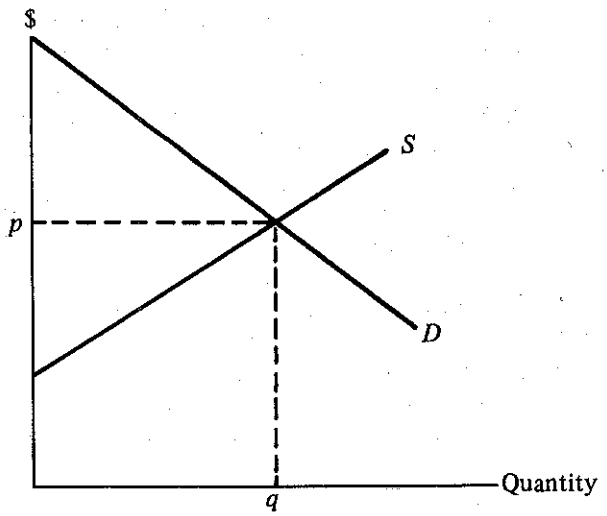

Figure 1.2

At this point, the market price will equal the cost of the marginal producer. Suppose that cost is $2.50. All corn farmers will be selling corn at $2.50 a bushel, but those with the best land will be incurring a (social) opportunity cost of only $1.

The difference between the total revenues of the industry depicted in Figure 1.2 (that is, $p \times q$) and the total opportunity costs of production (the area under S to the left of q) is called economic rent (not to be confused with rental). Rent for our purposes is a (positive) difference between total revenues and total opportunity costs. Who gets the rents in Figure 1.2? Not the producers of the corn, but the owners of the good land (of course, they may be the same people, but the roles of owner and producer are distinct). Competition between producers will eliminate any producer rents, leaving all the rents to be captured by the owners of the resource that generates them. If the quantity of ideal land could be increased without cost, competition would eliminate the scarcity that generates the rents, and with it the rents themselves. Thus under competition rents are earned only by the owners of resources that cannot be augmented rapidly and at low cost to meet an increased demand for the goods they are used to produce.

The very high incomes earned by a few singers, athletes, and lawyers include economic rents that are due to the inherent scarcity of the resources that these persons control—a fine singing voice, athletic skill and determination, the analytical and forensic skills of the successful lawyer. Their earnings may greatly exceed their highest potential earnings in an alternative occupation even if they sell their services in a fully competitive market. A different kind of economic rent, discussed in Chapter 9, is earned by the monopolist, who creates an *artificial* scarcity of his product.

Returning to the concept of an equilibrium, imagine that the government has placed a price ceiling on the good depicted in Figure 1.2, and the ceiling is below the equilibrium price (otherwise it would be ineffective), thus forcing down the dotted line p. As a result, p will now intersect the supply curve to the left of the demand curve—meaning that supply will fall short of demand. The reason is that the lower price simultaneously reduces the incentive of producers to make the good and increases the desire of consumers to buy it. The result is a shortage. How is equilibrium restored? By using a nonprice method of allocating supply to demand. For example, consumers might be required to queue up for the product; the cost of their time will determine the length of the queue. Queues are common in markets in which prices are regulated, and we shall discuss examples in this book. The removal of price regulation invariably reduces, and usually eliminates, queues—as the inhabitants of the formerly communist countries of Central and Eastern Europe have learned in recent years. (As an exercise, graph a glut caused by a price floor, and discuss its consequences.)

The third basic principle of economics is that resources tend to gravitate toward their most valuable uses if voluntary exchange—a market—is permitted. Why did the manufacturer of lawn mowers in an earlier example pay more for labor and materials than competing users of these resources? The answer is that he thought he could use them to obtain a higher price for his finished good than could competing demanders; the resources were worth more to him. Why does farmer A offer to buy B's farm at a price higher than B's minimum price for the property? Because the property is worth more to A than to B, meaning that A can use it to produce a more valuable output as measured by the prices consumers are willing to pay. By a process of voluntary exchange, resources are shifted to those uses in which the

value to consumers, as measured by their willingness to pay, is highest. When resources are being used where their value is highest, or equivalently when no reallocation would increase their value, we say they are being employed efficiently.

A methodologically useful although unrealistic assumption is that there are no unexploited profit (in the sense of rent, not cost of equity capital) opportunities. A profit opportunity is a magnet drawing resources into an activity. If the magnet doesn't work, the economist takes this as a sign not that people are dumb or have weird tastes or have ceased to be rational maximizers but that there are barriers to the free flow of resources. The barrier could be high information costs, externalities, inherent scarcities as in our rent-of-land example, or other economic conditions discussed in this book. If there are no such barriers, then in the market depicted in Figure 1.2 each seller will (as shown in Figure 1.3) confront a horizontal demand curve equal to p even though the market as a whole faces a downward-sloping demand curve (which can be viewed as the sum of a very large number of individual-firm demand curves, each of which is only trivially downward-sloping, i.e., approximately horizontal, but the aggregate of which is steeply sloped).[12] The significance of a horizontal demand curve is that if the seller raises, however slightly, his price above the market price, his sales will go to zero; for by raising his price and thereby opening a gap between price and marginal cost, he will create a profit opportunity that another seller will immediately snatch away from him.

§1.2 Value, Utility, Efficiency

The previous section bandied about some pretty highly charged words—value, utility, efficiency—about which we need to be more precise. The economic value of something is how much someone is willing to pay for it or, if he has it already, how much money he demands for parting with it. These are not always the same amounts, and this can cause difficulties, which we shall consider later.

Utility is used in two quite different senses in economics. First, it is used in analyzing the value of an uncertain cost or benefit as distinct from a certain one; utility (more precisely, "expected utility") in this sense is entwined with the concept of risk. Suppose you were asked whether you would prefer to be given $1 million or a 10 percent chance of $10 million. Probably you would prefer the former, even though the expected value of the two choices is the same: $1 million (= .10 × $10 million). Probably, then, you are risk averse. Risk aversion is a corollary of the principle of diminishing marginal utility of money, which just means that the more money you have, the less additional happiness you would get from another dollar. Diminishing marginal utility is more dramatically illustrated by less versatile commodities than money—it is easy to picture in the context, say, of chairs, or lamps,

12. It is not necessary to assume, however, that there is an infinitely large number of sellers in the market—only that entry is instantaneous if price exceeds marginal cost. This point is explained in Chapter 9. Notice in Figure 1.3 that the firm's marginal cost curve is shown upward-sloping, just like the industry's curve in Figure 1.2. The same things that cause the industry's marginal cost to rise will cause the individual firm's to do so; an additional consideration is the increasing cost of information and control as a firm grows larger and more complex. See §14.1 *infra*. Notice that if a firm did not encounter rising marginal costs at some point, its output would be indeterminate. The relationship between costs and demand is discussed more fully in Chapter 12.

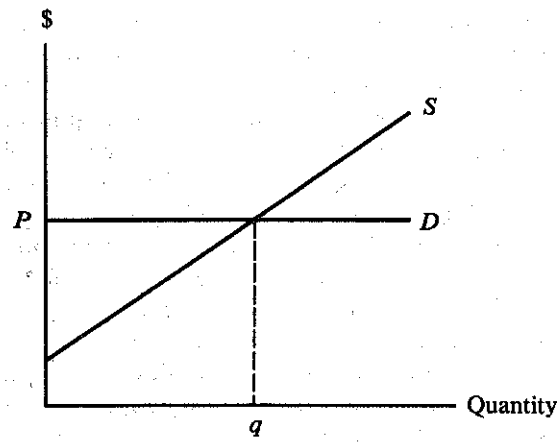

Figure 1.3

or pet gerbils. Nevertheless, it should be apparent on reflection that another dollar also will mean less to a person as his wealth increases. Suppose you had a net worth of $1 million: Would you be willing to stake it on a 50-50 bet to win $2 million? If not, it means that your first million dollars is worth more to you than a second million would be.

Risk aversion is not a universal phenomenon; gambling illustrates its opposite, risk preference (can you see why?). But economists believe, with some evidence (notably the popularity of insurance and the equity premium—the higher expected return of common stock than of bonds), that most people are risk averse most of the time, though we shall see that institutional responses to risk aversion such as insurance and the corporation may make people effectively risk neutral in many situations.

The use of the words "value" and "utility" to distinguish between (1) an *expected* cost or benefit (i.e., the cost or benefit, in dollars, multiplied by the probability that it will actually materialize) and (2) what that expected cost or benefit is worth to someone who is not risk neutral obscures a more dramatic distinction. This is the distinction between (1) value in a broad economic sense, which includes the idea that a risk-averse person "values" $1 more than a 10 percent chance of getting $10, and (2) utility in the sense used by philosophers of utilitarianism, meaning (roughly) happiness.

Suppose that pituitary extract is in very short supply relative to the demand and is therefore very expensive. A poor family has a child who will be a dwarf if he does not get some of the extract, but the family cannot afford the price and could not even if it could borrow against the child's future earnings as a person of normal height, because the present value of those earnings net of consumption is less than the price of the extract. A rich family has a child who will grow to normal height, but the extract will add a few inches more, and his parents decide to buy it for him. In the sense of value used in this book, the pituitary extract is more valuable to the rich than to the poor family, because value is measured by willingness to pay; but the extract would confer greater happiness in the hands of the poor family than in the hands of the rich one.

As this example shows, the term "efficiency," when used as in this book to denote that allocation of resources in which value is maximized, has limitations as an ethical

criterion of social decisionmaking. The concept of utility in the utilitarian sense also has grave limitations, and not only because it is difficult to measure when willingness to pay is jettisoned as a metric. First, most people don't believe—and there is no way to prove them wrong—that maximizing happiness, or contentment, or joy, or preference satisfaction, or the excess of pleasure over pain, or some other version of utility is or should be one's object in life. Happiness is important to most people, but it isn't everything. Would you be willing to take a pill that would put you into a blissfully happy dreamlike trance for the rest of your life, even if you were absolutely convinced of the safety and efficacy of the pill and the trance?

Second, by aggregating utility across persons, utilitarianism treats people as cells in the overall social organism rather than as individuals. This is the source of a number of well-known barbarisms of utilitarian ethics, such as the deliberate sacrifice of innocents to maximize the total amount of happiness in the society (or the world, or the universe); or the "utility monster," whose capacity for sadistic pleasure so far exceeds the capacity of his victims to experience pain that utility is maximized by allowing him to commit rape and murder. Defenders of utilitarianism seek to deflect such criticisms by pointing out that lack of trust in officials would defeat any effort to empower the state to attempt to maximize utility on a case by case basis. The only regime that would be utility maximizing in the real world would be a form of rule utilitarianism that limited the power of government. But practical objections to the logical implications of utilitarianism strike critics as missing the point. They regard the logic itself as repulsive. Even if all the problems of implementation are assumed away, such results as the inducement of blissful trances by utterly benign, democratically responsive officials and the sacrifice of innocents for the sake of the greater good are deeply disquieting.

Of course, it is possible that the practical objections to the logical implications of utilitarianism are what underlie the "repulsiveness" of those implications; that is, moral sentiments may have a pragmatic, conceivably an economic, basis.

But, third, utilitarianism has no boundary principles, except possibly sentience. Animals feel pain, and even more clearly do foreigners, so that utilitarianism collides with powerful intuitions that our social obligations are greater to the people of our own society than to outsiders and greater to human beings than to (other) animals.

The objections to utilitarianism and thus to tying the concept of efficiency to utilitarian ethics have turned many economists to a definition of efficiency that confines the term to outcomes of voluntary transactions. Suppose A sells a wood carving to B for $100, both parties have full information, and the transaction has no effect on anyone else. Then the allocation of resources that is brought about by the transaction is said to be Pareto superior to the allocation of resources before the transaction. A Pareto-superior transaction (or "Pareto improvement") is one that makes at least one person better off and no one worse off. (In our example, it presumably made both A and B better off, and by assumption it made no one worse off.) In other words, the criterion of Pareto superiority is unanimity of all affected persons.

Who can quarrel with unanimity as a criterion of social choice? Well, a liberal in the nineteenth-century sense—one who believes with John Stuart Mill that every person should be entitled to the maximum liberty consistent with not infringing anyone else's liberty—can quarrel with it. The problem arises when people have preferences concerning each other's consumption. Imagine a society composed of

two individuals (or two homogeneous groups, to make it a little more realistic). A, a Protestant, doesn't want B, a Catholic, to read the Catholic Bible. He would like the Catholic Bible to be banned. His second choice is that he, A, read the book, as he considers himself sufficiently inoculated against Catholic heresy. His last choice is that B read the book. B's first choice is that A be required to read the Catholic Bible—he needs it the most, being a Protestant—and his second choice is that only he himself, B, be allowed to read it. His last choice, obviously, is that the book be banned. So the only thing that A and B agree on is that it is better that A be allowed to read the book than that B be allowed to read it.[1] That is therefore the Pareto-superior choice. But it is also an illiberal choice, because it involves prohibiting B from reading a book that he wants to read.[2]

Another objection to the criterion of Pareto superiority is that it has few applications to the real world because most transactions (and if not a single transaction, then a series of like transactions) have effects on third parties, if only by changing the prices of other goods (how?). In the less austere concept of efficiency mainly used in this book—called the Kaldor-Hicks concept of efficiency, or wealth maximization—if A values the wood carving at $50 and B at $120, so that at any price between $50 and $120 the transaction creates a total benefit of $70 (at a price of $100, for example, A considers himself $50 better off and B considers himself $20 better off), it is an efficient transaction, provided that the harm (if any) done to third parties (minus any benefit to them) does not exceed $70. The transaction would not be Pareto superior unless A and B actually compensated the third parties for any harm suffered by them. The Kaldor-Hicks concept is also and suggestively called potential Pareto superiority: The winners could compensate the losers, whether or not they actually do.

The conditions for Pareto superiority are almost never satisfied in the real world, yet economists talk quite a bit about efficiency; the operating definition of efficiency in economics must not be Pareto superiority. When an economist says that free trade or competition or the control of pollution or some other policy or state of the world is efficient, nine times out of ten he means Kaldor-Hicks efficient.

The dependence of even the Pareto-superiority concept of efficiency on the distribution of wealth—willingness to pay, and hence value, being a function of that distribution—further limits efficiency as an ultimate criterion of the social good. We can illustrate with the earlier example of A's sale of a wood carving to B. A may have valued the wood carving at only $70 and B at $120 not because A likes wood carvings less than B—he may like them much more—and not because there is any appealing concept of desert to which B might appeal to validate his claim to be able to buy the wood carving. A may simply be destitute and have to sell his wood carving in order to eat, and B, while not passionate about wood carvings—while indeed, let us assume, indifferent to them—wishes to diversify his enormous wealth by holding a variety of collectibles. These circumstances (a variant of the earlier example of the market for pituitary extract) are not at all inconsistent with the sale's making both A and B better off; on the contrary, they explain *why* it makes both better off. But they undermine the moral foundations of a social system ori-

§1.2 1. If "a" means that only A reads the book, "b" that only B reads it, and "n" that neither reads it, A's order of preference is n-a-b, and B's is a-b-n; hence both agree—and agree only—that a is preferable to b.

2. See Amartya Sen, The Impossibility of a Paretian Liberal, 78 J. Pol. Econ. 152 (1970).

ented to Pareto superiority, let alone potential Pareto superiority. Even after all society's institutions have been brought into conformity with the requirements of efficiency, the pattern of consumption and production will be derivative from an underlying distribution of wealth. If that distribution is unjust, the pattern of economic activities derived from it will not have a strong claim to be regarded as just either. And insofar as the distribution of wealth is itself largely determined by the market, the justice of the market cannot be derived from some independent notion of the just distribution.

Much inequality of wealth undoubtedly reflects sheer luck, even if one's natural endowment of character and intelligence is considered an entitlement rather than a product of the random sorting of the genes. There is the luck of being born in a wealthy versus a poor country, the luck of being a beneficiary or casualty of unpredictable shifts in consumer demands and labor markets, the luck of inheritance, the luck of the financial markets, the luck of whom you know, and the luck of your parents' ability and willingness to invest in your human capital. Determinists think that it's all luck, that deservedness has nothing to do with how rich or poor anyone is. The greater the role of luck in the distribution of wealth and economic opportunities, the more difficult it is to defend the distribution as just in a strong sense (what might a defensible weak sense of "just" be in this context?).

A market system tends actually to *magnify* differences in innate ability, driving a wedge between the natural lottery and income. The cause is the "superstar" phenomenon.[3] Consider two concert pianists, one of whom (A) is slightly better than the other (B). Suppose that most of the income of a concert pianist nowadays derives not from performing or teaching but from recording. Since recordings of the same piece of music are close substitutes, a consumer has no reason to buy recordings made by B rather than those made by A unless there is a significant difference in price, and there need not be; even if A receives a higher royalty from his contract with the record company than B could command, the added cost to the record company may be offset by the economies of a larger output. A may thus end up with a very substantial income from recording and B with a zero income from it, though A may be only a 2 percent better pianist and the difference in quality may be discernible by only a small percentage of the music-loving public. There need be nothing "unjust" in this outcome; but neither can it be referred to the difference in the quality of the individuals. It illustrates, rather, the moral *arbitrariness* of many of the wealth differences among individuals.

The basic point is simply that if income and wealth were distributed differently, the pattern of demands might also be different and efficiency would require a different deployment of our economic resources. Economics does not answer the question whether the existing distribution of income and wealth is good or bad, just or unjust, although it can tell us a great deal about the costs of altering the existing distribution, as well as about the distributive consequences of various policies. Neither does it answer the ultimate question whether an efficient allocation of resources would be socially or ethically desirable. Nor can the economist tell us whether, assuming the existing distribution of income and wealth is just, consumer satisfaction should be the preeminent social value. Many philosophers and social scientists, including some economists, doubt the authenticity of many of the beliefs

3. See Sherwin Rosen, The Economics of Superstars, 71 Am. Econ. Rev. 845 (1981).

and desires that determine consumer and other individual choices. These behavioralists contend that certain cognitive quirks prevent people from behaving rationally (see §1.4 *infra*); we shall consider examples throughout the book.

For all these reasons, the economist's competence in a discussion of the legal system is limited. He can predict the effect of legal rules on value and efficiency, in their strict technical senses, and on the existing distribution of income and wealth, but he cannot issue mandatory prescriptions for social change. Yet this turns out to be less of a handicap to the normative use of economics than might appear. Economists can usually appeal to a generally accepted goal, such as maximizing the value of output, rather than having to defend the goal. By showing how a change in economic policy or arrangements would advance us toward that goal, they can make a normative statement without having to defend their fundamental premises. They can keep debate at the technical level, where reasoning is over means rather than ends. They can demonstrate for example that cartelization results in a reduction in the value of output (see Chapters 9 and 10), and since maximizing that value is a generally accepted goal of a commercial society, their demonstration provides, without more, a prima facie case for prohibiting cartels.

The qualification "prima facie" is important. Opponents of proposals for economic reform are quick to posit competing goals to that of efficiency or value maximization. This is especially the case when economists get into areas that are not traditionally economic, which happens often in economic analysis of law. To say that an area is not traditionally regarded as "economic" is to say that suggestions for orienting it toward efficiency or other economic values are likely to jar, because it is assumed that noneconomic values dominate issues that are not explicitly economic. And then what is the economist to do? Can he say more than that he has shown that policy X would increase efficiency but that he cannot speak to its ultimate merit? This is a question to be borne in mind throughout this book.

An important question, already alluded to, in the economic analysis of law is whether and in what circumstances an *involuntary* exchange can confidently be said to increase efficiency. Even if efficiency is not defined as something that only a voluntary transaction can create—even if the Kaldor-Hicks concept is used instead—willingness to pay can be confidently determined only by actually observing a voluntary transaction. Where resources are shifted pursuant to such a transaction, we can be reasonably sure that the shift involves an increase in efficiency.[4] The transaction would not have taken place unless both parties had expected to be made better off. This implies that the resources transferred are more valuable in their new owner's hands. But many of the transactions either affected or effected by the legal system are involuntary. Most crimes and accidents are involuntary transactions, and so is a legal judgment to pay damages or a fine. How is one to know when such transactions increase, and when they reduce, efficiency? If we insist that a transaction be truly voluntary before it can be said to be efficient—truly voluntary because all potential losers have been fully compensated—we shall have few occasions to make judgments of efficiency, for few transactions are voluntary in that sense. An alternative approach, used heavily in this book, is to try to guess whether, if a voluntary transaction had been feasible, it would have taken place. If, for ex-

4. We cannot be completely sure, however, because that would require that everyone affected by the transaction be a party to it, and (to repeat) this requirement is almost never satisfied.

ample, the question were whether clean water was more valuable as an input into paper production than into boating, we might try to determine, using whatever quantitative or other data might be available to help us, whether in a world of zero transaction costs the paper industry would purchase from the boaters the right to use the water.

This approach attempts to reconstruct the likely terms of a market transaction in circumstances where instead a forced exchange took place—to mimic or simulate the market, in other words. A coerced exchange, with the legal system later trying to guess whether the exchange increased or reduced efficiency, is a less efficient method of allocating resources than a market transaction—where market transactions are feasible. But often they are not, and then the choice is between a necessarily crude system of legally regulated forced exchanges and the even greater inefficiencies of forbidding all forced exchanges, which could mean all exchanges, as all have some third-party effects.

Both parties to a market transaction are compensated for entering into it; if one of them were not, the transaction would not be voluntary in even a loose sense. But when, for example, invoking the concept of "nuisance," the legal system orders a person to discontinue some use of his land on the ground that it creates less value than the decline it brings about in the value of the surrounding land, the defendant is not compensated. A legally coerced transaction is less certain to increase net happiness than a market transaction because the misery of the (uncompensated) losers may exceed the joy of the winners. And if legal efforts to simulate market results do not promote happiness, how can they be defended? What, in short, is the ethical basis of the Kaldor-Hicks concept, corresponding to the utilitarian, or preference-regarding, ethical basis of Pareto superiority? One answer is that the things that wealth makes possible—not only or mainly luxury goods, but leisure, comfort, modern medicine, and opportunities for self-expression and self-realization—are major ingredients of most people's happiness, so that wealth maximization is instrumental to utility maximization. This answer ties efficiency to utilitarianism. Answers that relate efficiency to other ethical concepts are discussed in Chapters 8 and 16.

The Kaldor-Hicks or wealth maximization approach runs into a special problem of the dependence of the efficient allocation of resources on the existing distribution of income and wealth in cases where the subject matter of the transaction is a large part of one of the parties' wealth. Suppose I refuse a $100,000 offer for my house but then the government condemns it, paying me $50,000, which is its market value. The government would happily sell the house back to me for $100,000—it is worth less than that to the government, though more than $50,000—but I neither have nor can borrow $100,000. In whose hands is the house worth more—mine or the government's? In considering this conundrum, remember that *wealth* as used by economists is not an accounting concept; it is measured by what people *would* pay for things (or demand in exchange for giving up things they possess), not by what they *do* pay for them. Thus leisure has value and is a part of wealth, even though it is not bought and sold. We can speak of leisure as having an implicit or shadow price (computed how?). Even explicit markets create value over and above the price of the goods sold in them. Go back to Figure 1.2, and notice that if the quantity sold were smaller, price would be higher; evidently consumers would be willing to pay more for some units of the product. So they must obtain value from being able to buy them at the competitive price. This value, called consumer surplus (see §9.3 *infra*), is also part of the wealth of society.

§1.3 The Realism of the Economist's Assumptions

The reader who lacks previous acquaintance with economics may be troubled by what appear to be the severely unrealistic assumptions that underlie economic theory. The basic assumption, that human behavior is rational, seems contradicted by the experiences and observations of everyday life. The contradiction is less acute when one understands that the concept of rationality used by the economist is objective rather than subjective, so that it would not be a solecism to speak of a rational frog. Rationality means little more to an economist than a disposition to choose, consciously or unconsciously, an apt means to whatever ends the chooser happens to have. In other words, rationality is the ability and inclination to use instrumental reasoning to get on in life. It does not assume consciousness; it certainly does not assume omniscience. Positive information costs are assumed throughout this book. (These costs are of two kinds—costs of acquiring information, and costs of absorbing or processing information.[1]) Even with these qualifications, the assumptions of economic theory are one-dimensional and pallid when viewed as descriptions of human behavior—especially the behavior of such unconventional economic "actors" as the judge, the litigant, the parent, the rapist, and others whom we shall encounter in the economic analysis of law. But abstraction is of the essence of scientific inquiry, and economics aspires to be scientific. Newton's law of falling bodies is unrealistic in its basic assumption that bodies fall in a vacuum, but it is still a useful theory because it predicts with reasonable accuracy the behavior of a wide variety of falling bodies in the real world. Similarly, an economic theory of law will not capture the full complexity, richness, and confusion of the phenomena—criminal or judicial or marital or whatever—that it seeks to illuminate. But its lack of realism in the sense of descriptive completeness, far from invalidating the theory, is a precondition of theory. A theory that sought faithfully to reproduce the complexity of the empirical world in its assumptions would not be a theory—an explanation—but a description.

A greater danger for positive economics in general, and the positive economic theory of law expounded in many places in this book (especially in Part II) in particular, is the opposite of reductionism: Call it complicationism. When the economic analyst seeks to make a very simple economic model more complex, for example by bringing in (as we shall do many times in this book) risk aversion and information costs, he runs the risk of finding himself with too many degrees of freedom: that is, with a model that is so flexible that no empirical observation can refute it—which means that no observation can support it, either.

All this is not to suggest that the analyst has a free choice of assumptions. An important test of a theory is its ability to explain reality. If it does a lousy job, the reason may be that its assumptions are insufficiently realistic; but we need not try to evaluate the assumptions directly in order to evaluate it. Judged by the test of explanatory power, economic theory is a significant (although only partial) success; so perhaps the assumption that people are rational maximizers of their satisfactions is not so unrealistic as the noneconomist might at first think. Economic theory can explain a vast number of market and nonmarket phenomena, such as the inverse correlation, mentioned in the first section of this chapter, between price ceilings

§1.3 1. Some economists use the term "bounded rationality" to describe the rationality of rational persons who face positive costs of using the information available to them to make decisions.

and queues; the inverse correlation between rent control and the stock of housing; the positive correlation in financial markets between risk and expected return; the relation between futures prices and spot-market prices; the dependence of college enrollment on the financial returns to a college education; the fact that the best goods tend to be shipped the farthest distances and the worst consumed at home; and many others. Much of this book is concerned with proposing economic explanations for legal phenomena modeled in economic terms.

Another test of a scientific theory is its predictive power, and here too economics has had its share of successes, most dramatically in recent years. The effects of deregulation, for example of the airline industry in the United States, and, more dramatically, of the communist economies of Central and Eastern Europe, have had the effects predicted by economists. In particular, the aftermath of communism has provided repeated confirmations of the predictions of economic analysis, such as that price regulation leads to queuing, black markets, and shortages, competition and free trade promote productivity, and private property encourages investment.

Still another test of a scientific theory is its ability to underwrite effective interventions in the world of action. The most dramatic example is the atomic bomb, which showed that modern atomic theory was not just another clever speculation about invisible entities. In this respect, too, economics has had successes, albeit more modest ones than the natural sciences have had. Economists have created new methods of pricing financial and other products, new financial trading strategies, new methods of employee and executive compensation, and new regulatory methods such as airwave auctions and salable pollution rights. These interventions have worked, suggesting that economic theory is more than just pretty math.

§1.4 Irrationality and Hyperrationality; Herein of Game Theory

The rational-choice model that underlies orthodox economic analysis has been challenged on several grounds besides the superficial one that it does not describe how people think about or describe their decisions. Another superficial objection, as we have seen, is that people lack the information to act rationally. People are not omniscient, but incompletely informed decisions are rational when the costs of acquiring more information exceed the likely benefits in being able to make a better decision. A fully informed decision in such circumstances—the sort of thing a person makes who cannot prioritize his tasks—would be irrational! Still another superficial objection to the rational-choice model is that some people are not fully rational and that all of us have lapses from rationality. Economics is concerned with explaining and predicting tendencies and aggregates rather than the behavior of each individual person; and in a reasonably large sample, random deviations from normal rational behavior will cancel out.

A more serious objection is that, perhaps because our brains evolved under different social conditions from those of modern life, human behavior exhibits *systematic* departures from rationality. Economists such as Richard Thaler,[1] as well as a

§1.4 1. See Richard H. Thaler, Quasi Rational Economics (1991).

number of psychologists, have presented evidence that most of us are prone to the "sunk costs fallacy," or throwing good money after bad. That is, in making decisions, we are unable to ignore costs that, having already been incurred, cannot be altered by the decision. They observe that we are prone to succumb to the "endowment effect"—valuing what we have more than we would value the identical thing if we didn't have it. For example, we might refuse to sell for $100 a wristwatch for which we would not pay more than $90.[2] We also engage in "hyperbolic discounting"; that is, we weight present pains and pleasures more heavily than future ones to a degree that is irrational, as when we overeat (present pleasure) knowing that we will soon regret it (future pain).[3] We also give undue weight to immediate vivid impressions relative to what we read about (the "availability heuristic").

Some of these apparent departures from rationality may be explicable in rational-choice terms. The endowment effect may be a consequence of some combination of (1) wealth effects (our preferences change when our wealth changes, and we are more or less wealthy depending on whether we own the good in question); (2) consumer surplus (see §1.2 *supra*) (a glance back at Figure 1.2 will remind the reader that everyone who owns a good, except the marginal purchaser of it, values it above the market price; so owners of the good as a class value it on average more than nonowners do); (3) rational adaptive preference—the fact that we adapt to what we have, and would incur new costs to adapt to something new.[4] A person who is blinded in an accident must incur costs to adapt to being blinded; but a blind person who through a doctor's negligence fails to regain his sight has already adapted to being blind, so his loss of (prospective) sight is less costly than the sighted person's loss of sight.

Hyperbolic discounting may be explicable by positing that the person is a succession of separate selves, "time sharing" the same body; each self is rational, but each has its own interests and they are not identical across the selves (see §6.12 *infra*). Yet, as in this example, apparent systematic departures from rationality are, at the least, a challenge to the rational-choice theorist to think more carefully about the nature of rational behavior.

Traditional economics generally assumed (except when speculating about cartel behavior, and in a few other examples) that people made decisions without considering the reactions of other people. If the price of some product falls, consumers buy more without worrying that by doing so they may cause the price to rise again. The reason they do not worry is that the effect of each consumer's decision on the

2. See, e.g., Elizabeth Hoffman & Matthew L. Spitzer, Willingness to Pay Vs. Willingness to Accept: Legal and Economic Implications, 71 Wash. U. L. Q. 59 (1993).

3. More subtly, our discount rates are excessive in relation to our mortality risk. A rational individual chooses among possible actions by using a discount rate to reduce future costs and benefits, whenever they are expected to be realized, to a present value, thus enabling comparison among the future states and between each of those states and the present. Impartiality between present and future consumption implies discounting future costs and benefits at a rate equal to the probability of still being alive when the future state in question arrives. For most people at most ages, this probability is much greater than is implicit in a discount rate of 2 to 4 percent, the usual range of estimates of the real (that is, inflation-adjusted) riskless discount rate. The present value of $1 to be received in 40 years is only 21 to 45 cents at discount rates of 2 to 4 percent. This would imply, on the assumption of impartiality between present and future consumption, that the average 30-year-old had only a 21 to 45 percent probability of living to 70. In fact that probability is 75 percent. A possible explanation for this discrepancy is that the 40-year-old's future-oriented self is unable to dominate his present-oriented self, is incompletely altruistic toward the individual's future 70-year-old self, or is both. On the individual as a locus of different selves, see text below and Richard A. Posner, Are We One Self or Multiple Selves? Implications for Law and Public Policy, 3 Leg. Theory 23 (1997).

4. For still another rationality-compatible explanation of the endowment effect, see §3.14 *infra*.

price is likely to be negligible (the consumer is a "price taker"), while the costs to the consumers of coordinating their action would be prohibitive. In some situations, however, a rational person in deciding how to act will consider the probable reactions of others; he will, in other words, act strategically. This is the domain of game theory, which is much used by current economists because of the importance of strategic behavior in many areas in which economists, including economic analysts of law, are interested.

Game theory presents a striking contrast to the challenges to the rational-choice model that we examined earlier, because it assumes, at least in its purest form, a degree of rationality even higher than that assumed in orthodox economics. Consider the following "game" (that is, strategic situation). A monopolist is faced with the prospect of the entry of another firm. If the monopolist charges a price below his (and presumably the entrant's) cost, it will deter that entry by forcing the new entrant to lose money. But the monopolist will lose a lot of money in the process of repelling entry by selling its product below cost. In fact, it may well—and let's assume it will—lose more money than it would if it maintained the monopoly price, reduced its output, and in effect divided the market with the new entrant.

Assume the market has room for only two firms. Might the monopolist reduce his price anyway, thinking: "If I charge a below-cost price, the entrant will know that I'm not bluffing, because he will realize that while I'll lose money in the short run I will be making a worthwhile investment in developing a reputation that will discourage other challengers to my monopoly position"? It turns out this may not be a rational concern for the monopolist to impute to a prospective entrant. Suppose there are 10 prospective entrants. Even if the first nine were to abandon all thoughts of entering this market because of the monopolist's threat of below-cost pricing, the tenth would not. For he would realize that when there is only one possible prospective entrant, the monopolist will be better off sharing the market than charging a price below cost. Remember that below-cost pricing made sense to the monopolist only when it was buying a reputation usable against future entrants. When there are no more future entrants—when the only possible entrant has entered—the monopolist has nothing further to gain from investing in a reputation for deterring entry, so he will not charge a below-cost price. Knowing this, the last prospective entrant will enter—but so will the first. For the first knows that the monopolist, foreseeing the collapse of his scheme when the last entrant comes in, will have no incentive to use the scheme against the ninth (for with the entry of the tenth guaranteed, there is nothing to gain from making a reputation by beating up the ninth), and hence against the eighth, and so on right down to the first. (This way of solving a game-theoretic problem, by starting with the last move in the game, is called "backward induction.")

What is striking about this example, and common in game theory, is its sensitivity to the assumption that everyone involved not only is behaving with complete rationality but also assumes, and is right to assume, that everyone else involved is behaving with complete rationality too. Orthodox economic theory does not depend on such hyperrationality. But neither do all applications of game theory. Consider this familiar example: Federal law forbids colleges to give out a student's transcript to a prospective employer or another educational institution without the student's permission. Such permission is almost never refused. Game theory can help us see why. Suppose no student gave permission, then an employer, faced with an application from a college student, would assume that the student had average grades—what else could he assume? Any student with above-average grades would

be hurt by this assumption, so these students would begin giving permission to their schools to release their transcripts. Eventually all students with grades above the midpoint would grant such permission. So now when an employer received an application from a student who had not released his transcript, the employer would assume that the student was in about the middle of the *lower half* of the grade-point distribution, because everyone in the upper half would have revealed his grades. So every student in the third quartile (that is, in the upper half of the lower half of the grade distribution) would be disadvantaged by nondisclosure and would reveal his grades. Eventually, only the student with the very lowest grades would have nothing to gain from disclosure—and his failure to disclose would reveal his rank as unerringly as if he had disclosed it. Simple game theory thus shows why the transcript-privacy law has been ineffective. The example illustrates what game theorists call a "pooling equilibrium," in which (in contrast to a "separating equilibrium") strategic behavior prevents people with different preferences from acting differently. The reasoning process required to achieve a pooling equilibrium in the student-transcript case is not so elaborate as to require hyperrationality.

Now go back to the below-cost pricing example and assume that each of the 10 potential entrants is equally capable of entering first. Each will have an incentive to hang back, knowing that the incumbent seller may have an incentive to sell below cost when the first entrant enters, to show that he "means business." It's like the case of a person who has 6 shots in his gun and is facing 10 assailants. None of the assailants may be willing to attack first, and if so there may be no attack at all, even though the attack would be certain to succeed. Actually this is a clearer case than that of below-cost pricing; the defender has nothing to lose from shooting the first 6 assailants, so anyone who attacks first *knows* that he'll be shot.

When economics is defined as the theory of rational choice, and given that game theory is the theory of *rational* strategic behavior, game theory becomes a part of economic theory. And a part with many potential applications to law, because much of law deals with strategic behavior, not only in the antitrust and student privacy examples just given, but also in contract negotiations, litigation and settlement, and many other areas. Nevertheless, this book does not attempt systematic instruction in game theory. Game theory involves complex analytical methods and, as we have begun to glimpse, a specialized vocabulary—it requires a textbook of its own.[5] But a textbook on the economics of law cannot avoid frequent encounters with strategic behavior, and when these occur we shall use simple concepts of game theory to inform the analysis and to prepare the reader for a more systematic study of this very important analytical technique.

Suggested Readings

1. Gary S. Becker, Nobel Lecture: The Economic Way of Looking at Behavior, 101 J. Pol. Econ. 395 (1993).

2. Harold Demsetz, Rationality, Evolution, and Acquisitiveness, 34 Econ. Inquiry 484 (1996).

3. Milton Friedman, The Methodology of Positive Economics, in his Essays in Positive Economics 3 (1953).

5. An excellent game-theory textbook for lawyers is Douglas G. Baird, Robert H. Gertner & Randal C. Picker, Game Theory and the Law (1994).

CHAPTER 2

THE ECONOMIC APPROACH TO LAW

§2.1 Its History

Until about 1960 economic analysis of law was almost synonymous with economic analysis of antitrust law, though there had been some economic work on tax law (Henry Simons), corporate law (Henry Manne), patent law (Arnold Plant), contract law (Robert Hale), and public utility and common carrier regulation (Ronald Coase and others). The records in antitrust cases provided a rich mine of information about business practices; and economists, who at the time were preoccupied with the question of monopoly, set about to discover the economic rationales and consequences of such practices. Their discoveries had implications for legal policy, of course, but basically what they were doing was no different from what economists traditionally had done—trying to explain the behavior of explicit economic markets.

The economic analysis of antitrust, and of other legal regulation of explicit economic markets, remains a prosperous field and receives considerable attention in this book. However, the hallmark of the "new" law and economics—the law and economics that has emerged since 1960—is the application of economics to the legal system across the board: to common law fields such as torts, contracts, restitution, and property; to the theory and practice of punishment; to civil, criminal, and administrative procedure; to the theory of legislation and regulation; to law enforcement and judicial administration; and even to constitutional law, primitive law, admiralty law, family law, and jurisprudence.

The new law and economics began with Guido Calabresi's first article on torts and Ronald Coase's article on social cost.[1] These were the first modern[2] attempts

§2.1 1. Guido Calabresi, Some Thoughts on Risk Distribution and the Law of Torts, 70 Yale L. J. 499 (1961); Ronald H. Coase, The Problem of Social Cost, 3 J. Law & Econ. 1 (1960).

2. Important work on the economics of criminal law was done in the eighteenth and early nineteenth centuries by Beccaria and Bentham—and remains well worth reading. Cesare Beccaria, On Crimes and Punishments (Henry Paolucci trans. 1963); Jeremy Bentham, An Introduction to the Principles of Morals and Legislation, in 1 Works of Jeremy Bentham 1, 81-154 (John Bowring ed. 1843); Bentham, Principles of Penal Law, in 1 id. at 365. Other precursors are discussed in Ian Ayres, Discrediting the Free Market, 66 U. Chi. L. Rev. 273 (1999); Charles K. Rowley, Law-and-Economics From the Perspective of Economics, in The New Palgrave Dictionary of Economics and the Law, vol. 2, pp. 474, 474-476 (Peter Newman ed. 1998); Barbara Fried, The Progressive Assault on Laissez Faire: Robert Hale and the First Law and Economics Movement (1998); Neil Duxbury, Robert Hale and the Economy of Legal Force, 53 Modern L. Rev. 421 (1990); Izhak Englard, Victor Mataja's Liability for Damages From an Economic Viewpoint:

to apply economic analysis systematically to areas of law that do not regulate avowedly economic relationships. One can find earlier glimmerings of an economic approach to the problems of accident and nuisance law that Calabresi and Coase discussed,[3] especially in the work of Pigou,[4] which provided a foil for Coase's analysis; but the earlier work had made little impact on legal thought.

Coase's article introduced the Coase Theorem, which we met in Chapter 1, and, more broadly, established a framework for analyzing the assignment of property rights and liability in economic terms, thus opening a vast field of legal doctrine to fruitful economic analysis.[5] An important, although for a time neglected, feature of Coase's article was its implications for the positive economic analysis of legal doctrine. Coase suggested that the English law of nuisance had an implicit economic logic. Later writers have generalized this insight and argued that many of the doctrines and institutions of the legal system are best understood and explained as efforts to promote the efficient allocation of resources—a major theme of this book.

A list of the founders of the "new" law and economics would be seriously incomplete without the name of Gary Becker. Becker's insistence on the relevance of economics to a surprising range of nonmarket behavior (including charity, love, and addiction), as well as his specific contributions to the economic analysis of crime, racial discrimination, and marriage and divorce, opened to economic analysis large areas of the legal system not reached by Calabresi's and Coase's studies of property rights and liability rules.[6]

§2.2 Positive and Normative Economic Analysis of Law

Subsequent chapters will show how the insights of the pioneers have been generalized, empirically tested, and integrated with the insights of the "old" law and economics to create an economic theory of law having explanative power and empirical support. The theory has normative as well as positive aspects. Although the economist cannot tell society whether it should seek to limit theft, he can show that it would be inefficient to allow unlimited theft and can thus clarify a value conflict by showing how much of one value—efficiency—must be sacrificed to achieve another. Or, taking a goal of limiting theft as given, the economist may be able to

---

A Centennial to an Ignored Economic Analysis of Tort, 10 Intl. Rev. Law & Econ. 173 (1990); and Herbert Hovenkamp, The First Great Law & Economics Movement, 42 Stan. L. Rev. 992 (1990).

3. See William M. Landes & Richard A. Posner, The Economic Structure of Tort Law, ch. 1 (1987), for examples.

4. A. C. Pigou, The Economics of Welfare, ch. 9 (4th ed. 1932).

5. The modern literature on property rights also reflects, however, the influence of Frank Knight's important early work, Some Fallacies in the Interpretation of Social Cost, 38 Q. J. Econ. 582 (1924); see §3.1 *infra*.

6. For the character of Becker's contributions to economics, see Gary S. Becker, The Economic Approach to Human Behavior (1976); Becker, A Treatise on the Family (enlarged ed. 1991); Becker, Accounting for Tastes (1996).

The new law and economics is now the subject of an extensive literature; besides the pertinent portions of this book, book-length treatments include Robert Cooter & Thomas Ulen, Law and Economics (2d ed. 1997); William M. Landes & Richard A. Posner, The Economic Structure of Tort Law (1987); A. Mitchell Polinsky, An Introduction to Law and Economics (2d ed. 1989); Richard A. Posner, The Economics of Justice (1981); Steven Shavell, Economic Analysis of Accident Law (1987).

show that the means by which society has attempted to attain that goal are inefficient—that society could obtain more prevention, at lower cost, by using different methods. If the more efficient methods did not impair any other values, they would be socially desirable even if efficiency were low on the totem pole of social values.

As for the positive role of economic analysis of law—the attempt to explain legal rules and outcomes as they are rather than to change them to make them better—we shall see in subsequent chapters that many areas of the law, especially but not only the great common law fields of property, torts, crimes, and contracts, bear the stamp of economic reasoning. Granted, few judicial opinions contain explicit references to economic concepts. But often the true grounds of legal decision are concealed rather than illuminated by the characteristic rhetoric of opinions. Indeed, legal education consists primarily of learning to dig beneath the rhetorical surface to find those grounds, many of which may turn out to have an economic character. (Remember how broadly economics was defined in Chapter 1.) It would not be surprising to find that many legal doctrines rest on inarticulate gropings toward efficiency. Many legal doctrines date back to the nineteenth century, when a laissez-faire ideology based on classical economics was the dominant ideology of the educated classes. And with the fall of communism there has been a strong resurgence of free-market ideology both in the United States and throughout much of the rest of the world.

What we may call the efficiency theory of the common law is not that *every* common law doctrine and decision is efficient. That would be highly unlikely, given the difficulty of the questions that the law wrestles with and the nature of judges' incentives. The theory is that the common law is best (not perfectly) explained as a system for maximizing the wealth of society. Statutory or constitutional as distinct from common law fields are less likely to promote efficiency, yet even they, as we shall see, are permeated by economic concerns and illuminated by economic analysis. Such analysis is also helpful in explaining such institutional features of the legal system as the role of precedent and the allocation of law enforcement responsibilities between private persons and public agencies.

But, it may be asked, do not the lawyer and the economist approach the same case in such different ways as to guarantee a basic incompatibility between law and economics? X is shot by a careless hunter, Y, and sues. The only question in which the parties and their lawyers are interested and the only question the judge and jury will decide is whether the cost of the injury should be shifted from X to Y, whether, that is, it is "just" or "fair" that X should receive compensation. X's lawyer will argue that it is just that X be compensated since Y was at fault and X blameless. Y's lawyer may argue that X was also careless and hence that it would be just for the loss to remain on X. Not only are justice and fairness not economic terms, but the economist is not (one might think) interested in the one question that concerns the victim and his lawyer: Who should bear the costs of *this* accident? To the economist, the accident is a closed chapter. The costs that it inflicted are sunk. The economist is interested in how to prevent future accidents that are not cost-justified and thus in reducing the sum of accident and accident-prevention costs. The parties to the litigation may have no interest in the future. Their only interest may be in the financial consequences of a past accident.

This dichotomy, however, is overstated. The decision in the case will affect the future, and so it should interest the economist, because it will establish or confirm a rule for the guidance of people engaged in dangerous activities. The decision is a warning that if one behaves in a certain way and an accident results, one will have

to pay a judgment (or will be unable to obtain a judgment, if the victim). By thus altering the shadow price (of risky behavior) that confronts people, the warning may affect their behavior and therefore accident costs.

Conversely, the judge, and hence the lawyers, cannot ignore the future. The legal ruling will be a precedent influencing the decision of future cases. The judge must therefore consider the probable impact of alternative rulings on the future behavior of people engaged in activities that give rise to the kind of accident involved in the case before him. If, for example, judgment is awarded to the defendant on the ground that he is a "deserving," albeit careless, person, the decision will encourage similar people to be careless, a type of costly behavior. Thus, once the frame of reference is expanded beyond the immediate parties to the case, justice and fairness assume broader meanings than what is just or fair as between this plaintiff and this defendant. The issue becomes what is just and fair for a *class* of activities, and it cannot be sensibly resolved without consideration of the future impact of alternative rulings on the frequency of accidents and the cost of precautions. The ex ante perspective is not alien to the legal process after all.

The "economic theory of law" and the "efficiency theory of the common law" should not be confused. The former tries to explain as many legal phenomena as possible through the use of economics. The latter (which is included in the former) hypothesizes a specific economic goal for a limited subset of legal rules, institutions, and so forth. The distinction will become clear in Chapter 11, which argues that federal labor law administered by the National Labor Relations Board, although explicable in economic terms, is not a system for maximizing efficiency; its goal, which is economic but not efficient, is to increase the incomes of union members by cartelizing the supply of labor in particular markets.

§2.3 Criticisms of the Economic Approach

Economic analysis of law has aroused considerable antagonism, and not only among academic lawyers who dislike the thought that the logic of the law might be economics. We have already examined the criticisms that economics is reductionist (a criticism not limited of course to economic analysis of law) and that lawyers and judges do not speak its language. Another common criticism is that the normative underpinnings of the economic approach are so repulsive that it is inconceivable that the legal system would embrace them. This criticism may appear to confound positive and normative analysis, but it does not. Law reflects and enforces fundamental social norms, and how could those norms be inconsistent with the society's ethical system? But is the Kaldor-Hicks concept of efficiency really so at variance with that system? Besides what was said in the first chapter, we shall see in Chapter 8 that, provided only that this concept is *a* component, though not necessarily the only or the most important one, of our ethical system, it may be the one that dominates the law as administered by the courts because of the courts' inability to promote other goals effectively. With the same proviso, two normative uses of economics mentioned earlier—to clarify value conflicts and to point the way toward reaching given social ends by the most efficient path—are untouched by the philosophical debate.

Moreover, economic analysis of law should not be rejected merely because one is unconvinced by the most aggressive version of that analysis. One could believe that

economics explained only a few legal rules and institutions but that it could be used to improve many of them, or that it explained many of them but regrettably so because economics is an immoral guide to legal policy, or even that economic analysis of law had little explanatory or meliorative significance but was intellectually fascinating—and in any of these cases one would not want to shut this book quite yet!

Another recurrent criticism of the economic approach to law —although it is better described as a reason for the distaste with which the subject is regarded in some quarters—is that it manifests a conservative political bias.[1] We shall see that its practitioners have found that capital punishment deters, legislation designed to protect consumers frequently ends up hurting them, no-fault automobile insurance is inefficient, and securities regulation may be a waste of time. Findings such as these provide ammunition to the supporters of capital punishment and the opponents of the other policies mentioned. Yet economic research that provides support for liberal positions is rarely said to exhibit political bias. For example, the theory of public goods (see §16.4 *infra*) could be viewed as one of the ideological underpinnings of the welfare state, but is not so viewed; once a viewpoint becomes dominant, it ceases to be perceived as having an ideological character. The criticism also overlooks a number of findings of economic analysts of law, discussed in subsequent chapters of this book—concerning right to counsel and standard of proof in criminal cases, bail, products liability, the application of the First Amendment to broadcasting, the social costs of monopoly, damages in personal-injury cases, the regulation of sex, and many others—that support liberal positions. Perhaps the best evidence that economic analysis of law is ideologically neutral or balanced is the significant number of prominent practitioners of it who are decidedly liberal, such as Ian Ayres, Guido Calabresi, John Donohue, Gillian Hadfield, Jon Hanson, Christine Jolls, and Daniel Rubinfeld.

The economic approach to law is criticized for ignoring "justice." One must distinguish between the different meanings of this word. Sometimes it means distributive justice, the proper degree of economic equality. Although economists cannot tell society what that degree is, they have much to say that is relevant—about the actual amounts of inequality in different societies and in different periods, about the difference between real economic inequality and inequalities in pecuniary income that merely offset cost differences or reflect different positions in the life cycle, and about the costs of achieving greater equality. These matters are discussed in Chapter 16.

A second meaning of justice, perhaps the most common, is—efficiency. We shall see, among other examples, that when people describe as unjust convicting a person without a trial, taking property without just compensation, or failing to make a negligent automobile driver answer in damages to the victim of his negligence, this means nothing more pretentious than that the conduct wastes resources (see further §8.6 *infra*). Even the principle of unjust enrichment can be derived from the concept of efficiency (§4.14 *infra*). And with a little reflection, it will come as no surprise that in a world of scarce resources waste should be regarded as immoral.

But there is more to notions of justice than a concern with efficiency. It is not obviously inefficient to allow suicide pacts; to allow private discrimination on racial, religious, or sexual grounds; to permit killing and eating the weakest passenger in

§2.3 1. Although not enough of one for some tastes! See, e.g., James M. Buchanan, Good Economics—Bad Law, 60 Va. L. Rev. 483 (1974); Richard A. Epstein, A Theory of Strict Liability, 2 J. Leg. Stud. 151, 189-204 (1973).

the lifeboat in circumstances of genuine desperation; to force people to give self-incriminating testimony; to flog prisoners; to allow babies to be sold for adoption; to allow the use of deadly force in defense of a pure property interest; to legalize blackmail; or to give convicted felons a choice between imprisonment and participation in dangerous medical experiments. Yet all these things offend the sense of justice of modern Americans, and all are to a greater or lesser (usually greater) extent illegal. An effort will be made in this book to explain some of these prohibitions in economic terms, but most cannot be; there is more to justice than economics, a point the reader should keep in mind in evaluating normative statements in this book.

The first edition of this book was published three decades ago. Even if that were taken to mark the beginning of economic analysis of law—which would date it too late, since there was already a significant scholarly literature—the field has now outlasted legal realism, legal process, and every other one of the twentieth century's new fields of legal scholarship, except those too recent to have yet reached their peak. And it shows no signs of abating. Like some of the other fields, it may some day become so deeply woven into the fabric of the law that it ceases to be visible as a distinct field. But for now, it is well worth studying as a fruitful, interesting, and influential body of insights and analytic techniques.

Suggested Readings

1. Ronald H. Coase, The Problem of Social Cost, 3 J. Law & Econ. 1 (1960).
2. Jules L. Coleman, Markets, Morals, and the Law (1988).
3. Harold Demsetz, The Primacy of Economics: An Explanation of the Comparative Success of Economics in the Social Sciences, 35 Econ. Inquiry 1 (1997).
4. Neil Duxbury, Patterns of American Jurisprudence, ch. 5 (1995).
5. David D. Friedman, Law's Order: What Economics Has to Do with Law and Why It Matters (2001).
6. Duncan Kennedy, Cost-Benefit Analysis of Entitlement Problems: A Critique, 33 Stan. L. Rev. 387 (1981).
7. William M. Landes & Richard A. Posner, The Economic Structure of Tort Law, ch. 1 (1987).
8. Arthur Allen Leff, Economic Analysis of Law: Some Realism About Nominalism, 60 Va. L. Rev. 451 (1974).
9. Thomas J. Miceli, Economics of the Law: Torts, Contracts, Property, Litigation (1997).
10. Chicago Lectures in Law and Economics (Eric A. Posner ed. 2000).
11. Ethics, Economics, and the Law, 24 Nomos (1982).
12. Foundations of the Economic Approach to Law (Avery Wiener Katz ed. 1998).
13. Symposium on Efficiency as a Legal Concern, 8 Hofstra L. Rev. 485, 811 (1980).
14. Symposium: Economists on the Bench, 50 Law & Contemp. Probs. 1 (Autumn 1987).
15. Symposium on Post-Chicago Law and Economics, 65 Chicago-Kent L. Rev. 3 (1989).