Courts, Legislatures, and the General Theory of Second Best in Law and Economics

Thomas S. Ulen

Follow this and additional works at: https://scholarship.kentlaw.iit.edu/cklawreview
Part of the Law Commons

Recommended Citation
COURTS, LEGISLATURES, AND THE GENERAL THEORY OF SECOND BEST IN LAW AND ECONOMICS

THOMAS S. ULEN*

I. INTRODUCTION ............................................. 190

II. FIRST BEST IN LAW AND ECONOMICS ..................... 192
    A. Welfare Economics, Imperfections, and the
    Economic Analysis of Law .......................... 192
    B. The Non-Uniqueness of Pareto Optimality ........... 194
    C. Legal Intervention on the Basis of Transaction
    Costs ............................................. 196
       1. The Coase Theorem ............................ 196
       2. Contract law .................................. 197
       3. Tort law ...................................... 198
       4. Public law .................................... 200

III. THE GENERAL THEORY OF SECOND BEST IN LAW AND
     ECONOMICS ......................................... 203
    A. The Theory's General Implications for Economic
    Policy ............................................. 203
    B. Second Best in Law and Economics ................... 208
       1. Private law .................................... 208
       2. Public law ..................................... 213

IV. THE IMPLICATIONS FOR LAW AND ECONOMICS: THE
    CHOICE OF INSTITUTION TO CONSIDER SECOND-BEST
    EFFECTS ........................................... 216

V. CONCLUSION ............................................. 219

* Alumni Distinguished Professor of Law, and Professor, Institute of Government and
  Public Affairs, University of Illinois at Urbana-Champaign. Robert Spencer and Christophe
  Boeraeve provided me with excellent research assistance. Russell Korobkin and E. Clay Ulen,
  Jr., were kind enough to read an earlier draft and to share their extensive and very helpful
  comments with me.
INTRODUCTION

Law and economics has thrived by applying microeconomic theory, particularly the insights of welfare economics, to the analysis of legal rules, processes, and institutions. The core recognition of law and economics is that legal rules can correct market imperfections, such as monopoly and the presence of external costs and benefits. Antitrust law and the regulation of natural monopolies, such as telecommunications (in the past) and water supply, were the early manifestations of this recognition. The more modern insights of law and economics result from the perception that private law areas such as property, contract, and tort can correct market imperfections, or more generally, encourage the efficient use of social resources. For example, an economic analyst would favor the general protection of intellectual property to correct for the public-good nature of information.

1. The early influence of economics in each of these areas of the law was particularized and not the thorough, comprehensive treatment that characterizes the modern law-and-economics literature. In antitrust law, economics was useful for such tasks as defining the social virtues of competition, enumerating the social costs of collusion, and characterizing the hallmarks of competitive markets. An older generation of antitrust scholars made only this particularized use of economics. See, e.g., Philip Areeda, Antitrust Analysis: Problems, Texts, Cases (1981). A more modern generation of antitrust scholars sees the field more comprehensively through the lens of economics. See, e.g., Richard A. Posner, Antitrust Law: An Economic Perspective (1976) (expressing sympathy for the use of economics in antitrust jurisprudence); Stephen F. Ross, Principles of Antitrust Law (1993) (indicating less sympathy). Law and economics counts the revolution in antitrust law in the 1970s as one of its significant early successes in influencing the law.

In the governmental regulation of natural monopolies, economics was useful to an earlier generation in providing a thorough justification of the special status of natural monopolies—namely that declining average and marginal costs over the relevant range of output make it unlikely, if not impossible, for competition to flourish in those industries; rather, it is economically most efficient for production to be in the hands of a single producer. The generation and distribution of water and electricity and of telecommunications services are examples of natural monopolies. For a classic treatment, see Alfred E. Kahn, The Economics of Regulation: Principles and Institutions (1970), (2 vols.).

Law and economics also counted its influence on the federal deregulation of industry in the late 1970s as another significant accomplishment. The arguments for deregulation had two different components. One was to point out that there were some regulated industries, such as the airlines, which had never been natural monopolies and should never have been regulated, at least on that ground. Another was to point out that there were other regulated industries, such as telecommunications and, more recently, electricity distribution, that had been natural monopolies under an older technology but were, under newer technologies, competitive industries for which traditional rate-of-return regulation was no longer appropriate. See generally Stephen G. Breyer, Regulation and Its Reform (1982).
and encourage the investment of real resources in the production and dissemination in inventive and creative activity.

Neither private nor public law economic analysts, however, have paid much attention to a well-known proposition in welfare economics known as the General Theory of Second Best (the "theory"). That theory suggests that the seriatim correction of market imperfections will not necessarily improve overall social welfare. Indeed, the theory holds that piecemeal correction, or correction of a subset of the imperfections, might reduce social well-being. The only circumstance in which the correction of market imperfections will inevitably improve social welfare is that in which society simultaneously corrects all of the imperfections. The clear implication of the theory for policymaking is that one ought to do nothing unless one can do everything.

Law and economics is not alone in ignoring the General Theory of Second Best. Economists, too, have generally paid little attention to the theory. For example, most modern intermediate and graduate microeconomics text do not even mention the theory. The reason for this is not that they have discovered a fatal flaw in the theory; rather, they simply do not believe it, or if they do, choose to act as if it is a figment of the ivory tower with little practical application. This curious selectivity, in which a learned profession chooses to believe some, but not all, of the results in the field, is a subject far beyond my abilities to explain.

Professor Richard Markovits has almost single-handedly shown the potentially devastating implications of the General Theory of Second Best for the general claim of efficiency in law and economics.

3. See Lipsey & Lancaster, supra note 2, at 17.
4. See id.
5. See id. at 11.
7. I, like most economists, clearly recognize the anomaly of the theory's reception in economics and am aware of how little I have, heretofore, chosen to take it seriously.
Thus far, the mainstream of law and economics has not responded. But it should: the criticisms of the theory are serious and deserve careful consideration. I try to begin that response in this Article.

I open with a brief overview of the connections between welfare economics, the theory of market imperfections, and the analysis of legal rules. I then briefly summarize the theory and show its implications for economic policy generally and law and economics specifically. I conclude by discussing how law and economics should take account of the insights of the theory in the future.

II. FIRST BEST IN LAW AND ECONOMICS

In this section, I show that in the absence of market imperfections there is no particular case to be made for legal intervention into private decisionmaking on the basis of efficiency. (An alternative case for intervention might be made, for instance, on the basis of distributive justice.) Rather, where there are no imperfections, private bargaining can achieve efficient outcomes without help. Then I show that when there are market imperfections, legal rules can correct the inefficiencies that arise from these imperfections.

A. Welfare Economics, Imperfections, and the Economic Analysis of Law

Welfare economics considers the ability of markets to achieve allocative (or "Pareto") and productive efficiency.\(^9\) According to the First Fundamental Theorem of Welfare Economics, perfectly competitive markets will lead to allocatively efficient outcomes.\(^10\) There are, however, identifiable circumstances that prevent markets from achieving Pareto-optimal outcomes: market power (i.e., monopoly and monopsony), external costs and benefits (i.e., costs or benefits involuntarily imposed or conferred upon third parties by the utility or profit-maximizing activities of others), public goods (i.e., goods for which the costs of excluding non-paying beneficiaries are so high that no profit-regarding supplier will be able to supply them in the optimal

---

\(^9\) Allocative efficiency obtains when it is impossible to reallocate holdings of goods and services so as to make some people better off without making other people worse off. See ROBERT COOTER & THOMAS ULEN, LAW AND ECONOMICS 12 (2d ed. 1997). Productive efficiency holds if it is impossible to reallocate inputs among productive processes so as to produce a greater total output (or to produce the same amount of total output by means of fewer inputs). See id. at 11-12.

\(^10\) See KREPS, supra note 6, at 286-87; VARIAN, supra note 6, at 490-93.
amount to the market), and severe informational asymmetries (i.e., the possession of information by one party that is of great value to the other party but that the first party, for reasons of strategic behavior, will not reveal to the second party).

Normative economic policy analysis identifies market imperfections and recommends corrections for those imperfections. For example, if market power prevents market efficiency, then the correction is to make that market more competitive by, for example, directly regulating output prices, encouraging entry into the market, and breaking up the existing monopoly into rival parts.\(^1\) An earlier phase of law and economics concentrated on taxation, governmental regulation of natural monopolies, and antitrust, in large part because the application of welfare economics in those fields was straightforward.\(^2\) As first applied to the analysis of law, this framework served the early phase of law and economics well.

Since Professor Coase first offered his famous theorem,\(^3\) however, the method by which welfare economics has been used to examine the law has changed. I refer not so much to the fact that Coase turned the analysis from the traditional areas of interaction between economics and law noted above, but more importantly for present purposes, to the fact that the style of analysis changed. Coase (and the literature that developed in explication of the Coase Theorem)\(^4\) began the analysis with the clear identification of transaction costs rather than market imperfections.\(^5\) There is, of course, a correspondence between the notions of market imperfections and transactions costs: I believe Coase intended the category of transaction costs to encompass all the sources of market imperfection, or "all things that prevent people from achieving their ends through voluntary exchange."

As a result of Coase's innovation, law and economics now begins its analysis of any legal rule or institution by identifying the impediments to a voluntary bargaining solution to potential or actual problems. If the impediments are minimal, then there is probably no

1. For the general economic argument about market imperfections and policy correctives, see Cooter & Ulen, supra note 9, at 38-41.
3. R. H. Coase, The Problem of Social Cost, 3 J.L. & Econ. 1 (1960). The Coase Theorem holds that if transaction costs are zero (i.e., if the costs of effectuating a bargain are zero), then resources will be used efficiently, regardless of the assignment of legal rights. See id.
4. For a review of that literature, see generally Cooter & Ulen, supra note 9, at 79-93.
5. See Coase, supra note 13, at 28-42.
need for social or legal intervention into private decisionmaking.\textsuperscript{16} Replacing the language of “perfect markets,”\textsuperscript{17} this alternative analysis stresses that Pareto-optimal outcomes will occur when the costs of effectuating bargains are so low that, whenever there is a cooperative surplus that could be split between any two potential transactors, they will find a means of splitting the surplus and thereby mutually benefiting. It follows that the law should intervene only when transaction costs frustrate a bargain solution to resource allocation. My point here is a simple one: the economic analysis of law is a straightforward application of welfare economics to the non-market setting of the law, substituting the concept “transaction costs” for that of “market imperfections.” I shall return to an example shortly.

\textbf{B. The Non-Uniqueness of Pareto Optimality}

One might think that the application of welfare economics to policy and to law and economics would, if done appropriately, lead to a unique Pareto optimum. That is, wise and well-meaning analysts could identify the market imperfections (or sources of high transaction costs) and apply the correctives to reach one and only one Pareto-efficient allocation of goods and services. From that point one could alter resource holdings to achieve whatever distributional state the analyst desired.\textsuperscript{18}

In terms of law, this view implies that there is one set of legal rules (and institutions) that would be first-best or Pareto-efficient. We might be willing to hedge that conclusion and say that this uniquely efficient set of legal rules would be true only for a specific jurisdiction, given its resources, technology, citizens’ tastes, history, and so on. That is, we need not go so far as to suggest that welfare economics or transaction-cost analysis posits that there is one set of efficient legal

\textsuperscript{16} Indeed, third-party intervention might frustrate private bargaining, by, for example, causing actors inefficiently to spend resources getting around the third-party-imposed constraints.

\textsuperscript{17} The replacement of that category is appropriate in the economic analysis of law because many of the matters on which the law focuses are non-market transactions.

\textsuperscript{18} This separation of efficiency and equity concerns is a common one in welfare economics. In fact, the Second Fundamental Theorem of Welfare Economics implies that public policy solutions to social problems should make this separation rather than seek to use, for example, price regulation to achieve distributional goals. See \textsc{Varian}, supra note 6, at 517-19. Thus, the Second Fundamental Theorem criticizes ceilings on rental rates for apartments as confounding efficiency and equity. Instead, the separation theorists would recommend very little regulation of rental rates so as to foster efficiency in the rental and closely related housing markets combined with income supplements for those burdened by housing costs.
rules for all societies or even for every sub-region of a particular society.

Even this circumscribed view would be mistaken. It is important to be clear about why it is mistaken. Certainly there will be reasonable differences of opinion about whether or not some situation is a market imperfection that needs correction, or whether or not transaction costs are so high as to make out a case for legal intervention. Furthermore, even if there is agreement about which situations are inefficient and need correction, there may be reasonable differences about the most appropriate method of intervention in order to achieve efficiency. Of course, those (and other pertinent) differences will exist. But more importantly, even if we agree on what counts as an imperfection or transaction cost and what the appropriate correctives are, there will still be no unique efficient outcome for a given economy. In terms of law, even for a specific jurisdiction there is no uniquely efficient set of laws. There is, as economists were once wont to say, no optimum optimorum. This is a well-known result among economists: in any given economy there is no uniquely optimal distribution of goods and services. Allocative efficiency is an inherently contingent concept, one that crucially depends on the initial endowment positions of the participants in the economy. Each different initial endowment will result in a different Pareto-optimal (or allocatively efficient) outcome so that there are as many Pareto-optimal outcomes as there are distinct initial endowments.

The fact that we are unable to specify an optimum optimorum has not brought the welfare analysis of economics to its knees. Economists have long sought a method for detecting which among the many allocatively efficient outcomes is best, but they have not thus far succeeded. Instead, the field sets distributional issues to one side and carries on with the welfare analysis. That is, all our prescriptions for allocative efficiency are best only in the constrained sense that we are

19. See Varian, supra note 6, at 499-501.
20. See id. at 496-98.
21. The story is familiar to any student of microeconomic theory. As indicated, there are many Pareto-efficient allocations, and economists have sought to discover a method of choosing which among them is best. The core issue is discovering a method of aggregating individual preferences over social outcomes into societal preferences. The Arrow Impossibility Theorem holds that there is no method of so aggregating individual preferences that does not violate one or more of five very simple characteristics that an aggregation method ought to have (such as that no one person's preferences should always determine society's preferences). See Kenneth J. Arrow, Social Choice and Individual Values (2d ed. 1964). For a modern introduction to the literature on social choice, see generally John Craven, Social Choice: A Framework for Collective Decisions and Individual Judgements (1992).
fully aware that we are leaving something out—namely, the distributional consequences of a different endowment. There is a sense in which this constraint should not be so troubling. It embodies the postmodern notion that everything is highly contextualized. First-best and uniquely optimal results are classroom exercises, not guides to the real world. History, political possibility, geography, custom, and all the other details matter too much to be ignored. There is no need to speculate on how things might have been if we had been able to start with a clean slate.

I have digressed to point out this standard result from welfare economics to stress the uncertainty and unavoidable indefiniteness at the heart of the welfare implications of modern microeconomics. This is, I believe, a very important point to make in light of the topic of this Article. As I later point out, one of the criticisms of the General Theory of Second Best has always been that it brings analysis to a halt. By asserting that a correction of an imperfection does not necessarily increase social welfare, the theory makes one skeptical of the advisability of any corrective. And yet correctives continue, as does the academic consideration of their utility. By analogy, the recognition that there is no unique Pareto optimum, only contingently optimal allocative efficiency, has not stopped the search for more efficient and more just states of the world. The willingness to soldier on in the face of this ambiguity is laudable; in any case, a policy of doing nothing to make the economy and society more efficient would be otiose.

C. Legal Intervention on the Basis of Transaction Costs

The argument above illustrates that we need legal intervention in private decisionmaking only when we find what traditional analysis calls "market imperfections" and what law and economics calls "transaction costs." The relationship between this observation and the standard results in law and economics are so well known that I shall not dwell on them here. Rather, I shall very briefly indicate the connection between transaction costs and a corrective legal rule.

1. The Coase Theorem

We have seen that the Coase Theorem holds that private parties will bargain to the mutually satisfactory division of a cooperative surplus when the cost to them of effectuating this bargain is low. That is, private bargaining is the most appropriate means of allocating soci-

22. For a thorough discussion, see COOTER & ULEN, supra note 9, at 87-89.
ety's scarce resources (by comparison to social or third-party allocation) when transaction costs are low. Law does not need to intervene in private decisionmaking where that decisionmaking is likely to be efficient. Utilizing the Coase Theorem to analyze the law consists of no more than skillfully recognizing the factors that lead to high transaction costs. These factors include such things as the number of people involved in bargaining and the uniqueness or fungibility of the item or performance at issue. Once identified, in theory an appropriate legal rule can correct for the high transaction cost.

2. Contract law

Consider contract law, which concerns the enforceability of consensual agreements and the appropriate remedies for the non-performance of those agreements. What role does "transaction costs" play in explaining the issues of enforceability and remedy? Law and economics will generally intervene with private bargaining to foster more efficient agreements by helping parties to do things that they cannot do for themselves. For example, consider that consensual agreements typically require the parties to forego other profitable opportunities until they fully perform their promises. Such a binding commitment is a tricky matter. In principle, both parties are better off making the commitment than not; otherwise, they would not be involved in this particular consensual agreement. However, it may turn out that one or both are even better off making the commitment and then breaking it. If both parties recognize that there may be an incentive to make commitments that will turn out to be more valuable in the breach than in the keeping, they must take steps to reduce any incentive to break the commitment. That is, they must make these commitments credible.

Perhaps they can do so by themselves. They might, for instance, leave sums of money with a third party (a form of surety bond) that would be forfeit to the innocent party if one should break his commitment to perform the promise. Or the parties might exchange hostages and agree to forfeit them in the event of non-performance. Short of these devices, parties wishing to enter into consensual agreements could invest resources to establish a reputation as a trustworthy person, a reputation that would suffer from an opportunistic breach of the commitment. But while such methods are hypothetically avail-

23. See id. at 180-85
24. See id. at 168-71.
able, they are expensive, both directly and in terms of opportunity cost. These costs may even be so high that they would discourage contract formation.

Law can help foster efficient consensual agreements by establishing a set of rules that will reduce the costs of forming and enforcing those agreements. These rules could specify how, when, with whom, and under what circumstances one may form such an agreement, what substantive terms are allowable, and the remedies that will be available to the innocent party if the other party breaks the agreement. These default rules, as they are called, may be modified by the parties within broad boundaries and allow the parties to specify (within limits) the damages for which they will be responsible in the event of non-performance. Other rules may be immutable, such as a rule that forbids enforcing a consensual agreement against someone who is not mentally competent to make agreements. Or they may be rules that a court will apply if the parties cannot themselves resolve a dispute within the terms of the contract, such as a rule that requires the innocent party to perform specifically. All of these rules may be generally characterized as rules to which the parties themselves would have agreed if the costs of transacting were low. I leave for a future section the consideration of a particular contractual rule in light of this general formulation.

3. Tort law

Consider also tort law, which from an economic standpoint concerns the efficient allocation of costs arising from accidents. If one stranger injures another accidentally and there is no law that specifies who is responsible for those injuries, there is no compelling incentive for the injurer to compensate the victim. Perhaps more importantly, there is no ex ante incentive for the potential injurer to take care not

25. This is the emerging consensus in law and economics about contract rules. See Cooter & Ulen, supra note 9, at 161-202; Symposium on Default Rules and Contractual Consent, 3 S. Cal. J. Interdisc. L. 1 (1993). The dominant view is that the default rules mimic what the parties themselves would have agreed to under low transaction costs. But there is another view of default rules (a supplementary, not substitute view) under which the law should design some default rules to solve severe informational asymmetries between the contracting parties. Thus, the rule would impose a penalty on one party for failing to disclose something material to the other party. An example is the rule from Hadley v. Baxendale, 156 Eng. Rep. 145 (Ex. 1854), under which the failure to disclose special losses arising from breach precludes the innocent party from recovering anything more than the reasonably foreseeable losses from breach. See generally Ian Ayres & Robert Gertner, Filling Gaps in Incomplete Contracts: An Economic Theory of Default Rules, 99 Yale L.J. 87 (1989); Ian Ayres & Robert Gertner, Strategic Contractual Inefficiency and the Optimal Choice of Legal Rules, 101 Yale L.J. 729 (1992).
to impose the injuries in the first place. The costs of taking care would be an out-of-pocket or opportunity cost borne by the potential injurer with no clear payoff. Therefore, rational potential injurers would not take care.

Can private parties correct this state of affairs themselves? Potential victims might contract with potential injurers to take additional care (preventing or lessening the severity of accidental losses), thereby avoiding accident losses. Victims should be willing to pay injurers a portion of the sum that they would save from not being injured; injurers should be willing to accept a sum from victims that is greater than the cost of purchasing precaution.26 Under these terms, both parties are better off than they would be by simply letting accidents happen and the costs fall where they may.

The problem is that such exchanges between potential victims and injurers are obviously too expensive to take place in many circumstances. Consider, for instance, that one does not know with whom one ought to transact in order to receive the contractual protection. Even aside from the costs of identifying all those parties with whom one ought to make an agreement, there may be holdouts or free riders. Many potential victims—let us assume that they can so identify themselves—would like to have someone else contract with the potential injurers to take care, on the theory that they will be the beneficiaries of the precaution taken by the injurers, even though they did not pay for it.27 Potential injurers (again assuming that they can so identify themselves) would be reluctant to enter into individual agreements to take care if there were costs of distinguishing among those victims entitled to precaution and those not so entitled; they would rather hold out to conclude an agreement with all the potential victims whom they might injure. These and other likely conditions make private consensual agreements between potential injurers and victims a

26. One can imagine far more complicated situations for these hypothetical transactions, such as one in which the parties are not sure whether they are going to be a victim or an injurer. For a favorable review of the possibilities of a contractual method of dealing with accidents between strangers, see Robert Cooter, Towards a Market in Unmatured Tort Claims, 75 VA. L. REV. 383 (1989).

27. Of course, this free riding is not likely to occur if the precaution can be taken specifically with respect to an identifiable victim. This is not an impossible state of affairs to imagine, but it is unlikely in the real world. Moreover, there are secondary problems that arise in a regime in which victims provide precaution tailored only to those with whom they have contracted. Consider, for instance, the problem of finding a way for those who have contracted for the injurer’s precaution to identify themselves in such a way as to prevent counterfeit identification by non-contracting potential victims.
highly unlikely means of reducing the probability and severity of accidents.

Law can correct this problem by creating a set of rules designed to minimize the social costs of accidents, which include the precaution, accident, and administrative costs. Tort liability creates an incentive for potential victims and injurers to internalize the costs of accidents. By holding out the possibility of imposing liability on the injurer (or not, in which case the accident losses become the responsibility of the victim or of her insurer), tort law seeks to induce parties to take all cost-justified precaution—i.e., to take precaution up to the point at which the marginal cost of the last unit of precaution taken equals its marginal benefit (which is itself equal to the reduction in the expected costs of the accident). I leave for discussion in a section below the consideration of a particular tort liability issue in light of this general formulation.

4. Public law

Public law is a vast area that encompasses crime, constitutional law, administrative law, bankruptcy law, labor law, and more. Those areas have received substantial treatment in the law-and-economics literature, but because I shall draw below what I think is an important distinction between the manner in which private and public law treat issues of second best, we should have an example of the economic analysis of a public law topic before us.

Law and economics has only a sketchily developed theory of when to deal with a market imperfection or transaction-cost problem by means of public law rather than by means of private law. Generally, law-and-economics theory treats public law and private law as interchangeable means to achieve social ends; society should choose the cheaper of those substitute methods of solving the social problem


29. The expected cost of an accident is equal to its probability of occurring times the losses that will be imposed if the accident occurs. This is, of course, the modern formulation of the famous Hand Rule from United States v. Carroll Towing Co., 159 F.2d 169, 173 (2d Cir. 1947). For an extended discussion of the application of this rule, see COOTER & ULEN, supra note 9, at 281-83.


at hand. Consider, for instance, the issue of automobile safety. We might define socially optimal automobiles as those for which the marginal cost of providing the last unit of precaution in the car are equal to the marginal benefit of the precaution provided. Installing armor plating on the sides of an automobile would not, therefore, be optimal if the cost of installing and maintaining that plating (including the increased cost of operating the car because of its increased weight) exceeded the benefit (which would include the reduction in the likelihood of an accident and in the severity of injuries to those riding in the car).

Society can hope to induce manufacturers and consumers of automobiles to produce and purchase socially optimal automobiles in one of several very general ways. First, it might rely entirely on the market. That is, one might argue that the transaction costs between manufacturers and consumers of determining the appropriate amount of safety to build into a car are so low that there is no warrant for intervention into this private decisionmaking. For example, if consumers perceive that the costs of passenger-side airbags exceed their benefits, they will not purchase them, and therefore manufacturers will not put them in every car. Because some consumers may perceive that those airbags are worth the increased cost of the car so equipped, manufacturers may want to offer those airbags as an optional purchase on their models of automobile. Notice that, under the assumption that many consumers perceive that the costs of airbags exceed the benefits, it would be socially wasteful to compel them to purchase passenger-side airbags.

Suppose, by contrast, that there are good reasons for believing that the transaction costs between manufacturers and consumers are so high that some cost-justified safety devices will not be produced and purchased. What legal corrective will induce the production and sale of these precautions? Private law might come into play by allowing those injured in automobile accidents to sue the manufacturer for negligence in not producing a more crashworthy car. If in the relevant jurisdiction the courts hold that it is negligent to produce an automobile without all those safety devices whose marginal benefits

32. For a discussion of the economic issues in choosing between administrative agency regulation and tort liability, see Charles Kolstad et al., Ex Post Liability for Harm vs. Ex Ante Safety Regulation: Substitutes or Compliments?, 80 AM. ECON. REV. 888 (1990). For an application of these matter to automobile safety, see generally JERRY L. MASHAW & DAVID L. HARFST, THE STRUGGLE FOR AUTO SAFETY (1990).

33. In the alternative formulation there is no market imperfection that needs correction.
exceed their marginal costs, then rational manufacturers will begin to produce cars that incorporate all such safety devices. Moreover, when technology causes the costs of newer safety devices to fall or changes in medical technology cause the benefits of some safety devices to rise, automobile manufacturers will have the appropriate incentives to incorporate new safety into automobiles.

Alternatively, society could deal with high transaction costs between automobile manufacturers and consumers by means of ex ante regulation. For example, the National Motor Vehicle Safety Administration ("NMVSA") could determine which safety devices are cost-justified and which are not and then compel all those who sell motor vehicles in the United States to include all cost-justified safety devices. Failure to do so would expose the manufacturer and retailer to fines.34

Whether by private law or by public law, the end is the same: socially optimal automobiles. Society may choose whichever of the policies seems more effective or cheaper. There may also be systemic reasons for preferring one or the other form of regulation. For instance, private law works most effectively when victims bring actions against their injurers and least effectively when victims do not bring private causes of action. If there is some reason for believing that victims of automobile accidents will not bring actions against manufacturers for the negligent design and manufacture of a car (as, for example, might be the case if the costs of litigation are high or the victims are fully insured),35 then there is a relatively strong argument for regulation through the NMVSA rather than through private causes of action. But if one believes that regulatory agencies are subject to capture by those whom it seeks to regulate, so that the NMVSA will operate not so much in the interest of producing safer automobiles as in the interest of producing more profitable automobiles, then there is a relatively stronger case for private causes

34. Many details and elaborations could be added to this sketch. One of the most important is whether or not compliance with the NMVSA regulations is a complete defense to a private cause of action against an automobile manufacturer for negligent design and manufacture.

35. However, I do not want to suggest that these factors necessarily interfere with the incentives of victims to bring private causes of action. For instance, while high litigation costs may discourage some victims, they also encourage potential injurers to take care, on the theory that additional precaution and the consequent lower number or severity of accidents are cheaper than litigation. The net result may be no change in the quality and quantity of accidents. With regard to insurance, the victim's incentive to bring an action is, of course, diminished by the extent of his first-party insurance. However, it is in the interests of insurers to recover from injurers who are liable. (If they can recover, their insurance rates will be lower.) Therefore, most insurers have a subrogation clause entitling them to pursue any legal claim that their customers may have under the insurance contract. This may correct for the diminution in the victim's incentive to sue.
of action as the principal guarantor of automobile safety. Finally, it is
worth remarking that the strongest case may be one that combines
administrative agency regulation and private liability so as to compen-
sate for the deficiencies of each. But that is a topic beyond the focus
of this Article.

For our present purposes, it is enough to have shown that public
law may correct market imperfections or high transaction costs. I
have suggested that the choice of private law, public law, or the mar-
ket as a method of achieving social efficiency may be made according
to the same economic considerations that dictate all our other choices
in life. That is, there is nothing fundamentally different about public
law, compared to private law, as a means to achieve Pareto efficiency.

III. The General Theory of Second Best in Law
and Economics

In this section, I consider the implications of the General Theory
of Second Best for law and economics. First, I will show what the
theory holds and how economists have responded to it. That accom-
plished, I shall turn to consider how the theory applies to the eco-
nomic analysis of law.

A. The Theory's General Implications for Economic Policy

We have seen that welfare economics identifies a limited set of
impediments to the attainment of social optima. Economists have al-
ways inferred that as a matter of policy, correcting market imperfec-
tions is always desirable. Because resources are limited, it would
generally be impossible to correct all the imperfections at once. But
such economic considerations would seem to suggest that social wel-
fare would be advanced by correcting as many of the imperfections as
possible, and indeed, that we might be able to rank the many imper-
fections in the order in which society should try to correct them. Pre-
sumably this ranking would be done on economic principles, so that
the first imperfections corrected would be those for which the ex-
pected net increment in social welfare (i.e., the difference between the
marginal social benefit of the correction minus the marginal social
cost of correction) is greatest. Corrections would continue down the
list until the last corrective is reached—namely, that for which the
marginal social benefit of correction equals the marginal social cost.

36. See Kolstad et al., supra note 32, at 889.
This happy picture of how to maximize social welfare by correcting market imperfections seriatim was challenged in the mid-1950s by Professors Lipsey and Lancaster. In their famous article, they argued that in general, correcting a subset of market imperfections does not necessarily improve social welfare.\textsuperscript{37} For example, suppose that there are 100 identifiable market imperfections. According to Lipsey and Lancaster, and contrary to intuition, society would not necessarily be better off (by comparison to the state of correcting none of the 100 imperfections) if it were to correct only the most egregious of those imperfections.\textsuperscript{38} No one doubts, and the theory demonstrates convincingly, that simultaneous correction of all the imperfections will fulfill the conditions of the First Fundamental Theorem of Welfare Economics. Nor does the literature seem to doubt that correction of some of the 100 imperfections will make society better off.\textsuperscript{39} But no one has demonstrated how or why this piecemeal correction will necessarily make society better off. It might improve things, but it might not. Nor are the circumstances known under which piecemeal correction will be welfare-enhancing or welfare-decreasing. The daunting conclusion is that economists can confidently recommend that some societal imperfections or inefficiencies should be corrected only when all of them are corrected at the same time. Partial correction may make matters worse and should therefore perhaps not be undertaken.

The reasons for the results of the General Theory of Second Best are technical and evident only to those with the patience for mathematical argument.\textsuperscript{40} A clear, intuitive example of the theory is not a staple of the microeconomic literature. So, the best I can do is to sketch a sense of what the theory suggests. Return to the example of 100 imperfections. Suppose that the most glaring of these imperfections is the condition of monopoly in the supply of water. Intuition, uninformed by the theory, would confidently predict that correcting the distortions of the water monopoly would improve society. The price of water would fall from a monopolistic to a competitive price, with a consequent increase in the quantity of water demanded and supplied. Other consequences would follow: those productive

\textsuperscript{37} See Lipsey & Lancaster, supra note 2, at 17.

\textsuperscript{38} The proof of this assertion is complicated. I shall shortly give an example of why it might be true.

\textsuperscript{39} The sad fact is that the profession has simply ignored this implication of the General Theory. No one has disproved the theory; nor has anyone shown that there are circumstances in which piecemeal corrective of market imperfections will improve matters. No economist today feels obliged to say anything at all about the matter.

\textsuperscript{40} See, e.g., Lipsey & Lancaster, supra note 2, at 19-20.
processes that used water would find it cheaper and might, therefore, experience an increase in profits, production, wages, or all three. Consumers who had economized on the use of water at the monopoly price would now be free to make wider use of it in many domestic tasks, including cooking, washing, cleaning, and gardening. Taken all together, these consequences of correcting the monopolistic distortion in the water market would obviously seem to be a societal improvement.

The theory holds that this obvious conclusion might be wrong. Can we imagine circumstances in which the correction of that one glaring imperfection might not improve social well-being? Consider that the conversion of the water monopoly into a competitive supplier may produce many losers. For instance, those firms who altered their production processes away from the use of water because of the monopolistic price may find themselves with inappropriate production technologies when the relative price of water falls. Those suppliers who had designed, manufactured, marketed, sold, and maintained equipment that economized on the use of water would now find their market niche at risk. Their investments, and those of their customers, in water-saving equipment would be in jeopardy. Consumers who had relandscaped their lawns and gardens to minimize water use when its price was relatively high might find their plantings inappropriate when water is competitively priced. The sum of all the losses among those who have made adjustments to monopoly water prices will not necessarily be greater than the sum of all the gains that would accrue to those who would benefit from lower water prices. We must at least admit that possibility and see why it is that correcting a single, glaring market imperfection might not necessarily make society better off.

Surely these are not the circumstances to which the General Theory of Second Best refers. The possibility that the losses from a change could exceed the gains is nothing more than the application of Kaldor-Hicks efficiency\(^4\) or cost-benefit analysis. There is a world of learning on what is likely to happen in circumstances in which gains

\(^4\) A change is said to be Kaldor-Hicks efficient when the gainers from the change could compensate the losers and still be better off than they were before the change. The compensation does not have to be effectuated, only possible. For that reason, a Kaldor-Hicks change is sometimes referred to as a "Pareto potential" change. Economists believe that the Kaldor-Hicks criterion for evaluating changes in the economy is more flexible than the Pareto criterion, which requires unanimity—i.e., the gainers must actually effectuate the payment to the losers. See COOTE\(R & ULEN, supra\) note 9, at 41; MATTHEW D. ADLER & ERIC A. POSNER, A DEFENSE OF COST-BENEFIT ANALYSIS (University of Pa. Sch. of Law Working Paper, 1998).
and losses must be calibrated and balanced. The theory must be asking us to consider something more than merely toting up costs and benefits of correcting market imperfections, laudable criterion though that may be for evaluating changes. Although this has never been entirely clear in the literature, my impression is that the theory imagines that the piecemeal correction of market imperfections may not result in a stable equilibrium, but rather in a series of adjustments (some occurring in markets far removed from those in which the corrective is at issue), none of which is stable, and none of which is clearly superior to another. In short, the theory holds that the correction of an imperfection in one market may lead to the creation of a new imperfection or to the exacerbation of an existing imperfection in another market. Piecemeal correction of market imperfections would thus decrease, not increase, social welfare.

How might this apply in our example of the correction of the water monopoly? Suppose that, as before, competitive prices replace monopoly prices for water. Suppose, also as before, that some producers now make greater use of water because of the lower price. Specifically, suppose that a production process in which water is an important input increases its output significantly because that important input is now cheaper. Finally, suppose that this particular production process pollutes the air and water around the factory and that this particular market imperfection (the polluter's external costs) has not yet been corrected (by, for example, forcing the polluter to internalize the external costs of production). The increase in pollution resulting from the increase in production, which itself resulted from the correction of the water monopoly problem, might reduce social welfare below its original level (i.e., before we corrected the monopoly problem). It is certainly possible that there may be many other responses to lower water prices that result in new or exacerbated market imperfections. For instance, the increased use of water could lead to an increase in irrigation by farmers, which could in turn cause an increase in chemical pollution from water runoff; it is not impossible to imagine that ending the water monopoly may create a monopoly position for the supplier of some other input that is in scarce supply and must be used in relatively fixed proportions with water. To the

42. Below I shall make a connection between this cost-benefit notion and the interpretation that I am about to propound of what the General Theory of Second Best means.

43. If a system of effluent charges calibrated to the amount of external cost were already in place, then the increase in the pollution attendant upon the increase in the amount of production because of the now-cheaper water would be fully internalized.
extent that the correction of the water monopoly may create new or exacerbated imperfections elsewhere in the economy, the correction of that one glaring imperfection in the supply of water clearly did not necessarily increase social welfare.44

Does this mean that our hands are tied, that we ought not correct those imperfections that we can because of prudential fear that we may be making things worse elsewhere? That is the central question posed to us by the theory. If we take the theory seriously, then we must pursue one of two policy courses. First, we may stay our hand and make no attempts to correct market imperfections, knowing that the cost society is bearing under these imperfections may be the best state of affairs for which we can hope. Second, we can direct our efforts at the correction of only manageable imperfections: those that are glaring, unlikely to have adverse consequences elsewhere, or whose correction we may plausibly couple with the corrections of other imperfections to avoid situations like that in which the correction of the water monopoly increases the external costs of air and water pollution. But in order to be confident about these latter methods of dealing with piecemeal correction (and piecemeal it must be because no society can afford to correct all imperfections at once), we must be fairly confident that we have identified all the impacts in other markets of making a correction in one market. That is, we should be aware that adjusting the water monopoly will increase pollution (and incidents of other undesirable effects elsewhere) and take the appropriate steps to implement optimal pollution-control policies.

One may be justly skeptical of the ability of economists and others to identify all the second-best ramifications of market-imperfection correctives. Such skepticism is a significant point to which the theory draws our attention: skepticism about global efficiency claims, skepticism about claims that correction of market imperfections will inevitability improve social welfare, and prudence about where we

---

44. Above I suggested that there was a difference between a cost-benefit analysis of the effects of the piecemeal correction of market imperfections and second-best considerations. There is, however, a connection between the two notions. The Coase Theorem might be read to suggest that second-best considerations—by which I mean new or exacerbated imperfections resulting from the correction of other market imperfections—are the result of high transaction costs. If those costs were low, then those whose situation is worsened by the correction of, say, the water monopoly—e.g., those who suffer harms from the increased water and air pollution—would pay those who are making increased use of the now-cheaper water to cut back their usage so as to reduce the pollution to tolerable levels. Only if it is too expensive for private parties to reach these bargained-for corrections of the unintended consequences of market corrections will there be the possibility of second-best effects.
must look before we confidently recommend welfare-enhancing correctives of social problems.

**B. Second Best in Law and Economics**

We turn now from the General Theory of Second Best in economic policy generally to its specific application in law and economics. In this section I am not interested so much in what law and economics should do in response to a recognition of the problem. That is the topic of the following section. Here I want merely to show how second-best considerations might arise in some of the areas that we have already considered as examples of the application of welfare economics to the analysis of the law.

1. **Private law**

In a previous section we looked at general examples of how private law could solve transaction-cost and market-imperfection problems. Here I want to show how particular common law contract doctrines designed to correct imperfections might have second-best effects. Above we considered contract law generally as a method to correct high transaction costs or market imperfections. Here I want to consider a particular contract doctrine to show, first, how that doctrine specifically corrects for a market imperfection (or situation of high transaction costs), and second and more importantly, how this correction may give rise to second-best effects.°

Standard-form contracts, offered on a "take-it or leave-it" basis, may or may not be efficient. If the contract has terms to which the parties themselves (say, A and B) would have agreed, had there been time and low transaction costs, then the terms may be efficient. The fact that it is a contract of adhesion is not inherently troubling because bargaining can be, in some circumstances, a waste of resources. However, if A imposes the terms in a contract of adhesion on B because of B's incompetence, or because of a severe informational asymmetry running in favor of A, or because A has a monopoly position with respect to B, then there is no reason to believe that the contract will be efficient. These considerations suggest the advisability of a case-

45. I am grateful to my colleague Russell Korobkin for extensive discussions on the matters of this section.
by-case analysis of standard-form contracts with no general presumption in favor of or against such contracts.46

As an example, consider *Tunkl v. Regents of the University of California*.47 Mr. Tunkl sought admission to the University of California at Los Angeles Medical Center in June, 1956. The Regents of the University of California operated the hospital as a nonprofit charitable institution; they admitted patients if they believed that the study and treatment of their illness would contribute to the medical research and education mission of the UCLA School of Medicine.48 As a condition of Mr. Tunkl's admittance, he signed a document that released the "Regents of the University of California, and the hospital from any and all liability for the negligent or wrongful acts or omissions of its employees, if the hospital has used due care in selecting its employees."49 When asked to execute this release, Mr. Tunkl "was in great pain, under sedation, and probably unable to read."50 Nonetheless, he (or his agent) signed the release. He was admitted and treated, but the medical treatment was unavailing, due, Mrs. Tunkl later complained, to the negligence of two physicians employed at the Medical Center. After her husband's death, Mrs. Tunkl sued the hospital, the trial court held the executed release to be valid, and it entered judgment for the Regents.51

On appeal, the California Supreme Court held that the exculpation clause violated California Civil Code section 1668, which stated, "All contracts which have for their object, directly or indirectly, to exempt anyone from responsibility for his own fraud, or willful injury to the person or property of another, or violation of law, whether willful or negligent, are against the policy of the law."52 The court recognized that exculpation clauses are generally valid "only if the public interest is not involved,"53 and held that the hospital was a public interest enterprise: it was suitable to public regulation, it engaged in an important and necessary public service, and willingly provided that service to any member of the public who sought it.54 Moreover, the hospital held a superior bargaining position with respect to Mr. Tunkl

46. The law seems to treat adhesion contracts as generally enforceable and unenforceable only in certain circumstances. See E. ALLAN FARNSWORTH, CONTRACTS § 4.26 (2d ed. 1990).
47. 383 P.2d 441 (Cal. 1963) (In Bank).
48. See id. at 443.
49. Id. at 442.
50. Id. at 442 n.1.
51. See id. at 442.
52. Id.
53. Id. at 443 n.6.
54. See id. at 445.
and took advantage of that position to press the exculpation clause in the contract of adhesion upon him without making "provision whereby [he might] pay additional reasonable fees and obtain protection against negligence."\(^5\)

The Regents defended the exculpation clause on two grounds. First, they argued that while the exculpation clause might be void with respect to paying patients, it should not be void with respect to charitable patients.\(^6\) The implicit economic justification for the Regent's position is sound. The Regents could take charitable cases only if the benefit of the study and treatment of their condition towards the hospital's educational and research mission exceeded the costs of treatment. One obvious method of keeping the cost of treatment down was to insist on the exculpation clause. Without the exculpation clause, the cost of treating charitable patients would rise, perhaps higher than the benefits of treating them for educational and research purposes, ultimately resulting in a decline in the number of charitable patients treated. If the charitable patients would prefer treatment under the exculpation clause to no treatment or to paying for the treatment without the exculpation clause, then the contract term would be mutually beneficial and, therefore, efficient. These arguments did not impress the California Supreme Court. The court felt that it would be unwise to treat patients differently with respect to exculpation clauses according to whether they were paying or charitable patients.\(^7\)

Second, and closely related to the first point, the Regents argued that the real purpose of the exculpation clause was to protect the hospital's employees. If the hospital could not immunize its employees from liability, it would be difficult for the hospital to put together a competent staff, particularly in a nonprofit, charitable teaching and research hospital. Competent staff would have sought employment in hospitals where they could be immunized from liability; otherwise, they would demand higher compensation to work in nonprofit charitable institutions that could not issue exculpation clauses. The court rejected this argument, too, as contrary to public policy.\(^8\)

Perhaps the court hoped that the decision in *Tunkl* would serve to induce future parties to craft more efficient contracts. The default rule—indeed, an immutable rule—became this: nonprofit charitable

\(^{55}\) Id. at 446.
\(^{56}\) See id.
\(^{57}\) See id. at 448.
\(^{58}\) See id.
COURTS, LEGISLATURES, & THE THEORY

hospitals could not insist on the patient's signing an exculpation clause as a condition of admission to the hospital. As with all common law rules, the additional circumstances to which this disfavor of exculpation clauses extended was not then clear but would be revealed only through future litigation. But the effects may have been widespread. Counsel to clients in other businesses that provided services that could affect the public interest might have advised their clients that they too could not ask their customers to sign exculpation clauses. One cannot be sure how much further the ripples of Tunkl spread.

Let us concede, for the sake of further argument, that the court's decision to void exculpation clauses like the one presented to Mr. Tunkl was efficient in the sense that it corrected an imperfection or a situation of high transaction costs in the market for medical services. The imperfection lay in the exercise by the hospital of its situational monopoly (Mr. Tunkl, recall, was in pain, had been sedated, and was temporarily unable to read) or in the exploitation of a severe informational asymmetry about the risks and costs of medical negligence (if Mr. Tunkl had been able to bargain). Alternatively, one might characterize the situation as one of generally high transaction costs: the costs to patients of informing themselves about the risks of malpractice, and alternative methods of protecting themselves against those risks, may be so high as to prevent patients from effectively bargaining with hospitals. One might suppose that the court's holding duplicates the contract terms that the parties themselves would have bargained to if transaction costs had been low.

Suppose that, as a result of the ruling in Tunkl, the efficiency with which nonprofit charitable hospitals deliver health care services increases. Unable to impose exculpation clauses upon their admittees, the hospital minimizes its potential liability by paying more attention to those whom it hires and to the quality of the services it provides to its patients, thus decreasing the number of iatrogenic injuries.


60. I ask for this concession because I am not certain that these clauses are efficiency-enhancing. However, for the purposes of explicating second-best effects in the common-law context, the true efficiency characteristics of the clause are not that important. For an interesting, economically sophisticated, and persuasive defense of the exculpation clause, see Richard A. Epstein, Mortal Peril: Our Inalienable Right to Health Care? 372-76 (1997); Glen O. Robinson, Rethinking the Allocation of Medical Malpractice Risks Between Patients and Providers, 49 Law & Contemp. Probs. 173, 188-95 (1986).

61. One might argue that the exculpation clause constitutes a form of moral hazard for the hospital and its physicians: relieved of the possibility of being held liable, they may take less care
patients of the hospitals would be relieved of the costs of providing for their own protection against medical malpractice. Formerly, when patients bore those costs, they may not have used hospitals as much as they should have, with adverse consequences for their health. Indeed, patients bore more of the costs of health care than they should have and therefore may have postponed going to the doctor or the hospital until their illness had progressed considerably. By the time they took themselves to the hospital, they were in a more deteriorated state of health than they would have been had they gone earlier in the illness and, as a result, it was more expensive to treat them. If there had been no exculpation clause, patients might have been willing to go to the hospital earlier, when the costs of treating them would presumably have been lower. All of these inefficiencies would be avoided if exculpation clauses were not allowed.

The General Theory of Second Best asks us to look beyond these efficiency gains to ask if this corrective adversely affects other market efficiencies (or transaction costs among private bargainers). If this corrective does have those second-best effects elsewhere, then social welfare may not have increased. To confine our attention to the issue of second-best effects, I shall continue to assume that the decision in Tunkl was efficiency-enhancing in the ways described above.62

What second-best effects might there be, in other markets, from this (assumed) improvement in the hospital services market? Suppose that voiding exculpation clauses in admissions agreements to non-profit charitable hospitals raises the costs of treating charitable patients (relative to those of treating paying patients) at research and educational hospitals. (This, recall, is not inefficient; by assumption, this is the socially optimal cost of treatment.) Consequently, the hos-

62. To make the matter tractable, we must make an additional assumption. We do not know how far the ruling in Tunkl spread—whether or not, that is, other businesses interpreted the holding to mean that they, too, should not offer contracts of adhesion with exculpation clauses to their customers. If this was the response (as it surely was), then whatever second-best effects we can identify as a result of changes in the hospital market would be amplified by changes in those other markets where the exculpation clause was now void. Trying to consider the collateral effects in other markets that responded to Tunkl by dropping their exculpation clauses would make things complicated. To remove this complication, let us assume that the immediate effects of the holding are confined to nonprofit charitable hospitals. I think that it is fair to say that if we were to consider ripple effects of the Tunkl holding on other businesses affected with a public interest, the second-best effects that I am about to describe would be exacerbated, strengthening my ultimate point.
hospital admits fewer charitable patients or cuts back on its staff. Ultimately, the result may be that, because a smaller range of patients appear in the educational and research hospitals, the opportunities for improvements in medical treatment decrease or the quality of medical education falls. Perhaps another response to the correction of the exculpation-clause imperfection is that the compensation of physicians in nonprofit charitable hospitals falls relative to their compensation elsewhere. Fewer physicians may decide to go into medical research and education, reducing the opportunities for new discoveries.

The external costs flowing from these (and other) outcomes of correcting the exculpation-clause imperfection in nonprofit charitable hospitals could be extraordinarily large. Assume for the moment that these reductions in the amount of medical research or in the rate of improvement in medical treatment are the only second-best, welfare-decreasing effects that follow from Tunkl. How can society guarantee that voiding the exculpation clause is welfare-enhancing? Clearly, on these assumptions, Tunkl's rule should have been accompanied by an increase in the public subsidy given to educational and research hospitals. That increase—or some equivalent policy, such as a change in the tax treatment of revenues or expenditures at nonprofit hospitals—would have offset the deleterious side effects of correcting the inefficiencies of the exculpation clause. Generally speaking, the efficiency-correcting aspects of private-law rules may create inefficiencies in other markets, requiring supplemental legal attention.

2. Public law

Health-care markets also help to illustrate the possibility of second-best effects in public law attempts to correct market imperfections or high-transaction-cost impediments to social optimality. One of the most notable developments in health-care markets of the last fifteen or so years is the spread of managed care. Managed care is offered in many different forms, and in those different forms covers

63. There are, of course, other policy instruments that could have been implemented to correct for the decline in the production of socially valuable medical information—such as income tax breaks for physicians willing to work in nonprofit charitable hospitals.

Recall that I have assumed that other businesses, besides nonprofit charitable hospitals, did not respond to Tunkl by dropping exculpation clauses from their contracts with customers. To the extent that other businesses affected with a public interest did take Tunkl as a warrant for correcting their own behavior, the second-best effects are likely to be magnified beyond those I have identified.

64. See Gail A. Jensen et al., The New Dominance of Managed Care: Insurance Trends in the 1990s, 16 HEALTH AFFAIRS 126, 126 (1997) (finding that, in 1995, 28% of employees were covered by HMOs, 25% by PPOs, and 20% by POSs).
some 65 million Americans.\textsuperscript{65} The principal factor motivating the re-organization of health-care markets into managed-care enterprises is an attempt to lower costs while maintaining a high quality of health-care services.\textsuperscript{66} In their attempts to do this, managed-care hospitals have been altering the manner in which some services are delivered. For example, some patients are spending less time in the hospital than has heretofore been customary,\textsuperscript{67} some treatments that had previously been offered almost exclusively on an in-patient basis are now being offered more and more on an out-patient basis, and some forms of treatment are changing dramatically.

These changes in the health care market have caused some to wonder if managed-care enterprises are taking advantage of their patients.\textsuperscript{68} Managed-care facilities have treatment protocols that may leave less scope for individual patient variation, less flexibility for physicians in their treatment options, and may disadvantage patients in their dealings with the managed-care enterprises because the terms of their agreement with the facilities are so complex and multi-faceted that no patient has the ability to understand them.\textsuperscript{69}

In response, many states have sought to restructure the managed-care arrangement with their physicians and their patients by statute. For example, the Illinois General Assembly has debated a "bill of rights" for managed-care patients and has mandated certain treatments by managed-care enterprises, such as a forty-eight-hour stay after a radical mastectomy.\textsuperscript{70} Connecticut recently passed a statute that


\textsuperscript{67} Post-natal stays are generally shorter in managed-care facilities than they are in other kinds of facilities. See Tracy Wilson Smirnoff, Note, "Drive-Through Deliveries": Indiscriminate PostPartum Early Discharge Practices Presently Necessitate Legislation Mandating Minimum Inpatient Hospital Stays, 44 CLEV. ST. L. REV. 231, 235 (1997) (reporting that a higher percentage of women with HMO coverage are discharged within 24 hours of giving birth than are women with non-HMO coverage). As has become well-known, post-operative stays for some procedures, such as a radical mastectomy, are shorter than has been the case previously.


\textsuperscript{69} For a discussion of these alleged problems in managed-care plans, see Epstein, supra note 60, at 420 (discussing the decline of the traditional physician-patient relationship in managed care), 422-33 (discussing patient-physician conflicts in managed care), 426-28 (discussing how patient complaints with managed care may result in legislative action), and sources cited therein.

\textsuperscript{70} See 215 ILL. COMP. STAT. 5/356t (West 1997).

The provisions of the proposed Illinois "bill of rights" for managed-care patients, introduced as The Managed Care Patient Rights Act, H.B. 603, 90th General Assembly, Regular Session (Ill. 1997), are identical to those of other state regulatory plans. For example, it bans "gag orders" that prevent doctors from telling patients about different treatment options, see id. § 5-25, re-
allows patients to appeal grievances with their managed-care providers to a state arbitration panel, largely at state expense.\textsuperscript{71} Congress is also proposing to impose constraints on the ability of managed-care providers to alter the services they provide.\textsuperscript{72}

quires managed-care providers to respond within three hours to a doctor's request to extend a patient's stay in the facility (with failure-to-respond to be construed as approval), see \textit{id.} § 5-45, and mandates that consumers be covered for emergency room visits, even if the consumer turns out to have no serious medical problem but had reasonable cause for concern for her health, see \textit{id.}


A particularly apt example of a legislatively enabled private law remedy comes from a recent statute passed by the Texas legislature. In early June, 1997, Texas passed a statute that would allow medical malpractice actions against health maintenance organizations ("HMOs"). \textit{See Tex. Civ. Prac. & Rem. Code Ann.} § 88.002 (West 1997); \textit{see also} Sam Howe Verhovek, \textit{Texas Allowing Suits Against H.M.O.s}, \textit{N.Y. Times}, June 5, 1997, at A16. Texas thereby became the first state to allow malpractice or negligence actions against HMOs. \textit{See Frank Bass, HMOs in Texas to Be Made Liable for Malpractice}, \textit{Wall. St. J.}, May 23, 1997, at B12. Thus far, HMOs have been able to avoid state common law liability on the theory that they do not practice medicine, but merely administer insurance policies. As HMOs and other managed-care providers increasingly specify particular procedures as a means of keeping costs down, they come closer and closer to practicing medicine. Those patients who have successfully sued managed-care providers have generally done so on a "breach of contract" theory. \textit{See Allison Faber Walsh, Comment, The Legal Attack on Cost Containment Mechanisms: The Expansion of Liability for Physicians and Managed Care Organizations, 31 J. Marshall L. Rev. 207, 228-32 (1997).}

The HMOs opposed the legislation and argued that it would drive up their premiums by 4.5 to 12%. \textit{See Verhovek, supra, at A16.} The Texas Medical Association, representing physicians, supported the bill on the theory that there was no reason that they should be liable for medical malpractice while the managed-care providers for whom they worked escape liability. \textit{See id.} The Florida legislature passed a similar bill several years ago, but Governor Lawton Chiles vetoed the act, citing concerns about its impact on costs of the health care services.

Can these statutes altering the terms and conditions of managed-care contracts be construed as correctives for market imperfections or high transaction costs? I believe that many of them can, but need not debate that issue here. As in our examples of common law holdings, let us presume that these acts regulating managed care do correct inefficiencies that the parties involved could not themselves correct.73

The neglected question upon which I want to focus is this: are there second-best effects from these regulations?

There may well be. That is, after all, the contention of the General Theory of Second Best, and it is a contention that everyone who purports to correct a market imperfection ought to take seriously. This is not to say that every correction has second-best effects or that, where there are second-best effects, they necessarily exceed the social benefit of the correction. One merely needs to be alive to the possibility that second-best effects may exist and may be substantial. This observation leaves open the question of whether or not we might do more than merely stay alive for the presence of second-best effects. That question is the focus of the next section.

IV. THE IMPLICATIONS FOR LAW AND ECONOMICS: THE CHOICE OF INSTITUTION TO CONSIDER SECOND-BEST EFFECTS

We are now in a position to elaborate on how law and economics ought to take account of the General Theory of Second Best. I shall focus on two particular effects: 1) how scholars in the field ought to take account of second-best considerations, and, more importantly, 2) how we ought to take account of second-best effects in the field's analysis of the relative desirability of private litigation and legislation as correctives for inefficiencies.

Law-and-economics scholars, like economists generally, have not paid sufficient attention to the General Theory of Second Best as an important constraint on the piecemeal, seriatim analysis of efficiency that is the convention in the field.74 Analysts should be aware that an efficient rule in one area of the law may create or exacerbate inefficiencies elsewhere in the legal system or the economy. This does not

73. I recognize that this is a particularly strong presumption in this context because the political element—by which I mean the ability of competing interest groups to frame and push the legislative process on any issue—is so strong. Interest groups are not so often interested in seeing to it that the legislature corrects market imperfections or lowers transaction costs as they are in transferring resources towards themselves.

74. But for the work of Professor Markovits, there has been very little attention to the topic of Second Best. For an exception, see Richard L. Schmalbeck, The Uneasy Case for a Lower Capital Gains Tax: Why Not the Second Best?, 48 TAX NOTES 195 (1990).
mean that second-best effects are everywhere. Sometimes they are present and large, sometimes they are present but small, and at other times they are non-existent. At a minimum, we ought to look for them routinely. But we ought to try to do more. Specifically, law-and-economics scholars should seek to develop a more general framework for identifying second-best effects. That is, we should seek to find theoretical regularities, such as the hypothetical possibility that tort rules are far more likely to produce second-best effects than contract rules, or that second-best effects are more pervasive in certain kinds of correctives (such as those for public goods, say, by comparison to those for monopoly), or that those effects are more likely to appear in public law than in private law areas.

There is a further, and I believe more important, implication of the General Theory of Second Best for law and economics: the choice of legal institution for dealing with inefficiencies. If we assume that there are instances of rule-creation in private litigation that give rise to second-best effects, we must then ask whether or not a common law court should take some account of those effects in resolving private disputes. The answer is, almost certainly not. Courts should not, and indeed, they do not do this. Their competency is limited by the arguments of the parties before them in light of legal precedent and social custom. Common law judges create rules incrementally and cautiously; in so doing, they focus on the local effects of the rule. By contrast, if common law judges were to take into account second-best effects, perhaps effects far removed from the instant controversy, they would be looking at global effects and doing so without the benefit of specific argument to guide them on the matter. Courts almost never do this; when they do, they are often accused of overreaching, or "legislating" from the bench. One of the unexpected consequences of the theory is this insight into the limited province of adjudication as a method of creating rules to correct inefficiencies. In sum, private litigation is not now, and should not become, a method of taking second-best effects into account.

In contrast, legislation, when it seeks to correct inefficiencies, does so in a context that allows for the explicit consideration of second-best effects. Legislators might discover these effects themselves,
but more likely, they will learn of them from their interested constituents. The groups who track legislation have a strong incentive to identify second-best effects (as well as other effects) and to lobby legislators to account for such effects in legislation. I am fully aware that legislation is subject to imperfections of its own (as noted in the extensive public choice literature), but at its best, Congress (and state legislatures) can most ably account for second-best effects.

Legislators are particularly adept at structuring statutes to minimize the social-welfare decreasing aspects of those effects. I do not mean that legislators are philosopher-kings. Rather, the political process is, relative to the process of private litigation, open and comprehensive. A far broader spectrum of interests can be heard and accommodated in legislation than in private litigation. Thus, parties both immediately and tangentially affected by any proposed efficiency-enhancing action contemplated by the legislature will have an opportunity to make their position known in the legislative forum. Recall the hypothetical effects of the *Tunkl* decision on medical education and innovation. The Regents of the University of California could have raised those matters in defense of the exculpation clause, but surely the California Supreme Court would have, quite rightly, dismissed those concerns as remote from the instant controversy, speculative, or beyond their competence. By contrast, if the California General Assembly had debated making exculpation clauses illegal, then the Regents and every other group that thought it might be adversely or positively affected would be on notice to make its views known to the legislature. Presumably, the legislators could then weigh these competing views and amend the legislation as they deemed appropriate. Put more concisely, the stronger the second-best effects of any proposed corrective legislation, the greater the likelihood that some interest group, eager to guard against those socially undesirable but unintended consequences of the proposed legislation, will seek to affect the legislation. I suggest here that the legislative process, precisely in the jostling among competing interests that occurs every time legislatures propose to take action, offers an opportunity for second-best effects to be incorporated explicitly.

These brief considerations of relative institutional competence with respect to second-best effects suggest that law and economics, in

choosing between private law and legislation in correcting inefficiencies, should add another factor to those that it has already identified: private law is more appropriate than legislation for addressing inefficiencies when second-best effects are trivial; legislation is more appropriate than private law for addressing inefficiencies when second-best effects may be large and widely distributed. Perhaps the division of responsibility between private adjudication and legislation already reflects this point. Put somewhat differently, it is possible that the discernible trend towards an increasing scope for statutory law and a narrowing scope for private law reflects either a broader (but unarticulated) recognition of second-best effects, or a secular increase in the extent of second-best effects. We cannot carefully evaluate the current division of responsibility between public and private law because we lack an understanding of the circumstances in which second-best effects are most extensive. Until we have a better theoretical and empirical grasp on those circumstances, we shall have to exercise our best judgment in deciding how best to deal with inefficiencies.

V. Conclusion

Law and economics is a straightforward application of traditional welfare economic theory to legal rules and institutions. Economic analysts perceive the need for legal rules, both generally and in particular, to correct market imperfections such as monopoly, public goods, external costs and benefits, severe informational asymmetries, or the presence of high transaction costs that prevent parties from bargaining to mutually beneficial gains. Analysts seek to identify those rules and institutions that would enhance social welfare in consensual agreements, the use of resources, or in the taking of precaution. The implicit assumption is that correctives that lead to more efficient

77. See Komesar, supra note 31.
78. The law-and-economics literature typically stresses normative factors for the secular increase in statutory law, not the positive factor to which I have drawn attention. See Stearns, supra note 76; see also Guido Calabresi, A Common Law for the Age of Statutes 72-80 (1985).
79. I leave for another day the possibility that social norms may address second-best effects and, if they do, how we might choose among private adjudication, legislation, and social norms as the most appropriate device for minimizing second-best effects. There is an extraordinary law-and-economics literature on the efficiency of social norms. See, e.g., Robert C. Ellickson, Of Coase and Cattle: Dispute Resolution Among Neighbors in Shasta County, 38 Stan. L. Rev. 623 (1986); Jon Elster, Social Norms and Economic Theory, 3 J. Econ. Persp. 99 (1989); Richard H. McAdams, The Origin, Development, and Regulation of Norms, 92 Mich. L. Rev. 338 (1997); Cass R. Sunstein, Social Norms and Social Roles, 96 Colum. L. Rev. 903 (1996); Symposium, Law, Economics, & Norms, 144 U. Pa. L. Rev. 1643 (1996).
resource use or more efficient consensual agreements do not elsewhere adversely impact the legal system or economy.

The General Theory of Second Best, however, holds that correction for one market imperfection will not necessarily be efficiency-enhancing unless we also simultaneous correct for all other market imperfections. This is because one correction may have unintended and unanticipated consequences that adversely impact market efficiency or transaction costs elsewhere. Applying that theory to the analysis of legal rules suggests the possibility that, for instance, a contract rule that corrects for an imperfection in the market for consensual agreements may induce welfare losses elsewhere.

The recognition of this possibility should give scholars and practitioners pause. The possibility of second-best effects should temper claims of general efficiency within the legal system. Because we are unlikely to perceive all the second-best effects of legal rules, we ought to be content to see only as far as we can and make more circumspect claims about the efficiency-enhancing aspects of our economic analysis of legal rules and institutions.

There are clearer implications of Second Best Theory for judges and legislators. Because common law judges have a relatively modest social mandate to make law, even those who are aware of second-best effects that might flow from their holdings should, nonetheless, confine themselves to considerations of local optimality. Common law holdings that seek to incorporate second-best concerns overreach. By contrast, legislators are in a position to inquire into the broadest possible social consequences of their actions—i.e., to take second-best effects into account. Typically, because of their sensitivity to political issues, legislators do pay attention to second-best effects. These observations suggest that in the division of responsibility between common law courts and legislatures, an additional factor to consider is this: to enhance welfare society should assign jurisdiction to common law courts over those matters for which there are no or few second-best effects and assign jurisdiction to legislatures over those matters for which there are or may be extensive second-best effects.