Essay

Economic Analysis of Contract Law After Three Decades: Success or Failure?

Eric A. Posner†

INTRODUCTION

Modern economic analysis of contract law began about thirty years ago and, many scholars would agree, has become the dominant academic style of contract theory. Traditional doctrinal analysis exerts less influence than it did prior to 1970 and enjoys little prestige. Philosophical work on the nature of promising has captured some attention, but petered out in the 1980s, with little to show for the effort other than arid generalizations about the nature of promising. Academic critiques from the left no longer stir up excitement as they did twenty years ago. Scholarship influenced by cognitive psychology has so far produced few insights. Only economic analysis seems to be on solid footing.

One way to validate a field's claims is to look at its history. Economically oriented scholars writing in the early 1970s had foundational insights, and then over time subsequent writers have criticized and refined them; because these refinements were derived from common premises, there has been a sense of forward movement in the subject, of the building of an increasingly sophisticated consensus. Although critics of economic

† Professor of Law, University of Chicago. Thanks to Bill Dodge, Mark Gergen, Richard Posner, Alan Schwartz, Steve Smith, David Snyder, Kathy Spier, and workshop participants at the University of Illinois at Chicago Economics Department for helpful comments, to Bryan Dayton and Tana Ryan for research assistance, and to The Sarah Scaife Foundation Fund and The Lynde and Harry Bradley Foundation Fund for financial support. Richard Craswell's comments on an earlier draft also improved the paper; I am also grateful for his and Ian Ayres's published comments.

Imaged with the Permission of Yale Law Journal

analysis deride its scientific aspirations, the steady accumulation of insights over time resembles scientific progress. Doctrinal, philosophical, and critical scholarship by contrast has been static. The authors agree or disagree, and about the same things, as much today as they did twenty or thirty years ago.

Yet there are grounds for concern about the economic analysis of contract law. Careful students of its history know that the sense of convergence ended years ago; in the last ten years, theory has become divergent, and impasses have emerged. The simple models that dominated discussion prior to the 1990s do not predict observed contract doctrine. The more complex models that emerged in the 1980s and dominated discussion in the 1990s failed to predict doctrine or relied on variables that could not, as a practical matter, be measured. As a result, the predictions of these models are indeterminate, and the normative recommendations derived from them are implausible.

For these reasons, I will argue that economic analysis has failed to produce an "economic theory" of contract law, and does not seem likely to be able to do so. By this, I mean that the economic approach does not explain the current system of contract law, nor does it provide a solid basis for criticizing and reforming contract law. This is not to say that the economic approach has not produced any wisdom, but that the nature of its accomplishment turns out to be subtle and will become clear only after an extended discussion.

This Essay has two purposes: to document the failures of economic models to explain contract law or to justify reform, and to provide an explanation for these failures. The explanation centers on the difficulty of developing a model of contractual behavior that can be tested and that does not make unreasonable assumptions about the cognitive abilities of contractual parties.

At the outset, a few comments must be made in order to avoid some possible misunderstandings of the argument. First, I will not argue that some other approach to contract law is superior to the economic approach, nor that economic analysis should be abandoned. If a moral must be extracted from the discussion, it is skepticism about how much additional value economics has to offer to understanding contract law today.

Second, I do not make claims about the value of economic analysis for understanding other areas of law. Indeed, my critique rests on empirical and methodological judgments about the contracts literature, judgments that do not necessarily apply to, say, torts or property. Nor do I take a position in

Imaged with the Permission of Yale Law Journal

this Essay on controversies over the welfarist foundations of economic analysis.[1]

Third, I want to avoid making general arguments about what counts as a good theory. One might argue that any methodology that yields surprises or insights about a familiar topic is valuable, and those surprises or insights should be counted as theories. To avoid these philosophical issues, I will focus on the original aspirations of the economic analysis of contract law: to provide an explanation of existing legal rules, and to provide a basis for criticizing or defending those rules.[2]

Finally, I want to avoid debates about what counts as "economic analysis of contract law" by stipulating that it did not exist before 1970. This is, of course, artificial. Many earlier scholars, including Holmes, Llewellyn, Hale, and Fuller, used economic analysis in the sense that from time to time they would assume that contracting parties are rational and then speculate about how different legal rules would affect these parties' incentives.[3] From a modern perspective, however, their insights seem banal, and that is because post-1970 economic analysis is more systematic and careful.[4] The interesting question is whether the post-1970 commitment to methodological individualism and the other premises of the rational actor approach provide the basis for a theory that can be used to explain or criticize contract law.

My plan is as follows. Part I describes various results from the economic analysis of contract law and compares them with the legal doctrine. In virtually every case, models make either false or indeterminate predictions about the doctrines of contract law. Part II discusses the closely related literature on incomplete contracts, a literature that attempts to

1. For a recent defense of the welfarist approach, see Louis Kaplow & Steven Shavell, *Fairness Versus Welfare*, 114 HARV. L. REV. 961 (2001).

2. This was recently described in RICHARD A. POSNER, ECONOMIC ANALYSIS OF LAW 26-29 (5th ed. 1998). As Ayres and Craswell point out, authors are more careful today about "explaining" legal rules, but there is no doubt that these authors proffer such explanations frequently, even where the normative project is emphasized. *See* ROBERT COOTER & THOMAS ULEN, LAW AND ECONOMICS 222 (3d ed. 2000) ("The courts reduce the costs of negotiating contracts by supplying efficient default terms."); Kaplow & Shavell, *supra* note 1, at 1163 (arguing that notions of fairness have not "led us seriously astray" from welfare economics, because "basic rules of damages do not seem to reflect such principles").

3. *See* BARBARA H. FRIED, THE PROGRESSIVE ASSAULT ON LAISSEZ FAIRE (1998); Avery Katz, *Reflections on Fuller and Perdue's* The Reliance Interest in Contract Damages*: A Positive Economic Framework*, 21 U. MICH. J.L. REFORM 541 (1988); Richard A. Posner, *Introduction* to THE ESSENTIAL HOLMES, at ix (Richard A. Posner ed., 1992); Alan Schwartz, *Karl Llewellyn and the Origins of Contract Theory*, *in* THE JURISPRUDENTIAL FOUNDATIONS OF CORPORATE AND COMMERCIAL LAW 12 (Jody S. Kraus & Steven D. Walt eds., 2000).

4. For defenses of the earlier work, see FRIED, *supra* note 3; and Herbert Hovenkamp, *The First Great Law & Economics Movement*, 42 STAN. L. REV. 993 (1990). Fried and Hovenkamp like the earlier work because it struggled with foundational issues and made liberal or progressive recommendations. Although the importance of this scholarship in intellectual history cannot be denied, its lack of continued vitality is almost certainly due to its failure to produce a tractable methodology.

Imaged with the Permission of Yale Law Journal

predict the content of contracts, as opposed to contract law. The separation of these two bodies of scholarship, now gradually disappearing, is an accident of history, but useful for seeing the general problems with the economic project. Part III speculates about what went wrong with economic analysis and argues that an ambiguity at the heart of the concept of transaction costs is to blame. Part IV looks at trends in contracts scholarship. Part V criticizes alternative approaches to contract theory.

I. THE ECONOMIC ANALYSIS OF CONTRACT LAW

A. *Premises and Basic Results*

The economic analysis of contract law is too familiar to warrant an extended discussion; there are also several excellent surveys.[5] Fundamental assumptions, common to nearly all efforts at economic analysis, are that individuals have preferences over outcomes, that these preferences obey basic consistency conditions, and that individuals satisfy these preferences subject to an exogenous budget constraint. Contracts scholars usually assume that individuals do not have preferences regarding the consumption or well-being of other individuals, nor regarding contract doctrine itself—there is no preference for expectation damages, for example.[6]

The standard approach assumes that the parties enter a contract in order to secure investment in a jointly beneficial project.[7] The project could be as simple as the sale of a good from Seller to Buyer—with one party (or both) enhancing the gains by an investment that reduces the cost of production for Seller or increases the value of the good for Buyer—or as complex as the construction of a skyscraper. If Buyer can increase the value of the good by making investments prior to delivery, Buyer will want a guarantee that Seller will not increase the price after Seller has observed Buyer's reliance. A contract can sometimes prevent Seller from holding up the Buyer in this way, and thus permit Buyer to invest with knowledge that he will enjoy the full return of his investment.

In their contracts, parties include terms describing performance and governing the main contingencies that affect the value of performance.

5. See the relevant entries in ENCYCLOPEDIA OF LAW AND ECONOMICS (Boudewijn Bouckaert & Gerrit De Geest eds., 2000) and THE NEW PALGRAVE DICTIONARY OF ECONOMICS AND THE LAW (Peter Newman ed., 1998). *See also* COOTER & ULEN, *supra* note 2, chs. 6-7; POSNER, *supra* note 2, ch. 4; Kaplow & Shavell, *supra* note 1, at 1102-64; Lewis A. Kornhauser, *An Introduction to the Economic Analysis of Contract Remedies*, 57 U. COLO. L. REV. 683 (1986).

6. This is not always true; scholarship on donative promises usually assumes that the promisor cares about the well-being of the promisee. *See, e.g.*, Eric A. Posner, *Altruism, Status, and Trust in the Law of Gifts and Gratuitous Promises*, 1997 WIS. L. REV. 567.

7. Another common rationale for entering a contract is risk sharing. *See* A. Mitchell Polinsky, *Risk Sharing Through Breach of Contract Remedies*, 12 J. LEGAL STUD. 427 (1983).

Imaged with the Permission of Yale Law Journal

Terms might describe the goods to be delivered, the date of delivery, or the identity of the party that bears the risk of an accident during the shipment. The terms might also release the seller from its obligation if a strike or similar event occurs. A theoretically complete contract would describe all the possible contingencies, but transaction costs—including the cost of negotiating and writing down the terms—and foreseeing low-probability events, render all contracts incomplete. In addition, parties might choose some terms or avoid others for strategic reasons, in order to exploit superior bargaining power or information asymmetries. Thus, contracts are usually quite incomplete. Parties rely on custom, trade usage, and, in the end, the courts to fill out the terms of the contract.

The terms that appear in contracts, then, depend on what the parties are trying to accomplish, shared understandings about the relevant industry, transaction costs, general characteristics of their interaction such as asymmetric information and unequal bargaining power, and the background legal regime. The last factor, the legal regime, is the focus of the economic analysis of contract law. The question is, broadly speaking, what rules of contract law would best serve the interests of the parties. This question is asked in two different ways, depending on whether the scholar takes a descriptive or a normative approach.

Descriptive analysis provides a "prediction" of contract doctrine. Built into this approach is the assumption that judges decide cases (and/or choose doctrine) in a manner that maximizes efficiency.[8] The question why judges would decide cases in this way, or whether it is necessary for them to do so in order to generate efficient law, is bracketed.[9] The author constructs a model in which parties would maximize their utility if they could enter an optimal contract. They cannot enter such a contract in the absence of legal enforcement, so the question becomes what legal rule enables the parties to enter the optimal contract. This hypothetical legal rule is then compared to

8. The models are usually written as though the concept of efficiency being used is Pareto efficiency: The decisionmaker chooses the rule that maximizes the surplus from cooperation, and, although this might involve the Kaldor-Hicks idea of transferring goods from the person who values them less to the person who values them more, all people who use contracts are better off with an efficient system because prices will reflect the risk of ex post transfers. To be sure, in many cases prices will not adjust, and the transition from an inefficient rule to an efficient rule would likely be a Kaldor-Hicks move, but these will usually be minor considerations.

9. A literature that analyzes this assumption is inconclusive. *See, e.g.*, George L. Priest, *The Common Law Process and the Selection of Efficient Rules*, 6 J. LEGAL STUD. 65 (1977); Paul H. Rubin, *Why Is the Common Law Efficient?*, 6 J. LEGAL STUD. 51 (1977). There is no reason why the prediction must be that contract law is efficient; this seems to be an accident of intellectual history. One could imagine a different theory, along the lines of public choice, that holds that contract law reflects the self-interested decisions of judges to implement policy preferences. Indeed, such an approach has been used by political scientists to explain judicial interpretation of statutes and constitutional provisions. *See, e.g.*, JEFFREY A. SEGAL & HAROLD J. SPAETH, THE SUPREME COURT AND THE ATTITUDINAL MODEL (1993).

Imaged with the Permission of Yale Law Journal

actual legal rules, and, if they are the same, the descriptive hypothesis is vindicated.

The normative position assumes that contract law should be efficient. As before, the author constructs a model in which parties can increase their welfare through a contract that is legally enforceable. The author first shows the optimal outcome—where, for example, performance occurs only when the buyer's valuation exceeds the seller's cost, and buyer and seller make efficient investments—and then the equilibrium outcomes under alternative legal rules. Typically, the author recommends one rule as efficient, or shows that different rules are efficient under different assumptions, or else criticizes various existing rules because they do not enable the parties to achieve the optimal outcome.

In the following Sections, I will show the ways in which contract doctrine diverges from the predictions of the descriptive hypotheses, and I will show that the normative implications of the models are weak or nonexistent. The reason for discussing normative and descriptive failures at the same time is that the two are closely connected. From a descriptive perspective, the models generate either false or indeterminate predictions. From a normative perspective, the models generate either implausible or indeterminate recommendations. The reason in both cases is that the determinate models omit important variables, but including these variables makes them indeterminate, or, in some cases, unrealistic, because they place too great a burden on courts. The nature and origin of these difficulties will become clearer as we examine the models.

B. *Remedies*

Much contract doctrine comprises background rules that parties can change, albeit within limits. The victim of breach, by default, receives expectation damages, but the parties can vary this outcome ex ante by providing for liquidated damages in the contract. Their ability to contract around the expectation damages rule in this way is circumscribed by the penalty doctrine, which forbids liquidated damages that are unreasonably high.

At an early stage, scholars argued that the default rule should maximize the ex ante value of the contract. Expectation damages were said to have this effect as a result of an attractive property: They give a party the incentive to breach if and only if the cost of performance for the promisor exceeds the value of performance for the promisee. Performance occurs if

Imaged with the Permission of Yale Law Journal

and only if it is efficient. For this reason, expectation damages seemed to be the right measure of damages.[10]

This conclusion was premature, however. First, the argument overlooks the ability of the parties to renegotiate prior to performance. If renegotiation costs are low enough, efficient performance will occur regardless of the remedy. If the remedy is less than expectation damages and performance is efficient, the promisee will bribe the promisor to perform. If the remedy is greater than expectation damages and performance is inefficient, the promisor will pay the promisee for a release.

Second, the argument overlooks the effect of the expectation measure on other incentives. Consider the promisee's incentive to rely or invest in anticipation of performance. Under the rule of expectation damages, the promisee's reliance investment is fully compensated. But if the promisee expects to recover the investment regardless of whether or not trade is efficient, the promisee will overinvest—that is, he will invest as though the return were certain rather than stochastic, externalizing the cost on the promisor.[11] A superior measure of damages would give the promisee the amount of damages that would compensate the promisee if he engaged in efficient reliance, not the amount that would compensate the promisee for the loss given whatever level of reliance was taken.[12]

The concept of efficient investment is subtle, and a numerical example might help. Suppose that Buyer and Seller enter a contract under which Seller promises to supply goods that Buyer needs for his factory. Buyer can increase the value of the goods for his use by investing in adjustments to the factory prior to delivery. Let's say that if Buyer invests 0, his valuation of the goods equals 100. If Buyer invests 5, his valuation of the goods equals 120. If Buyer invests 10, his valuation of the goods equals 128. If Buyer will obtain the goods with certainty, then efficiency requires that he invest 10: $128 - 10 > 120 - 5 > 100 - 0$. However, if Buyer will obtain the goods with only a 50% probability, then efficiency requires that he invest 5: $0.5(120) - 5 > 0.5(128) - 10$, and $0.5(120) - 5 > 0.5(100) - 0$. A person who invests money in some outcome will invest more if the outcome is certain than if the outcome is uncertain. Because expectation damages provide a return to the promisee whether or not breach is efficient, the promisee will invest as though the yield of the investment would occur with probability of 1 rather than with the probability (<1) that performance

10. John H. Barton, *The Economic Basis of Damages for Breach of Contract*, 1 J. LEGAL STUD. 277 (1972); Robert L. Birmingham, *Breach of Contract, Damage Measures, and Economic Efficiency*, 24 RUTGERS L. REV. 273 (1970).

11. *See* William P. Rogerson, *Efficient Reliance and Damage Measures for Breach of Contract*, 15 RAND J. ECON. 39 (1984); Steven Shavell, *The Design of Contracts and Remedies for Breach*, 99 Q.J. ECON. 121 (1984).

12. *See* Robert Cooter, *Unity in Tort, Contract, and Property: The Model of Precaution*, 73 CAL. L. REV. 1, 14 (1985).

Imaged with the Permission of Yale Law Journal

occurs. The promisee thus invests an amount greater than would be efficient.[13]

Third, the argument neglects the ability of the parties to design remedial provisions for their contract. If expectation damages are optimal, the parties can achieve the effect of this remedy by giving each side the option to perform or pay an amount that is the function of revealed ex post values. If expectation damages are not optimal, then the parties can choose some superior remedy that would, for example, take account of reliance incentives. These considerations suggest that specific performance of the remedial portion of the contract would be efficient, not expectation damages, which in essence convert the obligation to perform into an option to perform or pay an amount determined by a court.

There are numerous other problems with expectation damages. Expectation damages are also undesirable if courts have trouble determining the parties' valuations at the time of breach. The better remedy is specific performance, which a court can award without determining the promisee's valuation.[14]

Expectation damages are also undesirable when information is asymmetric, unless highly specific conditions are met. Consider the *Hadley* rule, according to which a victim of breach obtains compensation for average, rather than actual, loss, unless he has revealed his valuation to the promisor ex ante.[15] Thus, the shipper cannot recover fully compensatory damages from a carrier who has breached the shipment contract if the shipper does not reveal the specially high value of the goods shipped. The *Hadley* rule gives the shipper an incentive to disclose his valuation prior to contracting, so that the carrier will take optimal precautions given the shipment's value.

But it turns out that the argument can be reversed. Imagine an expansive liability rule that gave the victim of breach actual damages (that is, expectation damages). The defense of *Hadley* implicitly assumed that under the expansive liability rule the high-value shipper would not have an incentive to reveal his valuation: If he is to be fully compensated, he has no

13. For another example, see A. MITCHELL POLINSKY, AN INTRODUCTION TO LAW AND ECONOMICS 32-35 (1983).

14. Kronman argues that the common law efficiently reserves specific performance for disputes involving valuation problems such as those involving unique goods. *See* Anthony Kronman, *Specific Performance*, 45 U. CHI. L. REV. 351 (1978). Schwartz points out that information problems about valuation, enforcement, and so forth are always present, and therefore specific performance should be the default rule. *See* Alan Schwartz, *The Case for Specific Performance*, 89 YALE L.J. 271 (1979); *see also* Thomas S. Ulen, *The Efficiency of Specific Performance: Toward a Unified Theory of Contract Remedies*, 83 MICH. L. REV. 341 (1984). The two remedies also have different effects on reliance incentives. *See* Shavell, *supra* note 11. But, the simplest defense of specific performance is that if parties are rational, they will design an optimal contract, and courts should enforce their terms rather than give the parties an option (expectation damages) when they did not bargain for it.

15. Hadley v. Baxendale, 156 Eng. Rep. 145 (Ex. 1854).

Imaged with the Permission of Yale Law Journal

reason to reveal his valuation, which would enable the carrier to charge a higher price. But the expansive liability rule does give the low-value shipper the incentive to reveal his valuation. If he does not, he will be charged ex ante for average compensation, but he would prefer to be charged a lower price, even if this means that the carrier will take less care. If the low-value shippers reveal their valuation, then the carrier can infer that any shipper that does not reveal his valuation must have a high valuation. Both the *Hadley* rule and its opposite give parties incentives to disclose private information.

Authors who have pursued this argument point out that one rule could be better than the other, depending on the distribution of valuations, the cost of revealing information, the relative bargaining power of the party with private information and the uninformed party, and related factors. If there are more low-value shippers than high-value shippers, the expansive liability rule requires more bargaining around, and therefore more transaction costs, and thus might be suboptimal.[16] But the relevant variables are too complex and too hard to determine. We do not observe doctrine incorporating them, nor do we have enough empirical data to be able to guess which rule is based on assumptions that are closer to reality.[17]

---

16. *See* Ian Ayres & Robert Gertner, *Filling Gaps in Incomplete Contracts: An Economic Theory of Default Rules*, 99 YALE L.J. 87 (1989) [hereinafter Ayres & Gertner, *Filling Gaps*]; Ian Ayres & Robert Gertner, *Strategic Contractual Inefficiency and the Optimal Choice of Legal Rules*, 101 YALE L.J. 729 (1992); Lucian Arye Bebchuk & Steven Shavell, *Information and the Scope of Liability for Breach of Contract: The Rule of* Hadley v. Baxendale, 7 J.L. ECON. & ORG. 284 (1991); Charles J. Goetz & Robert E. Scott, *Enforcing Promises: An Examination of the Basis of Contract*, 89 YALE L.J. 1261, 1299-300 (1980); Jason Scott Johnston, *Strategic Bargaining and the Economic Theory of Contract Default Rules*, 100 YALE L.J. 615 (1990); Louis E. Wolcher, *Price Discrimination and Inefficient Risk Allocation Under the Rule of* Hadley v. Baxendale, *in* 12 RESEARCH IN LAW AND ECONOMICS 9 (Richard O. Zerbe ed., 1989).

17. For further epicycles, see Barry E. Adler, *The Questionable Ascent of* Hadley v. Baxendale, 51 STAN. L. REV. 1547 (1999) (showing that the results of earlier models change if the high type is likely, rather than certain, to suffer a large loss the event of breach). Adler overstates his argument as a critique of *Hadley v. Baxendale* when, in fact, he just shows that courts must take into account yet another factor when determining the optimal rule. More to the point is his skepticism about the possibility that lawmakers could take into account the factors that he identifies when formulating doctrine. *Id.* at 1582. As Bebchuk and Shavell observe in their reply, "Adler does not note any reasons for assuming that the consideration that he discusses involves less practical problems for lawmakers than the considerations on which our analysis has focused." Lucian Arye Bebchuk & Steven Shavell, *Reconsidering Contractual Liability and the Incentive To Reveal Information*, 51 STAN. L. REV. 1615, 1627 (1999). Adler does seem to realize that the accumulating complexities of the analysis undermine its practical value for lawmakers. Bebchuk and Shavell, by contrast, state:

> [O]ur analysis of *Hadley* enables one to recommend that rule with greater confidence than researchers are often able to endorse other legal rules in other contexts. As we explained in some detail, it seems that the *Hadley* rule is clearly desirable for cases (such as *Hadley* itself) in which a minority of buyers has valuations of performance that are substantially higher than the valuations of ordinary buyers.

Id. at 1625. But they do not provide a reason for believing that any of the relevant factors are measurable in general conditions, and one cannot evaluate their historical claim without further evidence.

Imaged with the Permission of Yale Law Journal

There are other considerations as well. The remedy that is chosen will affect the incentive of each party to search for the optimal partners prior to contracting, to reveal private information about the probability that performance will be possible, to take precautions against breach, and to renegotiate after information is revealed about the state of the world.[18] Remedies will also affect the ability of the parties to shift risk in a contract when one or both parties are risk-averse. And, as I discuss below, remedies affect the ability of the contracting parties to take advantage of third parties who come onto the scene after the parties have entered the contract and value performance more than either of the contracting parties.

Articles that discuss these various incentives typically bracket most of them for the purpose of analysis and focus on one or two. As a result, the optimal remedy derived from a model is optimal only under narrow conditions. If we are to put the models together and try to draw from them as a group their prediction about contract law, we could take two approaches.

First, we could argue that the models collectively show that different remedies are optimal under different conditions and therefore predict that contract law should incorporate these conditions in doctrine. For example, contract law will make expectation damages the remedy when the parties can make choices only about breach or performance and not about how much to invest. But there are two problems with this approach. The first of these problems is that contract law does not resemble the predictions of the models. Awarding expectation damages is the general rule in contract law, but this rule can be justified by the models only under narrow conditions. Furthermore, doctrine does not make the application of expectation damages turn on variables identified by the models, such as the degree of reliance by the promisee. The second of these problems is that the models taken together are probably indeterminate. To generate predictions, one would need a vast amount of information about the characteristics of the parties and the transactions. If one remedy is best when renegotiation costs are high, and another is best when renegotiation costs are low, we need some way to measure renegotiation costs. If the optimal remedy depends on the shape of probability distributions for sellers' costs and buyers' valuations, we need this information as well. Yet no one has attempted to collect this information, and it is difficult to imagine how this task could be accomplished.

Under the second approach, we could argue that the models collectively show that one particular remedial structure—the existing doctrine of contract law—is optimal given the "average" circumstances of the parties.

18. For a clear discussion, see Richard Craswell, *Contract Remedies, Renegotiation, and the Theory of Efficient Breach*, 61 S. CAL. L. REV. 629 (1988).

Imaged with the Permission of Yale Law Journal

We might think, for example, that on average pre-performance investment is not a significant issue, or, if it is, it is adequately controlled by the doctrine of mitigation.[19] The rule of expectation damages is optimal because the perform-or-breach decision matters most, with specific performance reserved for cases where valuation problems are insurmountable. But this view is unsupported by any evidence.

C. *Contract Interpretation*

Many contract disputes turn on questions of interpretation. Seller delivers the goods, but Buyer argues that the goods do not conform to the requirements of the contract. Suppose the contract says "chicken," and the delivery is a scrawny, stewing chicken. Buyer says that "chicken" refers to a plump, juicy broiler; Seller says that the word just identifies the species and leaves the quality of the bird to Seller's discretion.[20] How should the court resolve this dispute?

Economists have proposed a number of interpretive strategies for courts.[21] One is to choose a "majoritarian default," the meaning that most parties to chicken contracts would use, which will often be the same as the customary meaning or trade usage. If parties expect that courts will apply a majoritarian default when disputes arise over the meaning of the contract, they will know that most of the time the court will choose the term that maximizes the probability of efficient trade. Accordingly, they would be more willing to enter a contract in the first place, despite high transaction costs, than they would under an alternative rule. Choosing a majoritarian default rule reduces the negative consequences of high transaction costs.

Another strategy is to choose a "penalty default," a meaning that most parties to chicken contracts would *not* use.[22] This strategy, which would give parties an incentive to write a less ambiguous contract than they might otherwise, has two motivations. First, it discourages parties from externalizing the cost of interpreting the contract on the courts. If parties were clearer, courts would have less work to do. Second, it discourages parties from opportunistically concealing information from each other. If one party knows about the ambiguity of the word "chicken" and prefers the majoritarian meaning, and the other party does not know about the ambiguity, then the first party would have no incentive to disclose the

19. *See* Charles J. Goetz & Robert E. Scott, *The Mitigation Principle: Toward a General Theory of Contractual Obligation*, 69 VA. L. REV. 967 (1983).

20. *Cf.* Frigaliment Importing Co. v. B.N.S. Int'l Sales Corp., 190 F. Supp. 116 (S.D.N.Y. 1960) (holding that the broad meaning of "chicken" is correct).

21. *See, e.g.*, Charles J. Goetz & Robert E. Scott, *The Limits of Expanded Choice: An Analysis of the Interactions Between Express and Implied Contract Terms*, 73 CAL. L. REV. 261, 321 (1985).

22. Ayres & Gertner, *Fillings Gaps*, *supra* note 16, at 95.

Imaged with the Permission of Yale Law Journal

ambiguity to the second, unless a penalty default rule held the informed party to the less favorable meaning.

A third strategy is to enforce the contract in a literalistic way. If the party says "chicken" and the dictionary or common sense definition of "chicken" has a general meaning, then Seller has the right to deliver the stewing chicken. The court does not try to determine what most parties mean by "chicken," or what most parties do *not* mean. This strategy, like the penalty default strategy, gives the parties an incentive to be clear, or at least to anticipate how courts normally interpret terms.

A final strategy is for the court to enforce whatever term would be efficient in the particular case. One can derive this term by asking the question: Supposing that transaction costs had been zero at the time of contracting, what would the parties have done? Buyer and Seller would have anticipated their dispute about the meaning of "chicken" and either chosen a more precise term (if trade is still efficient) or not made a deal (if trade is not efficient). What they would have done depends on the costs and values of the various birds. The difference between this strategy and the majoritarian default is the difference between a standard and a rule. The court chooses whatever is efficient for the contract in dispute, rather than enforcing whatever term is efficient for the majority of parties who enter similar or identical contracts.

We have already examined a model comparing the first and second strategy, namely, Ayres and Gertner's model of the *Hadley* rule.[23] The *Hadley* rule, in Ayres and Gertner's argument, plays the role of a penalty default, for they assume that a majority of buyers prefer unlimited liability, which would thus serve as a majoritarian default. Choosing between limited liability and unlimited liability when the contract does not specify one or the other is like choosing between the ordinary meaning of chicken and a narrow meaning of chicken when the contract does not define the term. The choice between these two meanings depends on the same factors that determine the efficiency of the *Hadley* rule: the cost of bargaining around the default rule, the distribution of valuations in the population of buyers, the market power of the seller, the degree to which the seller's performance would improve with superior information, and other factors that are not likely within the grasp of a decisionmaker. Thus, the indeterminacy that afflicts the *Hadley* analysis undermines any effort to choose between a majoritarian and penalty default.

For this reason, one might argue that courts should engage in literalistic enforcement. Indeed, Schwartz makes just such an argument, claiming that the responsibility for choosing default rules puts an unrealistically high

23. *Id.* at 101.

Imaged with the Permission of Yale Law Journal

informational burden on the courts.[24] Literalism, by contrast, allows parties to direct courts to enforce obligations that arise under conditions that the courts can verify. But although it is true that literalism does put a lighter burden on courts, it does not follow that literalism is superior to the majoritarian (or penalty) approach. The choice between the two approaches is, as Schwartz acknowledges, an empirical question about which we have no evidence.[25] The most significant problem with Schwartz's analysis, however, is that it depends on the methodological assumption that cognitive limitations do not exist or are minimal. The majoritarian approach depends on the assumption that parties fail to anticipate the future; Schwartz simply assumes the opposite.

This point can also be made about Schwartz's criticism of the view that courts should choose the term that is most efficient in the particular case. Schwartz argues that if the evidence necessary to choose such terms ex post is verifiable, then parties will bargain to the efficient result, in which case judicial intervention is not necessary.[26] In our example, the parties will trade the chicken only if the buyer values it more than the seller does, so that if the buyer accepts the stewing chicken, the ex post interpretation of the contract is effectively that the general meaning of chicken holds. If the evidence is not verifiable, and indeed not observable either, they might bargain to an impasse, or to an inefficient term, in which case courts cannot help. However, if the parties are boundedly rational—again, outside Schwartz's model—we do not know how they would bargain with each other, and therefore whether a court could improve on the outcome.

Let me summarize. From a descriptive perspective, we can distinguish two bodies of work. The standard economic analysis of default rules is broadly consistent with judicial practices; courts employ a mix of majoritarian and penalty defaults. But it does no more than rationalize these practices, for there is no way to measure the variables that determine the relative efficiency of the rules. Schwartz's argument, which is simpler and truer to economic premises, fails to account for courts' refusal (for the most part) to rely on the literalistic approach.[27]

From a normative perspective, Schwartz's argument that courts should engage in literalistic interpretation should appeal to those steeped in

24. Alan Schwartz, *The Default Rule Paradigm and the Limits of Contract Law*, 3 S. CAL. INTERDISC. L.J. 389, 416 (1993); Alan Schwartz, *Incomplete Contracts*, *in* 2 THE NEW PALGRAVE DICTIONARY OF ECONOMICS AND THE LAW, *supra* note 5, at 277, 280 [hereinafter Schwartz, *Incomplete Contracts*]; *see also* Robert E. Scott, *The Case for Formalism in Relational Contract*, 94 NW. U. L. REV. 847, 875 (2000).

25. Schwartz, *Incomplete Contracts*, *supra* note 24, at 280.

26. *Id.* at 282.

27. *Cf.* Alan Schwartz, *Relational Contracts in the Courts: An Analysis of Incomplete Agreements and Judicial Strategies*, 21 J. LEGAL STUD. 271 (1992). In this earlier article, Schwartz argued that courts interpret contracts aggressively when bargaining defects exist and when the interpretation can be justified on the basis of verifiable information.

Imaged with the Permission of Yale Law Journal

economics, but the appeal derives from the methodological decision to treat individuals as rational and courts as hampered by information asymmetries.[28]

D. *Unconscionability and Consumer Protection*

The premises of economics push in the direction of freedom of contract, and this current can be resisted only with difficulty. If parties are rational, they will enter contracts only when it is in their self-interest, and they will agree only to terms that make them better off. Courts that refused to enforce these terms would make it more difficult for future parties to use contracts to enhance their joint well-being. Therefore, courts should enforce the terms of the contract.

And yet courts do not always enforce the terms of contracts. They often refuse to enforce terms that seem harsh, oppressive, or improper: strict liquidated damages provisions, expansive security arrangements, alienation of the equity of redemption, restrictive arbitration provisions, broad covenants not to compete, wagers, choice-of-forum clauses and disclaimers of warranties in fine print or confusing language, and even price terms that seem too high or too low. Some of these practices derive from statutes (for example, usury laws); others arose in equity or the common law. The catchall term is unconscionability, but the relatively unusual application of this doctrine by courts only deflects attention from the widespread judicial scrutiny of transactions involving consumers, much of it in the form of interpretive presumptions that can interfere as much with freedom of contract as prohibitions do.

Economics has been better at deflating standard explanations for unconscionability and related doctrines than at explaining these doctrines. Let me say a few words about these standard explanations.

1. *Unequal Bargaining Power*

Courts sometimes say that a contract is unconscionable because of the unequal bargaining power of the seller and buyer. It is not always clear what courts mean when they use this term, but the closest economic concept is that of market or monopoly power. A seller has market power if it can increase the price of the good above its marginal cost by restricting supply. As is well known, such behavior is inefficient in the Kaldor-Hicks sense, and forcing the seller to sell at marginal cost would in theory eliminate a deadweight cost.

28. For a parallel argument, see *infra* Section II.B.

Imaged with the Permission of Yale Law Journal

Nonetheless, economists typically argue that courts should not avoid contracts because of the unequal bargaining power of the parties. When contracts appear to have very high price terms, a court could determine only with great difficulty whether the high price is due to market power or fluctuations in the costs of inputs. A high interest rate, for example, could result from the creditor's judgment about the risk of default posed by a particular debtor, and generally courts should defer to such judgments. A determination that the creditor has market power requires an evaluation of the structure of the market, a notoriously difficult enterprise usually reserved for antitrust litigation. A seller or creditor with temporary market power as a result of a patent, or some innovation that other market participants have not had a chance to imitate, should (arguably) be permitted to reap above-market returns, for that is how innovation is encouraged in a market economy.

When contracts appear to have harsh nonprice terms, there is another reason for thinking that these terms are unobjectionable. Even if the seller or creditor has market power, it has the right incentive to supply the terms that parties desire. For example, a debtor might be willing to consent to a harsh remedial term in return for a low interest rate.[29] And a supplier might be willing to give the buyer the power to terminate the contract with little notice, if that is the only way to get the buyer's business. The party with market power will supply terms if the other parties want them and will charge them a fee, but will not force terms on parties that do not want them, for generally the most efficient way to exploit market power is through the price term.[30] Although there are models in which a combination of market power and asymmetric information can result in inefficient terms, they justify nonenforcement only under complex and hard-to-identify conditions.[31]

These theories do not describe what courts do. Courts permit the harshness of nonprice, and occasionally price, terms to influence them, and they seem to attach significance to unequal bargaining power. For this reason, most economic work is cast as a normative critique of the judicial practice.

29. Richard A. Epstein, *Unconscionability: A Critical Reappraisal*, 18 J.L. & ECON. 293, 305-06 (1975).

30. Alan Schwartz, *A Reexamination of Nonsubstantive Unconscionability*, 63 VA. L. REV. 1053, 1071-76 (1977); Alan Schwartz & Louis L. Wilde, *Imperfect Information in Markets for Contract Terms: The Examples of Warranties and Security Interests*, 69 VA. L. REV. 1387, 1458 (1983).

31. *See* Alan Schwartz & Louis L. Wilde, *Intervening in Markets on the Basis of Imperfect Information: A Legal and Economic Analysis*, 127 U. PA. L. REV. 630, 666-71 (1979). For a discussion of the literature in the context of consumer finance, see Richard Hynes & Eric A. Posner, *The Law and Economics of Consumer Finance*, 4 AM. L. & ECON. REV. 168 (2002).

Imaged with the Permission of Yale Law Journal

2. *Lack of Information*

Courts sometimes say that a contract is unconscionable because one party, usually a consumer, lacks sophistication. Lack of sophistication is not the same thing as lack of information, but lack of information does seem to play a role in the cases. When terms are harsh and complex or hard to read, and consumers are unsophisticated, courts often express doubt that the consumers understood their obligations under the contract. This has led economists to investigate the role of information deficiencies in contract enforcement.

The topic is too complex to discuss here in any detail, but let me make a few observations. Consumers who lack information have incentives to acquire information. Some consumers will acquire information more easily than others; these are the people who read *Consumer Reports*. But the other consumers can free-ride on the efforts of the first group. If sellers cannot easily distinguish informed and uninformed consumers, they cannot exploit the latter by charging them a higher price. Thus, information deficiency alone does not justify judicial intervention.[32]

In addition, sellers have incentives to provide information to otherwise uninformed consumers. If seller *X* has lower costs than seller *Y*, and thus can charge lower prices and obtain a profit, *X* will invest in advertising in order to attract consumers from *Y*. There are limits, however, to the amount of information *X* will provide. If *X*'s cars are cheaper than *Y*'s cars, *X* has the right incentives; but if *X* knows that its cars in general are more dangerous than consumers believe, *X* has no incentive to provide that information.[33] Supplying such information is costly, both intrinsically and in the form of lost sales, and *X* does not internalize the benefits when he honestly warns of the dangers of automobile travel and consumers refrain from buying cars and avoid being injured.[34]

3. *Summary*

In sum, a simple model of the consumer-goods market implies that courts should not use the unconscionability doctrine to strike down contracts. More complex models suggest that courts should ignore bargaining power or should take it into account only under narrow conditions. Yet courts frequently criticize the inequality of bargaining

32. Schwartz & Wilde, *supra* note 30, at 1422-23.

33. *See* Howard Beales et al., *The Efficient Regulation of Consumer Information*, 24 J.L. & ECON. 491, 505-06 (1981).

34. *X*'s incentives are suboptimal even if cars are safer, rather than more dangerous, than consumers think because *X* would not internalize gains to *Y* that would result if *X* revealed this information to consumers. Monopolists might gain more from information disclosure than competitors, but there are further complications. *See id.* at 507-08.

Imaged with the Permission of Yale Law Journal

power between consumer and seller, imply that this fact may justify avoidance of the contract, and do not elaborate any further on the question of why bargaining power matters in some cases but not others. Other models suggest that courts might improve information asymmetries when consumers do not engage in enough comparison shopping or when competitive pressures do not force sellers to reveal information. Yet courts rarely pay attention to these factors when applying the unconscionability doctrine.

E. *Mistake*

In some circumstances, courts avoid contracts that are the result of mistakes. If the parties committed a mutual mistake as to a basic assumption of the contract, or if one party committed a mistake that the other party could have detected, the adversely affected party will sometimes have the right to avoid the contract.

In theory, parties could design contracts that released one or both parties who made a mistake. Consider a contract between Buyer and Seller for the sale of a cow.[35] Buyer and Seller might believe that the cow is barren when in fact she is fertile, in which case Seller will want to avoid the contract. Or Buyer and Seller might believe that the cow is healthy when in fact she is ill, in which case Buyer will want to avoid the contract. In either event, the parties can design the contract accordingly. The parties could enter a contract giving the Seller the right to withdraw from the contract if the cow proves to be fertile, or make performance contingent on subsequent confirmation that the cow is barren. And in the second case, Seller could give Buyer a warranty against illness or not, depending on how they want to allocate the risk. The general point is that if parties are rational, they know that they can make mistakes, and they will design the contract in a way that assigns this risk in the appropriate manner.

One might respond that because the parties are, by hypothesis, mistaken, it does not occur to them to build these contingencies into the contract. This is what courts mean when they say that the mistake was about a basic assumption of the contract. But rational parties always know that something could happen that makes performance more or less costly to Seller, and more or less valuable to Buyer. It could be that the cow has a hidden characteristic, good or bad; it could be that market conditions will change, so that a cow gains or loses value relative to other goods. From an economic perspective, there is nothing special about the cow being fertile or ill, nothing that distinguishes this contingency from a change in the price

35. Sherwood v. Walker, 33 N.W. 919 (Mich. 1887).

Imaged with the Permission of Yale Law Journal

caused by a shift in market conditions. Parties can design contracts that take account of all these contingencies.

If this argument is correct, there is no reason for courts to release parties when one or both of them make mistakes. It would be like releasing an insurance company from a fire-insurance contract on the ground that the insurance company mistakenly believed that a fire would not occur. From an economic perspective, parties cannot make mistakes: They have probability distributions that reflect information they have about the world. They know that they do not possess the absolute truth and would not believe otherwise.[36]

In order to explain the mistake doctrines, then, we need to make additional assumptions. One possible assumption, which by now should be familiar, is that "transaction costs" prevent parties from designing optimal contracts. This is the implicit route taken by Rasmusen and Ayres in an article on the mistake doctrines.[37] Before we turn to their argument, we should observe that using this assumption makes the analysis of the mistake doctrine the same as the analysis of any problem of contractual interpretation, where, as law and economics assumes, transaction costs prevent parties from defining a crucial term, like "chicken" in the *Frigaliment* case. The parties do not make a "mistake" in the ordinary sense of the term; they rationally choose to leave a contract incomplete in light of the costs of completing it. If one thinks that courts should use majoritarian defaults to determine such terms, then one should think that majoritarian defaults should also determine the parties' obligations if the cow is fertile or ill. On this view, the mistake doctrine should also have the same remedial implications as contract-interpretation disputes—namely, enforcement of the judicial interpretation rather than rescission and restitution—but of course it does not.

To evaluate the mutual and unilateral mistake doctrines, Rasmusen and Ayres assume that parties can set a price but that they cannot make performance contingent on the occurrence of the desired states of the world (either directly or through the use of an optimal incomplete contract). It is in this sense that they operationalize the concept of mistake, taking their cue from the work on contract interpretation. Seller expects an average cost to perform, c, and Buyer expects an average valuation, v, such that $v > c$, on average, but in some states of the world $v < c$, and trade should not occur. Ayres and Rasmusen investigate the question whether the mistake doctrine should release the parties from the contract (presumably, at the request of

36. Eric Rasmusen & Ian Ayres, *Mutual and Unilateral Mistake in Contract Law*, 22 J. LEGAL STUD. 309, 315 & n.13 (1993).

37. *Id.* at 315.

Imaged with the Permission of Yale Law Journal

Buyer if the price is higher than his realized valuation, or of Seller if the price is lower than her realized cost).

Their analysis is complex, so let me focus on their conclusions for mutual mistake. The doctrine, by excusing performance when both parties are uninformed or mistaken, would seem to create incentives for parties to avoid performance and to gather information. The question is whether these incentives are efficient.

One possible advantage of the mutual mistake doctrine is that it enables the parties to avoid a second transaction in order to reverse the initial contract when it turns out that $v < c$. But the reversal will occur only when both parties are also uninformed; if one party is informed, then the other party cannot avoid an inefficient contract by claiming mutual mistake. Further, the doctrine enables either party to avoid the contract when (despite the mistake) $v > c$: Buyer will avoid the contract when the price $p > v$, and Seller will avoid the contract when $p < c$.[38] If you permit parties to breach without paying damages, then inefficient breach will result.

The mutual mistake doctrine also affects the parties' incentives to gather information prior to entering contracts. But it does not affect the incentives in a desirable way. Suppose that parties can acquire information about the value of the good at some cost. If the acquisition of information does not increase the value of the good, and the cost of acquiring information is high enough, the mutual mistake doctrine properly encourages the parties to remain ignorant but also results in too many rescissions. When the cost of information is lower, the mutual mistake doctrine encourages each party to acquire information in order to prevent the other party from invoking mutual mistake, even though the additional information does not increase the value of the good. On the other hand, if the acquisition of information *does* increase the value of the good, then the mutual mistake doctrine gives too little incentive to acquire information, compared to a rule of enforcement.[39]

The mutual mistake doctrine is hard to reconcile with economic premises, both because the doctrine by its terms appeals to cognitive errors excluded from economic analysis, and because the doctrine encourages behavior that would be suboptimal if people did not make those errors, as economics assumes.[40] But even putting aside cognitive errors, Ayres and

38. *Id.* at 320.

39. *Id.* at 331-32.

40. Indeed, most contracts scholarship on information asymmetries assumes that the parties know about their information advantage or disadvantage and make a strategic decision to reveal information, demand a price adjustment, refuse to enter a contract, and so forth. The articles usually focus on the proper damages remedy given these assumptions. In this context, Rasmusen and Ayres's argument can be understood as an investigation of the conditions under which a "zero damages" rule would be superior to expectation damages and alternative rules. For examples of this literature, see Richard Craswell, *Performance, Reliance, and One-Sided Information*, 18 J.

Imaged with the Permission of Yale Law Journal

Rasmusen do not offer a clear replacement for the mutual mistake doctrine. They show that other doctrines, such as no excuse and unilateral mistake, dominate mutual mistake in different contexts, but not that one of these doctrines or some alternative would be optimal in general, nor that a court could distinguish the conditions—whether, for example, acquisition of information would increase the value of a good or not—under which the different rules have advantages.

F. *Impossibility*

Courts sometimes release promisors from performance when performance is "impossible" or "impracticable." Posner and Rosenfield argue that these doctrines efficiently shift risk from the promisor when the promisor is more risk-averse than the promisee.[41] Suppose a seller cannot insure itself against a strike by its workforce, but that the buyer can easily arrange for deliveries from alternative sellers if supply from the first is cut off. If the seller subsequently cannot make deliveries because of a strike, a court might excuse the seller from its contractual obligations on the grounds of impossibility, with the real reason being that buyer could have insured against this contingency more easily than the seller could have.

Subsequent work casts doubt on this argument. First, Posner and Rosenfield's argument neglects the other incentives of the parties. If the seller pays no damages or a limited amount like restitution, it has no incentive to perform when it is efficient to do so. The argument assumes that the court can determine whether the cost of performance exceeds its value to the buyer. But in other contexts, the justification of expectation damages, for example, it is assumed that the court cannot make this determination.[42]

Second, and more important, the impossibility and impracticability doctrines do not spread risk in the efficient way. To see why, imagine a risk-averse seller and a risk-neutral buyer. The optimal resolution of a

---

LEGAL STUD. 365 (1989) (analyzing the effect of damages rules on the promisor's incentive to disclose private information about the probability of performance); Richard Craswell, *Precontractual Investigation as an Optimal Precaution Problem*, 17 J. LEGAL STUD. 401 (1988) (analyzing the effect of damages rules on the parties' incentives to acquire information, prior to contracting, about their ability to perform); Peter A. Diamond & Eric Maskin, *An Equilibrium Analysis of Search and Breach of Contract, I: Steady States*, 10 BELL J. ECON. 282 (1979) (analyzing the effect of damages rules on parties' incentives to search for contract partners); and Steven Shavell, *Acquisition and Disclosure of Information Prior to Sale*, 25 RAND J. ECON. 20 (1994) (analyzing the effect of disclosure rules on the incentive to acquire and disclose information).

41. Richard A. Posner & Andrew M. Rosenfield, *Impossibility and Related Doctrines in Contract Law: An Economic Analysis*, 6 J. LEGAL STUD. 83 (1977).

42. *See* Michelle J. White, *Contract Breach and Contract Discharge Due to Impossibility: A Unified Theory*, 17 J. LEGAL STUD. 353, 360 (1988).

Imaged with the Permission of Yale Law Journal

dispute will place all the risk on Buyer; in effect, the court writes an insurance policy of which Seller is the beneficiary. Such an insurance policy would give Seller the same payoff in both states of the world—the breach state and the performance state. The impossibility and impracticability doctrines, however, do no such thing. They, at best, give Seller a zero payoff in the breach state (and less if Seller has incurred some costs or must make restitution), instead of giving Seller some amount between zero and its profits. Indeed, the risk-sharing argument implies that in some cases Seller should pay negative damages, something which is, of course, never observed.[43]

The excuse doctrines are hard to understand from the economic perspective. Sophisticated parties know that contingencies might occur that will make performance impossible or extremely costly. If they want to share the risk of these contingencies, they can write excuses into the contract. We observe this behavior not just in the use of force majeure clauses; excuses are frequently built into the central terms of the contract. Insurance contracts contain exclusions; ordinary sales contracts shift risk by tying the price to market indices. Firms can buy general insurance policies, or self-insure, and are usually risk-neutral with respect to run-of-the-mill contracts. It might be true that the cost of describing the parties' obligations prevents parties from assigning all the risks, but it remains doubtful that courts have the information necessary to repair the insurance market.

G. *Consideration and Promissory Estoppel*

Economics assumes that people exchange promises when both benefit from the exchange, but it does not follow that the law should enforce all promises. Courts make errors, and legal sanctions are sometimes clumsier than nonlegal sanctions. As a result, people who make and receive promises often do not expect, and would not want, courts to provide legal remedies if the promisor breaks the promise. But when the promisor wants the promise to be legally enforceable, and the promisee expects the promise to be legally enforceable, courts should enforce promises.[44]

---

43. *See* Alan O. Sykes, *The Doctrine of Commercial Impracticability in a Second-Best World*, 19 J. LEGAL STUD. 43, 48 (1990); White, *supra* note 42, at 375; *see also* Victor P. Goldberg, *Impossibility and Related Excuses*, 144 J. INSTITUTIONAL & THEORETICAL ECON. 100 (1988) (expressing skepticism about the risk-sharing aspect); Polinsky, *supra* note 7 (arguing that endorsement of liquidated damages is the optimal remedy for risk sharing); George G. Triantis, *Contractual Allocations of Unknown Risks: A Critique of the Doctrine of Commercial Impracticability*, 42 U. TORONTO L.J. 450 (1992) (criticizing the excuse doctrines).

44. This conclusion excludes the possibility of a market failure. One might also stipulate a de minimis requirement: Courts should not enforce promises when the cost to the legal system exceeds the gains to the parties. Charging the litigants a fee would be a better response.

Imaged with the Permission of Yale Law Journal

Economics, then, implies that courts should enforce promises when parties want their promises to be enforceable, and not otherwise. Consistent with this view, courts both routinely enforce promises and respect terms of agreements that disclaim legal enforceability.[45]

But these simple ideas do not explain the main doctrines that draw a line between the legally enforceable promise and the unenforceable promise, namely, the consideration and promissory estoppel doctrines.

The consideration doctrine holds that a court cannot enforce a promise if it was not exchanged for "consideration," a legal benefit to the promisor or detriment to the promisee (or, in the modern formulation, a promise or performance that was bargained for). In essence, the doctrine knocks out of court promises that are not part of a quid pro quo. Such promises include option contracts, promises to give a gift, and open-ended agreements that bind one party but not the other.

Yet these promises are unobjectionable from an economic perspective. An option contract—for example, a promise to keep open an offer to sell something while the offeree investigates its value—might be the only way to attract the interest of a prospective purchaser. A promise to give a gift enables the promisee to rely in anticipation of receiving the benefit and enables the promisor to defer performance until the funds or goods are acquired. Open-ended contracts—where, for example, one side commits itself to purchase goods produced by the other side—are often efficient methods for shifting risk, with the legally unconstrained party bound by reputational concerns and nonlegal sanctions.[46]

The courts, possibly because they recognize the force of these arguments, have whittled down the consideration doctrine. Its main function is now to deny enforcement of promises to give gifts.[47] The consideration doctrine also serves, under the *Restatement*, as a formality: Options are unenforceable unless the parties "recite" consideration.[48] But there is no reason to require parties to recite a consideration as opposed to reciting that they want their option to be enforceable. The same can be said for Holmes's argument that the consideration doctrine was always just a formality, so gift promises would be enforced if the promisee gave nominal consideration to the promisor.[49] Efforts to rationalize this practice as a way

45. *See, e.g.*, Empro Mfg. Co. v. Ball-Co Mfg., 870 F.2d 423 (7th Cir. 1989).

46. *See* Charles Goetz & Robert E. Scott, *Principles of Relational Contracts*, 67 VA. L. REV. 1089, 1149-50 (1981).

47. The exception for charitable gifts only complicates the puzzle. If promises to give gifts are socially desirable, then the exception makes sense, but the general unenforceability of gift promises does not; if promises to give gifts are not socially desirable, then the exception does not make sense.

48. RESTATEMENT (SECOND) OF CONTRACTS § 87 (1981) [hereinafter RESTATEMENT].

49. OLIVER WENDELL HOLMES, THE COMMON LAW 294-95 (Little, Brown & Co. 1946) (1881); *see also* Lon L. Fuller, *Consideration and Form*, 41 COLUM. L. REV. 799 (1941) (criticizing Holmes's argument).

Imaged with the Permission of Yale Law Journal

of ensuring that courts can distinguish enforceable and unenforceable promises fail because they do not explain the "form" of the formality.

The doctrine of promissory estoppel might seem consistent with economics because it does not forbid courts to enforce value-enhancing gratuitous promises. The doctrine does place a limit on the enforcement of promises, however, and that limit is the requirement of promisee reliance. This limit is not consistent with economics. If a person wants to make a gratuitous promise, it must be because he wants to make the promisee better off. The promisor can make the promisee better off regardless of whether the promisee relies on the promise or can be proved to have done so. Economic analysis therefore suggests that enforceability of a promise should not depend on whether the promisee relied, or relied reasonably.[50]

Promissory estoppel can also be understood as a device for relaxing the consideration doctrine's prohibition on liability for precontractual reliance. Parties often spend some time negotiating the terms of a contract before entering it; frequently, one or both parties will make investments during this period in anticipation of the eventual success of the negotiations. An example is the relationship between a franchisor and a franchise applicant, which can extend for months or years before the granting of the franchise.[51] During this time, the franchisor might require the applicant to acquire experience as an employee in another franchise business, or undergo training. To induce the applicant to make these investments of time and effort, the franchisor might make vague or contingent promises that the franchise will be awarded. Even when these promises are not definite enough to form contractual commitments, applicants who are not awarded the contract can sometimes obtain damages for their reliance costs, on the basis of promissory estoppel.

Several scholars have considered the possibility that promissory estoppel is efficient because it protects the promisee's investment.[52] This view is at first sight attractive because the promisor's behavior, when not justified by the discovery that the promisee is unfit, seems opportunistic. In theory, the promisor can hold up the promisee after the promisee has invested and demand from the promisee additional fees or obligations that extract all the surplus generated from the promisee's investment.

But if courts could reliably verify the promisee's behavior, and thus distinguish the promisee who proves merely to be unfit and the promisee who is the victim of holdup, then the parties could enter a contract at the

50. *See* Steven Shavell, *An Economic Analysis of Altruism and Deferred Gifts*, 20 J. LEGAL STUD. 401, 420 & n.24 (1991).

51. *E.g.*, Hoffman v. Red Owl Stores, 133 N.W.2d 267 (Wis. 1965).

52. Richard Craswell, *Offer, Acceptance, and Efficient Reliance*, 48 STAN. L. REV. 481, 489-95 (1996); Avery Katz, *When Should an Offer Stick?: The Economics of Promissory Estoppel in Preliminary Negotiations*, 105 YALE L.J. 1249, 1270-77 (1996).

Imaged with the Permission of Yale Law Journal

beginning of their relationship, one that specified what the promisee must invest and how the promisee will be evaluated. As far as I know, parties do not enter such contracts, presumably because they do not believe that courts can make these distinctions.[53] But if courts cannot, then the use of promissory estoppel to protect reliance is not justified.

In addition, an efficient promissory estoppel doctrine would not require courts to compensate all of the promisee's reliance. If it did so, promisees would overinvest in reliance. Courts would need to determine how much reliance is efficient in each case, and then award damages only equal to efficient reliance, undercompensating parties that rely too much, or not compensating them at all. The proper award would depend on such factors as the cost of, and return on, investment; the probability that the preliminary relationship would yield a franchise;[54] and the parties' incentives to reveal information to each other.[55]

Craswell studied a group of cases in which equitable estoppel or promissory estoppel arguments were advanced by offerees in order to prevent an offeror from withdrawing an offer, and found that courts were more likely to rule in favor of the offeree when reliance on the offer is efficient.[56] But he disclaims any intention to show that the outcomes of the cases are themselves efficient, for just the reasons given above: "[T]here are several factors other than the efficiency of [the offeree's] reliance that can affect the desirability of a commitment."[57] The methodological difficulty of showing that contract doctrine is efficient dissuades Craswell from making the attempt.

H. *Summary: Descriptive Versus Normative Failure*

The charge of descriptive failure will not surprise scholars familiar with the literature on economic analysis of contract law. The inefficiency of contract law is a theme of Shavell, Goetz and Scott, and Schwartz on expectation damages; Epstein and Schwartz on the unconscionability doctrine; Ayres and Rasmusen on the mistake doctrine; Sykes and White on

53. In response to the growth of precontractual liability, franchisors now require franchise applicants to sign waivers.

54. *See* Craswell, *supra* note 52, at 499-501; *see also* Lucian Arye Bebchuk & Omri Ben-Shahar, *Precontractual Liability*, 30 J. LEGAL STUD. 423 (2001); Jason Scott Johnston, *Communication and Courtship: Cheap Talk Economics and the Law of Contract Formation*, 85 VA. L. REV. 385 (1999); Katz, *supra* note 52.

55. Avery Katz, *Contract Formation and Interpretation*, *in* 1 THE NEW PALGRAVE DICTIONARY OF ECONOMICS AND THE LAW, *supra* note 5, at 425, 429-30.

56. Craswell, *supra* note 52, at 531-36.

57. *Id.* at 507.

Imaged with the Permission of Yale Law Journal

the impossibility doctrine; Craswell on promissory estoppel; and others.[58] None of the authors goes so far as to deny that contract law is efficient. Each examines only a small slice of contract law, and normal methodological practice cautions against making exaggerated claims. But for the observer who looks at the steady accumulation of failures over thirty years, the conclusion is inescapable.

Some readers might agree with Ayres and Craswell that since most law-and-economics authors do not claim that they explain doctrine, and instead make normative recommendations, it is not worth making too much of this failure. Perhaps not, but I do not think that the distinction between descriptive and normative scholarship is so clear. The doctrinal structure of contract law exerts force on the scholarly analysis. That is why so many authors try to rationalize the doctrine, or propose incremental changes, rather than coming to the austere conclusions of Schwartz and others influenced by the incomplete contracts perspective. Courts and legislatures are more likely to pay attention to scholarly recommendations that follow naturally from the logic of contract law, than those that float down from the ether, for courts and legislatures have no good reason—no economic crisis, no foreign contract-law system that is clearly superior, no chorus of complaints like those heard about the tort system—to think that there is anything wrong with the system of contract law that we have. When economics was able to keep the descriptive and normative together, when it was able to say that contract law was essentially efficient but for some tweaking here and there, it had the potential to influence decisionmakers, for it worked with the past, not against it, and did not force decisionmakers to reject the past on the basis of conjectures founded on empirical postulates that could not be verified.

The descriptive failure of the models takes two forms. Simple models, which examine only one or two margins of contractual behavior, fail to predict contract law as it exists. The other models are more complex because they examine a greater variety of behavior, or because they rely on more complicated ideas, such as information asymmetry. These models sometimes fail because they make predictions that are inconsistent with contract law. But more often they fail because they are indeterminate. The models incorporate variables that cannot be measured, and to which one cannot with any confidence attach general ranges or distributions.

To repeat one example, recall that the choice between the *Hadley* rule and the rule of expansive liability depends, among many other things, on the shape of the distribution of buyer valuations. If, in terms of numbers of

58. *E.g.*, Avery Katz, *Transaction Costs and the Legal Mechanics of Exchange: When Should Silence in the Face of an Offer Be Construed as Acceptance?*, 9 J.L. ECON. & ORG. 77, 90 (1993) (finding acceptance-by-silence doctrines' efficiency to be "imperfect at best").

Imaged with the Permission of Yale Law Journal

buyers or the magnitudes of their valuations, the distribution is lopsided toward low valuations, then the *Hadley* rule is more likely to be superior. No one has tried to determine the shape of this distribution through empirical research, and indeed it is hard to imagine how this could be done. It is also foreign to the fact-finding activities of courts and legislatures. (One must also fix the relevant population, take a stab at guessing the cost of communicating, and so forth.) Accordingly, it might be best to assume that the distribution is uniform or normal, in which case neither rule is superior.

A further point is that the descriptive approach has not been fruitful in the way that it is in other areas of economics. In these other areas, the thing to be explained is always partly hidden. It makes sense to develop a hypothesis because in the process of testing it one learns new things about the world, resulting in a productive dialectic between theory and data. By contrast, the thing to be explained by the economic analysis of contract law—contract doctrine—is known, or thought to be known. And although at one time some scholars thought that outcomes of cases might diverge from contract doctrine, with the outcomes reflecting efficiency—in which case, generating and testing hypotheses would make sense—today this view has few adherents.[59] Judges have no reason to describe doctrine in a way that misrepresents the outcomes of cases.

Rather than arguing that their models explain contract doctrine, most authors argue that their models can be used to criticize or defend contract doctrine. But the normative weaknesses of their models follow as a matter of course. Simple models do not justify legal reform because these models exclude relevant variables. Complex models do not justify legal reform because the optimal rule depends on empirical conditions that cannot be observed.[60]

One might respond that even if economic models cannot generate a determinate optimal contract law, they helpfully identify the costs and benefits of different legal rules. Before the economic analysis of expectation damages and specific performance, a court trying to decide whether to push the doctrine in one direction or the other had little to go on. Economic analysis identified factors of which judges should take account, factors that include the cost of renegotiation and the advantages of permitting breach. Even if economic analysis cannot determine the magnitude of these costs and benefits, and the extent to which they offset or

59. There has, however, been a recent effort to defend this proposition. *See* Fred S. McChesney, *Tortious Interference with Contract Versus "Efficient" Breach: Theory and Empirical Evidence*, 28 J. LEGAL STUD. 131, 176-84 (1999).

60. Others have made similar observations about normative law and economics. *See, e.g.*, Lewis A. Kornhauser, *Constrained Optimization: Corporate Law and the Maximization of Social Welfare*, *in* THE JURISPRUDENTIAL FOUNDATIONS OF CORPORATE AND COMMERCIAL LAW, *supra* note 3, at 87, 112-13.

Imaged with the Permission of Yale Law Journal

interact with each other, the judge who knows about them is more likely to make a wise decision than a judge who does not.

This defense has an air of plausibility but also distressingly open-ended and unambitious implications. The last decade has witnessed a piling on of relevant factors, but no increasing clarity about the function of contract law, and a wise judge might, in order to avoid paralysis, simply ignore them. But the scholarship itself is mute about its own weaknesses. Part III will provide some methodological reasons for skepticism. Before we get there, however, we can gain additional insight by examining the literature on incomplete contracts.

II. THE THEORY OF INCOMPLETE CONTRACTS

The literature on incomplete contracts diverges from the law-and-economics literature, though they overlap in many ways. The theory of incomplete contracts was motivated primarily by descriptive curiosity about the nature of private contracting, not about contract law. As a result, contract law is usually treated in an exceptionally simple manner, as a system that specifically enforces contractual terms when the underlying behavior can be verified by courts.[61] This assumption enables scholars to focus on the parties' choice of contractual form. By contrast, law and economics generally assumes that parties choose simple contracts—contracts with a fixed price and quantity and sometimes a liquidated damages clause—and focuses on the effect of different legal rules on contractual behavior.

The incomplete contracts literature poses the following question to law and economics: Why would rational parties choose noncontingent contracts when more sophisticated contracts would enable parties to achieve better results? And if parties did choose more sophisticated contracts, why would courts need to do anything other than enforce the terms of these contracts? If courts only enforced the terms of contracts, much of contract doctrine, and much of the law-and-economics literature, would be irrelevant.

The following discussion of the theory of incomplete contracts serves two purposes. First, it allows us to examine whether the descriptive failure of law and economics is the result of economic methodology in general, or

61. *See, e.g.*, OLIVER E. WILLIAMSON, MARKETS AND HIERARCHIES: ANALYSIS AND ANTITRUST IMPLICATIONS (1975); Oliver Hart & John Moore, *Incomplete Contracts and Renegotiation*, 56 ECONOMETRICA 755 (1988). *Incomplete Contracts and Renegotiation* is generally considered the seminal article, though the literature has roots in Williamson's work. My reference to the literature on "incomplete contracts" is intended to encompass articles that are formally based on complete contract models but are concerned with the question of why contracts are incomplete, and not just articles that include transaction costs as an element in the model.

Imaged with the Permission of Yale Law Journal

of the law-and-economics approach in particular. Second, it sheds light on the methodological difficulties hidden in the concept of transaction costs.

A. *Premises and Basic Results*

The incomplete contracts literature focuses on two of the kinds of incentives we have been discussing: the incentive to invest (or "rely") and the incentive to perform or breach. An efficient or "first best" contract does two things: It ensures that (say) Seller performs when her cost is less than Buyer's valuation, and not otherwise ("efficient trade"), and it ensures that Buyer (and/or Seller) invests the right amount ("efficient investment").

As we have seen, there is a tension between efficient trade and efficient investment. A simple way of ensuring efficient trade requires Seller to pay Buyer's valuation if she does not perform. For example, under the rule of expectation damages, if Seller fails to perform, she must pay Buyer's valuation. Thus, she performs if and only if her cost is less than Buyer's valuation, the condition for efficient trade. However, Buyer will expect to receive his valuation whether performance occurs (in which case he gets the good) or not (in which case he gets damages equal to his valuation). Expecting to receive his valuation in both states of the world, Buyer will overinvest.[62]

All of this should be familiar from our discussion of the law and economics of remedies. The difference between the two literatures is in the next step. Where law and economics evaluates alternative legal rules according to their impact on the efficiency of contractual behavior, the incomplete contracts literature analyzes how the parties might design their contract in order to achieve efficiency. If Seller and Buyer are rational, they will want to prevent Buyer's overinvestment, with the parties sharing the surplus that comes from eliminating this inefficiency. They can do so through correct contractual design, assuming courts will specifically enforce it.[63]

The simplest solution would be to write a contract that says that the parties must trade if $v > c$, and that Buyer must make optimal investment r. If the court could observe the valuations and the investment, then it could use specific performance or a penalty in order to force the parties to engage in efficient behavior. But then we would have a complete contract, and such contracts, the argument goes, are never used. In fact, contracts are

62. *See supra* text accompanying notes 11-13.

63. A few authors have also examined rules such as expectation damages, but they produce optimal incentives only under narrow conditions. *See, e.g.*, Aaron S. Edlin & Stefan Reichelstein, *Holdups, Standard Breach Remedies, and Optimal Investment*, 86 AM. ECON. REV. 478, 495 (1996).

Imaged with the Permission of Yale Law Journal

incomplete, because transaction costs prevent the parties from putting all relevant, that is, value-maximizing, obligations in the contract.

The literature stipulates that transaction costs mean that the investment is not verifiable by a court, so the parties gain nothing by putting the optimal investment in the contract. For various reasons,[64] this assumption is thought to be a more satisfactory way of capturing the concept of transaction costs than, say, stipulating that there is a cost to writing an obligation down or that the parties must write fixed price or noncontingent contracts—the two preferred strategies in the law-and-economics literature. In any event, arguments about damages rules in the law-and-economics literature probably do not turn much on exactly how transaction costs are characterized.[65]

Even if the investment is nonverifiable, the parties could design a contract that provides efficient incentives. The contract would give Buyer the right to set the price at the time of performance and make an offer to Seller.[66] Seller would have the right to reject the transaction (and receive liquidated damages, assumed for the sake of the example to be zero) or to accept the transaction and accept the price announced by Buyer, in which case Buyer must accept delivery at that price.

This contract would achieve first-best efficiency. To see why, suppose first that $v > c$. Buyer will set the price equal to (or slightly higher than) Seller's cost, which Buyer observes. If Buyer set the price lower than Seller's cost, Seller would refuse to trade, and Buyer would gain nothing. But Buyer has no reason to set the price higher than Seller's cost, which would only reduce Buyer's own return. Thus, Buyer sets the price equal to c. But if $v < c$, Buyer will set the price at some low level in order to prevent Seller from demanding trade, for Buyer does not want to receive less than he pays. Seller will thus trade if and only if $v > c$, so the conditions for efficient trade are met.

When Buyer makes his investment, he knows that he will receive the goods only if $v > c$, and not otherwise. Thus, Buyer makes his investment with an eye to obtaining a return only in the performance state of the world. Indeed, Buyer will obtain the full residual of the investment because he will

64. *See, e.g.*, Ilya Segal, *Complexity and Renegotiation: A Foundation for Incomplete Contracts*, 66 REV. ECON. STUD. 57, 72-73 (1999).

65. The assumption that investments are not verifiable would not, for example, undermine Shavell's arguments about the relative efficiency properties of expectation and reliance damages, *see* Shavell, *supra* note 11, at 147, but would undermine Cooter's argument that the optimal damage measure is the amount that would compensate the victim if efficient investment had occurred, *see* Cooter, *supra* note 12, at 14.

66. I follow the discussion in Benjamin E. Hermalin & Michael L. Katz, *Judicial Modification of Contracts Between Sophisticated Parties: A More Complete View of Incomplete Contracts and Their Breach*, 9 J.L. ECON. & ORG. 230 (1993).

Imaged with the Permission of Yale Law Journal

set price equal to *c*. Thus, Buyer has the incentive to engage in efficient investment.

The contract's trick, in this case, is to give the party with the investment decision the residual from trade. As a result, the party has the correct incentives both to trade and to invest. The other party, Seller, must of course be compensated for her expected costs, and an ex ante transfer from Buyer to Seller accomplishes this task.

This is only the first step in a literature that has become very lengthy and complicated. Authors have discussed such problems as two-sided investment, where Buyer can increase his valuation and Seller can reduce her cost;[67] cooperative investment, where Buyer can reduce Seller's cost and Seller can increase Buyer's valuation;[68] third party effects;[69] and so forth.

The most interesting thing about these models is that they predict that contracts will contain descriptions not of "physical" contingencies but of the bargaining procedures that parties must follow at the time of performance. Lawyers think of contracts as either providing absolute obligations (Seller must deliver widgets by December 1st) or conditional obligations, with the conditions referring to events that occur in the world such as a strike or price change (Seller must deliver widgets unless Seller experiences labor difficulties, etc.). In models of incomplete contracts, the bargaining procedures specified in the predicted contracts are designed to force parties to divulge, and act efficiently on the basis of, their realized valuations, Seller's cost, and Buyer's value, and so references to events in the world are unnecessary. If Seller suffers a strike, for example, and her costs rise above Buyer's valuation, Seller will exercise an option to pay money rather than produce and deliver the goods. The contract does not need to refer explicitly to Seller's obligations in case of a strike. Because the parties can foresee that their valuations might change, and can design bargaining procedures that elicit efficient behavior (or behavior no more inefficient than that which would occur under a simpler fixed price contract), they do not have to write down countless contingencies in their contract. For this reason, the guiding premise of law and economics, that

67. Philippe Aghion et al., *Renegotiation Design with Unverifiable Information*, 62 ECONOMETRICA 257 (1994); Tai-Yeong Chung, *Incomplete Contracts, Specific Investments, and Risk Sharing*, 58 REV. ECON. STUD. 1031 (1991). For a survey, see Klaus M. Schmidt, *Contract Renegotiation and Option Contracts*, *in* 1 THE NEW PALGRAVE DICTIONARY OF ECONOMICS AND THE LAW, *supra* note 5, at 432.

68. Yeon-Koo Che & Donald B. Hausch, *Cooperative Investments and the Value of Contracting*, 89 AM. ECON. REV. 125 (1999).

69. *See infra* Section II.C.

Imaged with the Permission of Yale Law Journal

transaction costs render contracts incomplete and justify court provision of default rules, seems too strong.[70]

And yet the incomplete contracts literature does not provide a promising alternative. The contracts that the models predict do not exist in the world. Instead, we see simple fixed price contracts or contracts that are conditional on a relatively small number of real world contingencies. Intuitively, the problem with the predicted contracts is that they are too complex for parties to design.[71] To write such contracts, parties would need to imagine their bargaining position if a breach should occur, and then work their way via backward induction to the optimal terms of the contract. People are not very good at backward induction. Yet the rationality assumptions of economics hold that they can do it perfectly.[72] This problem has led to some discussion among economists about whether a theory of contracts can avoid relying on a model of bounded rationality,[73] an issue to which I will return in Part IV.

B. *Freedom of Contract and Asymmetric Information: The Penalty Doctrine*

The incomplete contracts literature was motivated by the desire to explain contracting, not contract law; it is a branch of industrial organization, not of law and economics. But authors writing in this tradition have tried to explain some contract doctrines, and their efforts are worth examining because they shed light on the law-and-economics literature.

Hermalin and Katz show that as long as parties are symmetrically informed, courts cannot increase welfare by modifying, or refusing to enforce, contractual terms.[74] The logic should be familiar by now, and is indeed identical to longstanding defenses of freedom of contract. Parties have more information than courts about their preferences, and even if courts can obtain superior information ex post, at the time of performance or dispute, the parties will anticipate this behavior and design their contracts

70. For similar skepticism, on methodological and empirical grounds, see Schwartz, *Incomplete Contracts*, *supra* note 24. For example, if transaction costs were the cost of describing obligations for future states of the world, we would probably observe more complete contracts than we do. *Id.* at 277.

71. For a discussion, see Karen Eggleston, Eric A. Posner & Richard Zeckhauser, *The Design and Interpretation of Contracts: Why Complexity Matters*, 95 NW. U. L. REV. 91, 122-25 (2000).

72. *See* George J. Mailath, *Do People Play Nash Equilibrium? Lessons from Evolutionary Game Theory*, 36 J. ECON. LITERATURE 1347 (1998).

73. Oliver Hart, *Is "Bounded Rationality" an Important Element of a Theory of Institutions?*, 146 J. INSTITUTIONAL & THEORETICAL ECON. 696, 700-01 (1990); Oliver Hart & John Moore, *Foundations of Incomplete Contracts*, 66 REV. ECON. STUD. 115, 134-35 (1999); Eric Maskin & Jean Tirole, *Unforeseen Contingencies and Incomplete Contracts*, 66 REV. ECON. STUD. 83, 107-08 (1999); Segal, *supra* note 64, at 74.

74. Hermalin & Katz, *supra* note 66, at 245.

Imaged with the Permission of Yale Law Journal

accordingly. Parties might, for example, leave the price term blank, to be filled in by the court ex post. But parties would not want a court to change the price term, or any other contractual term, based on its own judgment about what is ex post efficient.

If Hermalin and Katz are correct, then the instances where contract law does authorize courts to interfere with contract terms become a puzzle. The penalty doctrine is just one example, as we saw above, and indeed Hermalin and Katz criticize the penalty doctrine for the usual reasons.

Hermalin and Katz's argument is based on the assumption that the parties are symmetrically informed when they enter the contract. This assumption is not always true, and they acknowledge that if the parties are asymmetrically informed, judicial restrictions on contracts could increase welfare. Indeed, Hermalin made just such an argument with another coauthor three years earlier.[75] In the later article, however, Hermalin backs away from the asymmetry information argument.

In the earlier article, Aghion and Hermalin show that when the parties are asymmetrically informed, "legal restrictions on private contracts can enhance efficiency," as their title puts it.[76] The argument is best made by example. Imagine that a contractor has private information about the likelihood that it will perform a project on time. There are two types of contractors: The "good" type is more likely to perform on time than the "bad" type is. Buyers prefer good types to bad types, and so good types want to distinguish themselves from bad types. They do so by offering to pay an extremely high late fee or penalty if the performance is late. The bad type might mimic this signal, or not, but in either event the equilibrium, in which one or both types agree to the penalty, can be inferior to an equilibrium in which courts refuse to enforce penalties so that neither party can credibly agree to them.[77]

If we take this argument seriously, we should apply it not only to remedial terms. The same logic applies to the price term or, indeed, any other term of a contract. The contracts that emerge as a result of asymmetric information are simply inefficient contracts—it's not just that one term is inefficient, and the rest of the contract is efficient once that inefficient term is severed—and courts should refuse to enforce them even when a penalty clause is not activated by a breach.

To see why, imagine an employer who prefers educated applicants not because she cares about their education but because she believes that people

75. Philippe Aghion & Benjamin Hermalin, *Legal Restrictions on Private Contracts Can Enhance Efficiency*, 6 J.L. ECON. & ORG. 381 (1990).

76. *Id.* at 381.

77. The conclusion depends on the parameters of the model. Under some parameters, the separation of the types is superior to pooling; under other parameters, the opposite is the case. Aghion and Hermalin's point is that an inefficient equilibrium without judicial interference is possible, not certain. *See id.* at 399.

Imaged with the Permission of Yale Law Journal

who graduate from college work harder than those who do not. The employer offers two employment packages, a low salary for those without diplomas, and a high salary for those with diplomas. Because a potential job applicant's decision to obtain an education has an external effect—it increases or reduces the employer's information about the quality of other potential applicants—the resulting equilibrium could involve inefficient signaling. For that reason, courts or legislatures might want to prohibit the employer's discriminatory behavior.[78] The Aghion and Hermalin argument implies that courts should scrutinize all contracts for inefficiency, and not just liquidated damages terms.[79]

Aghion and Hermalin, then, cannot distinguish the common law's treatment of remedial terms and non-remedial terms such as price terms. As a descriptive theory, it is a failure. As a normative theory, it is also not successful, as it assumes that courts have sufficient information to distinguish signaling equilibria where judicial intervention will increase welfare, and other equilibria where it will not. It is for this reason that Hermalin and Katz back away from the conclusions of Aghion and Hermalin. The former article, as I noted, expresses skepticism about the ability of courts to improve on parties' contracts even when asymmetric information is present. This is another descriptive failure because Hermalin and Katz cannot explain judicial restrictions on remedial terms.

C. *Freedom of Contract and Externalities: The Penalty Doctrine Again*

When two parties design a contract, they will choose terms that are optimal for themselves; they will not take account of the interests of third parties who might be affected by the contract. But there are such third parties. Consider a contract in which Seller must pay Buyer liquidated damages if Seller breaches. If liquidated damages are set very high, they might interfere with the effort of a third party (TP) to purchase the good from Seller, even though TP might value the good more than Buyer does.

To understand why, imagine that different potential TPs have different valuations. Among those who value the good more than Buyer does, some value it slightly more and some value it considerably more. When Seller and Buyer agree to relatively high liquidated damages, the latter clause prevents the low-value TPs from buying the goods (Seller won't sell to

78. The diploma example is taken from A. MICHAEL SPENCE, MARKET SIGNALING 14-15 (1974).

79. I am puzzled by Ayres's response to this argument. Ian Ayres, *Valuing Modern Contract Scholarship*, 112 YALE L.J. 881, 889-90 (2003). Information asymmetry is the standard justification for regulation of the insurance market, and regulation always involves the imposition of mandatory terms and other restrictions. Aghion and Hermalin themselves use the example of mandatory employment benefits to illustrate their argument. Aghion & Hermalin, *supra* note 76, at 401-03. I use the diploma example only because it is famous.

Imaged with the Permission of Yale Law Journal

them because she has to pay high liquidated damages if she does), but it also enables Seller to extract a very high price from the high-value TPs, who must pay an amount at least as high as the already high liquidated damages. Under plausible conditions, Buyer and Seller jointly gain more by extracting the surplus from the high-value TPs than they lose by failing to sell to the low-value TPs. But this is inefficient, and the law should deter such behavior by refusing to enforce high liquidated damages provisions.[80]

This inefficiency disappears if renegotiation is possible: Ex post, the three parties will renegotiate so that TP will end up with the good, and efficient trade is achieved. But then the inefficiency shifts to the parties' investment incentives. The parties will choose inefficiently high liquidated damages to improve Seller's bargaining position vis-à-vis TP. Seller will use this bargaining power to extract some of the surplus generated by TP's high valuation. But this means that Seller and Buyer jointly enjoy a return on, say, an investment in Buyer's valuation even in the state of the world in which TP, not Buyer, acquires the good. So Buyer will overinvest.[81]

Do these arguments show that the penalty doctrine is efficient? They do show that, under certain conditions, enforcement of a liquidated damages clause can produce negative externalities. But the arguments do not travel the distance from this modest observation, to the conclusion that the penalty doctrine is justified. Indeed, the fit is poor. The penalty doctrine does not incorporate any of the variables identified in the literature: The cost of renegotiation, the distribution of valuations among potential TPs, the incentives to overinvest, and so forth. Further, the penalty doctrine effectively substitutes expectation damages for the invalid liquidated damages provision, but the literature we have been discussing does not establish that expectation damages are optimal.[82] Finally, parties can harm TPs in the way we have examined even without using liquidated damages, simply by overinvesting, which raises the expectation damages that the breacher would have to pay.[83] If courts care about efficiency and can detect this kind of strategic behavior, they should limit expectation damages. If they care about efficiency and cannot detect this kind of strategic behavior, they should not necessarily subject liquidated damages clauses to special scrutiny.

80. *See* Tai-Yeong Chung, *On the Social Optimality of Liquidated Damage Clauses: An Economic Analysis*, 8 J.L. ECON. & ORG. 280, 282-83 (1992).

81. *See* Kathryn E. Spier & Michael D. Whinston, *On the Efficiency of Privately Stipulated Damages for Breach of Contract: Entry Barriers, Reliance, and Renegotiation*, 26 RAND J. ECON. 180, 182-83 (1995).

82. *See* Chung, *supra* note 80, at 299.

83. *See* Tai-Yeong Chung, *Commitment Through Specific Investment in Contractual Relationships*, 31 CANADIAN J. ECON. 1057 (1998).

Imaged with the Permission of Yale Law Journal

D. *Summary*

Our diversion through the literature on incomplete contracts has taken us through formidable terrain. This literature is flourishing, and I do not consider myself knowledgeable enough to criticize it.[84] Instead, I want to make a few points about its relevance for understanding contract law.

First, so far the literature has failed to predict the content of either contracts or legal doctrines such as the penalty doctrine.[85] Like the law-and-economics literature, the incomplete contracts literature founders on the ambiguity of contractual behavior and the difficulty of empirical investigation of this behavior. With no empirical basis for endorsing some assumptions and rejecting others, the models tend toward indeterminacy.

Second, if the literature has any normative implications, they are that courts should always specifically enforce all terms of every contract. Although we have seen highly stylized arguments that information asymmetries and externalities might provide reasons for judicial intervention, these arguments depend on implausible assumptions about the amount of information that is available to judges. Thus, we are left with a sterile normative defense of freedom of contract, one that is closely tied to its premises that parties know more about their interests than courts do.

Third, the problems with the literature suggest methodological complications for the theory of contract law. The literature takes more seriously than law and economics the premise that parties are rational, and permits them to design complicated contracts. But the premise of full rationality does not seem right, for it predicts contractual structures that bear little resemblance to the contracts designed by real parties. We will discuss these problems more fully in Section III.B.

84. But it should be noted that the literature is not free from controversy. Maskin and Tirole argue that if parties can commit not to renegotiate, the nonverifiability assumption does not explain the existence of incomplete contracts: Parties will use the same kind of contract regardless of whether the investment is verifiable. Maskin & Tirole, *supra* note 73, at 84. Hart and Moore reply by showing that if parties cannot commit not to renegotiate, nonverifiability does matter under certain conditions. Hart & Moore, *supra* note 73, at 116. The debate thus turns on whether parties can commit not to renegotiate, about which neither set of authors is able to marshal a decisive argument.

85. Or, for another example, courts are not likely to order parties to send messages (for example, name a price or pay a penalty) as required by some proposed contractual devices, such as that in Schwartz, *Incomplete Contracts*, *supra* note 24, at 281. We don't know this for sure, however, as these devices are not used, and thus are not the subject of judicial orders.

Imaged with the Permission of Yale Law Journal

III. WHY ECONOMICS FAILS TO EXPLAIN CONTRACT LAW

A. *The Problem of Methodological Indeterminacy*

Richard Craswell has taken philosophical approaches to contract law to task for failing to provide fine-grained explanations of contract doctrine.[86] He points out that philosophical theories might explain in a general way why promises should be enforced—typically, by restating the moral intuition that promises should be kept, and then assigning the government a role in encouraging people to keep them—but never explain the details of doctrine. A theory that people should keep their promises does not tell us whether expectation damages, reliance damages, specific performance, or some other remedy is the appropriate response when a contract is broken. Indeed, when philosophers turn to these matters, they usually engage in implicit economic analysis or make assertions about the role of custom or other factors that are unrelated to their theories.[87]

Craswell's critique is methodological, not substantive. He argues that even if the philosophical theories capture some aspect of the truth about why contracts are enforced, they have no determinate implications for the phenomena that their authors purport to study—the doctrines, or the vast majority of the doctrines, of contract law. Although Craswell does not assert that economic analysis avoids this methodological problem, many readers will understand him to be implicitly making this claim. One cannot avoid being impressed by the contrast between the large and ingenious economic body of work on default rules, and the small and vapid body of work produced by philosophers.

But even if we accept Craswell's critique of philosophical theories of contract law, we must still ask whether economics really enjoys any advantages. By now, the answer should be familiar. Economists have proposed some models of contract behavior that make determinate but wrong predictions about the law. These models avoid Craswell's charge of indeterminacy, but they are still wrong. Determinate but wrong predictions enjoy a little more intellectual respectability than indeterminate predictions, but they get us not much closer to an understanding of contract law.

Economists have proposed other models of contract behavior that make predictions that are indeterminate. These models enjoy some intellectual advantages over the philosophical theories, for they would enable us to make complex and interesting predictions about contract law if we had

86. *See* Richard Craswell, *Contract Law, Default Rules, and the Philosophy of Promising*, 88 MICH. L. REV. 489 (1989).

87. *See, e.g.*, CHARLES FRIED, CONTRACT AS PROMISE: A THEORY OF CONTRACTUAL OBLIGATION (1981); Randy E. Barnett, *A Consent Theory of Contract*, 86 COLUM. L. REV. 269 (1986).

Imaged with the Permission of Yale Law Journal

sufficient information about empirical conditions. But because we do not have such information, and it is—in my view, though others might disagree—unlikely that we ever would, the complex economic theories do not get us much closer to an understanding of contract law than the philosophical theories do.

Because this state of affairs could change with further research, one should not discount a renaissance in the law and economics of contracts. But another view is that theory and doctrine are mismatched, operating at different levels of generality. Welfare economics might be able to provide persuasive reasons for the superiority of a free market to, say, a planned economy. A free market can function only if people can trade, and trading almost always requires the making of binding promises. But there are many ways that promises can be made binding: through the operation of ordinary reputational mechanisms, through the creation of institutions like firms and trading associations that establish commitment mechanisms for members, and through contract law. And then there are many different rules of contract law that will be equally good at enabling people to make binding promises.[88] Specific performance is about the same as damages; literalistic interpretation is about the same as purposive interpretation. Individual contract doctrines, then, could be like rules of the road: sufficient as long as, within limits, everyone obeys them, and thus not susceptible to prediction on the basis of fine-grained theories of optimal interaction.

B. *The Problem of Rationality*

The economic scholarship on contract law purports to assume that individuals are rational in the sense of neoclassical economics. Their preferences obey certain consistency requirements, and their cognitive capacity is infinite. But on inspection, the nature of the rationality assumptions made by this scholarship is not so clear.

If individuals were rational, with no cognitive limits, and if transaction costs were zero, the role of contract law would be simple and uninteresting. Parties would foresee every possible future state of the world, and—the story goes—their contract would describe each party's obligation in each of these possible future states. For example, a contract for the sale of widgets would describe Seller's obligation if the cost of widgets increases or declines, and could make Seller's obligation turn on whether Seller invested in the right way, and so forth. Courts would specifically enforce the terms of the contract. In general conditions, efficiency would be obtained.

88. Richard A. Epstein, *Contracts Small and Contract Large: Contract Law Through the Lens of Laissez-Faire*, *in* THE FALL AND RISE OF FREEDOM OF CONTRACT 25, 26 (F.H. Buckley ed., 1999).

Imaged with the Permission of Yale Law Journal

Economic analysis of contract law assumes that contracts cannot be designed to describe every future state of the world. The usual statement is that transaction costs prevent the parties from achieving such a detailed and complex contract. Some authors seem to mean the cost of negotiating and writing a contract; other authors seem to refer to cognitive limits of the parties, which include the inability to foresee future events and maybe something more.[89] In any event, one needs some such assumption to get the economic analysis of contract law off the ground; if the parties entered complete contracts, the law would not need to supply default terms such as expectation damages. Instead, the parties would choose expectation damages whenever they anticipated the need, namely, when they wanted to give the promisor the option to perform or pay a sum of money to the promisee.

Let us examine the two main ways that authors use the idea of transaction costs. The first approach assumes that parties are rational but that entering a contract involves some special cost. Some authors assume that this cost refers to time spent negotiating or the time and materials needed to draft a document.[90] Others assume that the cost results from problems of asymmetric information, and, in particular, the inability of a court to verify a subset of the contract-related actions in which the parties engage.[91] As discussed in Part II, rational parties would minimize these costs by entering contracts that incorporate complex ex post bargaining mechanisms. Yet there is little evidence that such mechanisms are used in real contracts.

This problem is less clear in the law-and-economics literature than in the incomplete contracts literature, but that is only because the law-and-economics models constrain the types of contracts that parties may enter rather than formally modeling the transaction cost. The models generally permit parties to choose prices and quantities in noncontingent contracts, and not to choose contracts that stipulate ex post bargaining procedures, though these are likely to be superior. One defense of the methodological approach of law and economics is that the latter contracts are too complex to be useful, so it is a justified simplification to assume that only fixed price contracts are available to parties.

This brings us to the second approach to the idea of transaction costs, which is, in fact, to treat it as a metaphor for bounded rationality, the idea that parties are rational to the extent permitted by limits on cognitive capacity.[92] Although law-and-economics scholars rarely put their argument in this way, the assumption is reflected, as just noted, in the modeling

89. *See, e.g.*, COOTER & ULEN, *supra* note 2, at 100, 236, 240.
90. *E.g.*, Ayres & Gertner, *Filling Gaps*, *supra* note 16, at 93.
91. Like investment levels. *See* Hart & Moore, *supra* note 61.
92. *See, e.g.*, WILLIAMSON, *supra* note 61, at 21.

Imaged with the Permission of Yale Law Journal

device of permitting parties to choose only noncontingent contracts.[93] To understand the problem with this strategy, consider the common claim that default rules should be designed to give a party an incentive to reveal information about the party's cost or valuation. If "transaction costs," meaning bounded rationality, prevent the parties from choosing a sophisticated contract in light of future events, then they should also prevent parties from anticipating the effect of legal rules (which would be applied only in the contingent future) on the simple contract that they design. Instead, the model simultaneously assumes that individuals can foresee remote events and make complex calculations (otherwise they would not be motivated by the default rule to release information) and cannot engage in a perfect cognitive response (otherwise the cost of entering the bargain would be zero). The assumptions are jointly implausible.[94]

Economists reject bounded rationality arguments for two reasons. The first is methodological: They cannot agree on a standard, mathematically tractable formulation of bounded rationality. This might be a good reason for economists, but it is a bad reason for lawyers. The second is empirical: If the rationality assumptions of economics are close enough to the reasoning of individuals, or to the institutionalized reasoning implemented by firms, then the conclusions of the economic models are also good enough for predictive and normative purposes. My view is that the failure of contracts to include the mechanisms identified by the incomplete contracts literature is evidence that the rationality assumptions are not good enough. Others might disagree, claiming either that transaction costs—that is, high writing costs, severe information asymmetries, etc.—explain the absence of these mechanisms, or that better modeling will lead to different conclusions in the future. This seems to me a dodge, especially in the absence of an empirical test of the role of transaction costs in preventing the use of mechanisms.[95] But it cannot be dismissed out of hand. The question, then, is whether one should have optimism or pessimism about future research.

93. I do not mean that the authors self-consciously made this modeling choice as a way of capturing bounded rationality. I mean that they usually do not explicitly model transaction costs as, for example, the cost of drafting the contract and, instead, treat transaction costs as an (informal) explanation for why they assume that parties can choose only the contract price and quantity (and sometimes a liquidated damages provision). *E.g.*, Shavell, *supra* note 11. Even when a variable is used to refer to a transaction or communication cost in a formal model, the parties are constrained in their choice of contractual form. *E.g.*, Ayres & Gertner, *Filling Gaps*, *supra* note 16, at 108.

94. A similar point has been made about the incomplete contracts literature. *See* Maskin & Tirole, *supra* note 73, at 84; Segal, *supra* note 64, at 74.

95. But, of course, if there are no contracts that use these mechanisms, the techniques of econometrics are hardly necessary.

Imaged with the Permission of Yale Law Journal

C. *A Way Out?*

One way out of this impasse involves greater consideration as to what parties can realistically be expected to foresee. Contrast the efficient breach theory and the *Hadley* theory. If Seller experiences higher than expected costs and would like to avoid the deal, she would probably consult a lawyer. The lawyer would tell her that if she breaches the contract, she will probably pay expectation damages. Comparing the cost of performance and the expectation damages, Seller will decide whether or not to perform. This decision seems well within the cognitive abilities of an ordinary Seller.

By contrast, the *Hadley* theory applies to a decisionmaking process that occurs well before performance. At the time of contracting, Buyer (for example) must anticipate that Seller might breach rather than perform; that Seller must pay damages if she breaches; that these damages depend on the valuations of other buyers as well as Buyer's revelation of his private valuation; that Seller will (or maybe won't) anticipate these damages when deciding how much care to use when performing; that Seller will (or maybe won't) use Buyer's information to price discriminate; and so on. This chain of reasoning seems likely to exceed the cognitive capacities of an ordinary Buyer, whether or not a lawyer is consulted.

One way out of the impasse, then, requires incorporation of cognitive limitations into a theory of the relationship between contract law and contract-related behavior, so that one can distinguish incentives that are likely to influence behavior, and those that are too remote to influence behavior. No widely accepted theory of bounded rationality exists, however. The likelihood that such a theory could be developed is discussed in Section V.D.

IV. THE INFLUENCE OF ECONOMICS ON CONTRACTS SCHOLARSHIP AND LAW

A. *The Influence of Economics on Contracts Scholarship*

Defenders of economic analysis of contract law point to the significant influence of economics on contracts scholarship. Indeed, this influence can be documented in many ways. Economic analysis played almost no role in contracts scholarship prior to 1970, whereas it has played a dominant role in much of the contracts scholarship published in the major law reviews in the 1990s. It has influenced some of the analysis in the treatises. It shows up in the casebooks. Economic articles on contracts are frequently cited in noneconomics articles. And contracts scholars at the top law schools are frequently identified with the law-and-economics approach.

Imaged with the Permission of Yale Law Journal

But this rosy picture does not tell the whole story. The most influential economic articles, with one exception, were published in the 1970s and early 1980s. The exception, Ayres and Gertner's 1989 article on default rules, is usually cited for the useful distinction between majoritarian and penalty defaults and not for the economic analysis of this distinction. On the whole, economics-of-contracts articles published in the last fifteen years are cited no more frequently than noneconomics articles are.[96]

Contract-law casebooks and treatises show the influence of economics, but it is the influence of pre-1980 economics. Most casebooks and treatises mention the idea of efficient breach, but not the equally important idea of efficient reliance.[97] Casebooks generally treat the economic approach as an exotic "perspective," as an object at which to marvel, and not as the underlying logic of contract law. To be sure, most casebook authors are noneconomists, but what is important is that these authors have apparently concluded that greater economic content would not expand the market share of their casebooks.[98]

B. *The Influence of Economics on Contract Doctrine*

The influence of economic analysis on contract law is harder to discern. Let us start with the common law. Judicial opinions occasionally cite economic articles, and occasionally use economic concepts such as transaction costs and risk aversion. But it is hard to find cases where the judges self-consciously rely on an economic argument in order to justify a result. One such case is the *Van Wagner* case,[99] which relies heavily on Anthony Kronman's analysis of specific performance.[100] Many opinions

96. This conclusion is based on a regression of the annual citations of articles published since 1980 in leading law journals (California, Chicago, Columbia, Harvard, Michigan, Northwestern, NYU, Pennsylvania, Stanford, Texas, UCLA, and Yale) and faculty-edited journals (*Journal of Law and Economics*, *Journal of Law, Economics, and Organization*, and *Journal of Legal Studies*) on a dummy variable equal to one if the article uses economic analysis, and zero otherwise. The mean annual citation for economics articles was 3.8 (71 articles); for noneconomics articles, it was 4.1 (52 articles). If you exclude all the faculty-edited journals, however, then economics articles are cited more often at a statistically significant level; but if you include the *Journal of Legal Studies* alone, then they are not. There are many ways that one could conduct this test, and for that reason I can conclude only that citation evidence does not exclude the hypothesis that economics articles are no more influential than noneconomics articles.

97. A Westlaw search of Farnsworth's treatise yields the following results: "efficien!" = 32 results; "economic!" = 113 results; "transaction! cost!" = 5 results; "Coase" = 2 results. The "economic!" search caught many concepts unrelated to economic analysis, such as "economic waste."

98. The Scott and Leslie casebook is the exception; compare the contracts casebooks by Barnett, Dawson, and Eisenberg. *Compare* ROBERT E. SCOTT & DOUGLAS L. LESLIE, CONTRACT LAW AND THEORY (2d ed. 1993), *with* RANDY E. BARNETT, CONTRACTS: CASES AND DOCTRINE (1995), JOHN P. DAWSON ET AL., CONTRACTS: CASES AND COMMENT (7th ed. 1998), *and* LON L. FULLER & MELVIN ARON EISENBERG, BASIC CONTRACT LAW (6th ed. 1996).

99. Van Wagner Adver. Corp. v. S & M Enters., 492 N.E.2d 756, 760 (N.Y. 1986).

100. Kronman, *supra* note 14.

Imaged with the Permission of Yale Law Journal

cite Posner and Rosenfield's article on the impossibility doctrine, but almost always for the proposition that contracts shift risk, an idea that predates economic analysis by many decades. The notes to the *Restatement* contain only a handful of references to economic ideas.[101]

To examine the influence of economic analysis on contract doctrine more systematically, I read the state and federal court opinions that cite an economics article that appeared in a major law review or faculty-edited law journal after 1980. Only thirty-six such opinions were issued. Of these, few discussed rather than cited the article, and none was clearly influenced by an article.[102]

Economic analysis has also had little influence, as far as I can tell, on statutory and regulatory law. Statutes and regulations in the 1970s incorporated common-law developments that economic analysis criticized, and although the consumer-protection movement crested in the 1970s, economics did not spur deregulation of the consumer-product and consumer-finance markets as it did for so many other markets such as trucking and air travel.

V. THE FUTURE

If the limits of economic analysis are becoming visible, can some other methodology take us beyond them? A brief survey of the other contenders provides grounds for concern.[103]

A. *Philosophy*

For a long time, legal scholars have sought a philosophical explanation for contract law. Fuller and Perdue argued that contract law is based on corrective justice.[104] Fried argued that contract law is based on the morality

101. RESTATEMENT, *supra* note 48. Searches on Westlaw yielded the following results: "efficien!" = 8 results; "economic!" = 38 results; "transaction! cost!" = 1; "Coase" = 0 results. In fact, few of the "economic!" results were related to economic analysis (as opposed to "economic waste," etc.).

102. Richard Posner's *Economic Analysis of Law*, however, has been cited in 116 cases in the LEXIS contracts database. A sample suggests a mix of meaningful analysis and meaningless citation. Cooter and Ulen's textbook was cited only twice. It might be the case that economically minded judges such as Calabresi, Easterbrook, and Posner have influenced contract doctrine, and I have not tried to measure their influence by looking at whether other judges accept their views about contract law. William Dodge credits the theory of efficient breach as provoking a judicial backlash against awarding punitive damages for breach of contract. *See* William S. Dodge, *The Case for Punitive Damages in Contracts*, 48 DUKE L.J. 629, 642-44 (1999).

103. For a more optimistic view, see ROBERT A. HILLMAN, THE RICHNESS OF CONTRACT LAW (1997).

104. Lon L. Fuller & William R. Perdue, Jr., *The Reliance Interest in Contract Damages*, 46 YALE L.J. 52 (1936).

Imaged with the Permission of Yale Law Journal

of promising.[105] These theories remain the most influential despite their inadequacies.[106]

Let us first consider each theory from a descriptive perspective. Fried argues that contracts should be enforced because individuals have a moral obligation to keep their promises. Fried's theory has the virtue of simplicity but cannot explain the many ways that contract law refuses to enforce promises. Unreciprocated offers, promises that lack consideration, promises that violate the Statute of Frauds, promises that lack specific terms—all of these promises are, in ordinary cases, not enforced. Finally, as Craswell has pointed out, Fried's theory cannot explain the default terms the law uses to fill out promises that otherwise are ill-defined.[107]

Fuller and Perdue argue that contracts should be enforced in order to prevent one party, the promisor, from benefiting at the expense of the other party. Corrective justice demands that the breaching promisor make the promisee whole. The reliance measure is ideal for this purpose, but because reliance costs are hard to measure and the expectation measure approximates the reliance measure in competitive markets, the expectation measure is the appropriate rule. But as Craswell shows, Fuller and Perdue's theory cannot explain why the appropriate baseline for exercising corrective justice is the promisee's position prior to the making of the promise, as opposed to after the making of the promise.[108] Others—notably Grant Gilmore and Patrick Atiyah—tried to generalize Fuller and Perdue's analysis and claimed that contract is being absorbed into tort.[109] But these efforts have gone nowhere.

The theories fare no better when conceived as normative arguments for the reform of contract law. As Craswell shows, they are indeterminate over nearly all aspects of contract doctrine.[110] Fried's theory justifies the enforcement of promises, but sheds no light on which of many remedies—expectation damages, reliance damages, specific performance, even nominal damages—is the right one. Fuller and Perdue's theory, as just mentioned, cannot solve the baseline problem.

---

105. FRIED, *supra* note 3.

106. The present discussion concerns normative theories, not "analytic" or "interpretive" theories (as Craswell uses the terms) that have been advanced by legal philosophers. For the distinction, see Stephen Smith, An Introduction to Contract Theory (2002) (unpublished manuscript, on file with author). Philosophers have goals that are different from those of economists, *see* Jody S. Kraus, *Philosophy of Contract Law*, *in* THE OXFORD HANDBOOK OF JURISPRUDENCE AND PHILOSOPHY OF LAW 687 (Jules Coleman & Scott Shapiro eds., 2002), and here I am asking only whether philosophy can succeed where economics has failed.

107. Craswell, *supra* note 86, at 521-23.

108. *Id.*

109. P.S. ATIYAH, ESSAYS ON CONTRACT (1990); GRANT GILMORE, THE DEATH OF CONTRACT (Ronald K.L. Collins ed., 1995).

110. Craswell, *supra* note 86.

Imaged with the Permission of Yale Law Journal

None of this is to say that philosophy has nothing to offer contract law. Philosophical reasoning, if not necessarily the reasoning of philosophers, has a significant accomplishment: the critique of the "will theory" of contract. The will theory, which derives contract doctrine from the premise that a contract is the coming together of two wills, is not just a once popular legal theory; it is also an intuitive, common sense approach to understanding contract law, instinctively adopted by generation after generation of first-year law students, and a happy target for philosophical criticism. The celebrated critique of the will theory of the duress doctrine—namely, contracts entered under duress and contracts entered "voluntarily" involve the same kind of coercion—is powerful and important, and this critique owes something to philosophical reasoning. But one must also understand that this critique is much older than those who are credited with it: It can be found not only in Dawson, and Hale, and Holmes, but also in Hume and probably earlier.[111] Basic philosophical ideas about the nature of the will, of agreement, and of contract go back for centuries, and their implications for contract law are well understood.

B. *Psychology*

Several scholars have recently argued that cognitive psychology holds promise for explaining the law, including the law of contracts. This view has superficial attractiveness.[112] If, as I have argued, economic models of the law are undermined by their rationality assumptions, then psychologically accurate models of human cognition might fill in the gaps left by the economic explanation.

Let us focus on the example of the penalty doctrine because it is the topic of a recent debate about the value of using cognitive psychology to understand the law. We have already seen that economics fails to explain the penalty doctrine; can cognitive psychology?

The question is, why do courts give less deference to liquidated damages clauses than they do to other provisions of a contract, including choice-of-forum clauses, which will become relevant only if a dispute arises? In response to Robert Hillman's skepticism about whether cognitive psychology can explain this practice,[113] Jeffrey Rachlinski argues that (1) biases that cause overoptimism justify scrutiny of liquidated damages provisions; (2) the status quo bias (contrary to Hillman's claim) does not

111. DAVID HUME, A TREATISE OF HUMAN NATURE 525 (L.A. Selby-Bigge ed., Oxford Univ. Press 1978) (1740).

112. *See, e.g.*, Russell B. Korobkin & Thomas S. Ulen, *Law and Behavioral Science: Removing the Rationality Assumption from Law and Economics*, 88 CAL. L. REV. 1051 (2000).

113. Robert A. Hillman, *The Limits of Behavioral Decision Theory in Legal Analysis: The Case of Liquidated Damages*, 85 CORNELL L. REV. 717 (2000).

Imaged with the Permission of Yale Law Journal

justify deference because the increased effort to bargain around the damages rule does not necessarily eliminate the effects of overoptimism; and (3) although aversion to ambiguity justifies deference to liquidated damages, courts actually use this insight under the penalty doctrine by giving more deference to liquidated damages clauses when damages are hard to calculate (and thus ambiguous).[114]

Even accepting these arguments, which will strike many as ad hoc, Rachlinski cannot explain why the biases justify judicial scrutiny of liquidated damages terms but not other terms. Breach is not the only low-probability event that occurs within contractual relations; a contract might make any number of obligations conditional on events that occur with a low probability. Think of bond covenants that give creditors the right to accelerate repayment when the debtor's asset-debt ratio falls below a threshold, employment compensation packages that provide payoffs only when market conditions are favorable, and sales contracts that allocate the risk of the destruction of the goods during delivery. If parties overlook low-probability events, then any of these provisions could be defective, but because they are not liquidated damages provisions, courts do not subject them to scrutiny. Indeed, Rachlinski concedes the explanatory failure of cognitive psychology when he says that the field "might cause scholars to question much of contract law's foundations."[115] Rachlinski slips from a descriptive claim to a normative claim in the face of the poor fit of cognitive psychology and the penalty doctrine.

C. *History*

Historical explanations of contract law once held promise, but early enthusiasm has given way to skepticism. Consider the attempt to link trends in contract-law doctrine to the rise of the welfare state.[116] Scholars claimed that the increasing informality of contract law over the last century, and especially the rise of promissory estoppel, showed courts moving away from laissez faire and toward statism and the enforcement of community standards. The convenient link to other trends in political economy, and specifically the rise of the welfare state, obscured the poor fit between the theory and doctrinal trends.[117] The rise of promissory estoppel, for example, could be interpreted as reflecting judicial impatience with a formality—the consideration doctrine—that interfered with, rather than promoted, private

114. Jeffrey J. Rachlinski, *The "New" Law and Psychology: A Reply to Critics, Skeptics, and Cautious Supporters*, 85 CORNELL L. REV. 739, 762 (2000).

115. *Id.* at 763.

116. P.S. ATIYAH, THE RISE AND FALL OF FREEDOM OF CONTRACT (1979); GILMORE, *supra* note 109.

117. *See* Epstein, *supra* note 88, at 25.

Imaged with the Permission of Yale Law Journal

contracting.[118] Contract doctrine can coexist with many different political systems, and broad trends, such as the decline of formalism, do not necessarily reflect changes in politics or morality.

Similar criticisms apply to Simpson's argument that the eighteenth-century shift from judicial accommodation of penalties to hostility toward penalties was due to "social evolution" away from tolerance for the private use of terror to the monopoly of force held by the state.[119] Simpson argues that once courts deprived private parties of the right to use force, the courts vindicated their longstanding commitment to the compensation principle by banning penalties.

The argument raises more questions than it answers. Both before and after the "social evolution," parties depended on courts to enforce their contracts. Before, a party could not collect a penal bond without first obtaining a judgment from a court. The other party had a number of legal defenses: not only full performance of the underlying promise, but such conventional defenses as duress and impossibility. The doctrinal change did not reflect a shift away from tolerance for private use of terror; it reflected a shift in the degree of deference given to remedial terms in contracts. Simpson's argument boils down to the assertion that judges stopped deferring to remedial terms because they wanted to control remedies, but he does not explain why they would want to treat contractually stipulated remedies differently from other terms in the contract.

Even if we accepted the "social evolution" argument—the shift from private to public remedies—we need to understand why judges would think that the "compensation principle" should control remedies and thus exclude penalties.[120] Simpson argues that judges had a longstanding belief that the law should provide compensation (not overcompensation) for injuries. But judges also believed that "the real function of contractual institutions is to make sure, so far as possible, that agreements are performed."[121] Simpson acknowledges that the two principles—compensation and respect for agreement—are in conflict, but does not explain why the first prevailed over the second.

Simpson's argument shares the flaws of the historical scholarship described above—the use of macro trends to explain micro phenomena that are consistent with other trends, the casual appeal to such long-term trends to explain a change that occurred at a particular time, and the arbitrary resolution of tensions between different principles or ideas in favor of one

118. Eric A. Posner, *The Decline of Formality in Contract Law*, *in* THE FALL AND RISE OF FREEDOM OF CONTRACT, *supra* note 88, at 61.

119. A.W.B. SIMPSON, A HISTORY OF THE COMMON LAW OF CONTRACT 124 (1975).

120. Recall also the ambiguity of the notion of "compensation," which presupposed a baseline. *See supra* text accompanying note 108. But we will assume that whatever compensation means, a penalty requires something beyond compensation.

121. SIMPSON, *supra* note 119, at 123.

Imaged with the Permission of Yale Law Journal

rather than the other. Historical scholarship is often illuminating, and any good theory of contract law would need to account for aspects of its historical evolution, but thus far theories of contract law emerging from historical research have not resolved basic puzzles about modern contract doctrine.

D. *Advances in Economics and Game Theory*

I argued in Section III.C that the failures of the economic analysis of contract law derive, in part, from the bounded rationality of individuals who enter contracts. If people were rational, then their contracts could be predicted. And if people's behavior could be predicted in this way, then firm recommendations about contract doctrine could be made. But because the models press rationality to its limits, the world falls short of the predictions, and so the natural consequence is to incorporate cognitive limitations into models of behavior.

The economic literature on bounded rationality is complex and large, and I cannot do justice to it. None of the models of bounded rationality that have been proposed has achieved canonical status,[122] and thus it is difficult to discuss in general what bounded rationality means for contract law. An example will illustrate the problems, and, I think, justify skepticism about the ability of future models of bounded rationality to shed light on contract doctrine.[123]

In the beauty contest game, the experimenter asks each member of a group to write down a number between 0 and 100 that is, say, 2/3 of the average number (between 0 and 100) that everyone else writes down. The person who writes the correct answer wins a prize; if there is a tie, the prize is divided among the people with the right answer.

Game theory—that is, game theory that assumes "unbounded" rationality—predicts that everyone will write down 0 and share the prize. The intuitive explanation for this prediction follows. Imagine that you are one of the people asked to write down the number. You might start by imagining that everyone else will pick a number at random. If so, you might expect a uniform distribution from 0 to 100, with a mean of 50. Thus, you would pick 2/3 of the mean, which is about 34. But then it might occur to you that everyone will have reasoned in the same way that you have. Thus, everyone else will have written down 34. But if everyone else has written down 34, then you can win only by writing down 2/3 of 34, which is about 23. Yet everyone else knows this as well, so they will write down 23, and

122. *E.g.*, David M. Kreps, *Bounded Rationality*, *in* 1 THE NEW PALGRAVE DICTIONARY OF ECONOMICS AND THE LAW, *supra* note 5, at 168.

123. For other efforts, see *id.* at 168-73.

Imaged with the Permission of Yale Law Journal

you should write down 2/3 of 23, and so on. Continuing with this reasoning, we reach 0. If everyone chooses 0, then you can do no better than choosing 0 as well, for then you will share the prize, whereas if you choose a higher number, you will receive nothing.

The explanation can be given more formally. If everyone chooses 0, everyone gets a share of the prize. If one person decides to choose a number different from 0, that person will no longer receive a share of the prize, and thus will receive a lower payoff. No one can do better than choosing 0 given that everyone else chooses 0. The outcome 0 is what game theorists call a Nash equilibrium: It is an outcome from which one has no incentive to deviate, given everyone else's choice, because one cannot increase one's payoff by deviation.[124] By contrast, if one person knows that everyone else will choose a particular number $n > 0$, then that person can do better by choosing a number different from n, namely 2/3 of n. Thus, any $n > 0$ cannot be a Nash equilibrium.

Nash equilibrium does a poor job predicting behavior in the initial rounds of play. When people play the beauty contest game, typically there is a distribution as follows. For a sufficiently large group, most or all numbers will be chosen, with spikes around 34, 23, and perhaps 14, and then 0.[125] A natural explanation of this pattern is that some people choose numbers randomly, or misunderstand the game; others are able to think one, two, or three steps ahead, or even more, but it never happens that everyone figures out, and plays, the Nash equilibrium in the initial rounds.

Aside from the empirical disconfirmation, the experiments pose a conundrum. Suppose we imagine a perfectly rational person, X. What do you predict that X will do? In Nash equilibrium X chooses 0, but if X is smart (rational?) enough, X will realize that not everyone else will play the Nash move, in which case X should choose a number greater than 0, its magnitude depending on the distribution of cognitive ability in the population. So our purely rational X will not act purely rationally as defined by game theory.

What to do? One idea recently investigated by Teck-Hua Ho and his coauthors proceeds as follows.[126] Imagine that a person engages in a number of cognitive steps when thinking about how to play the beauty contest game. In step 0, he randomly chooses a number between 0 and 100. In step 1, he thinks that everyone else has engaged in step 0 and only step 0—that is, everyone else has chosen a number at random—and chooses

124. *See* DOUGLAS G. BAIRD ET AL., GAME THEORY AND THE LAW 19-23 (1994).

125. *See* Teck-Hua Ho et al., *Iterated Dominance and Iterated Best Response in Experimental "p-Beauty Contests,"* 88 AM. ECON. REV. 947, 953-58 (1998). Their version of the game uses a different range of numbers from the example in the text, which is simpler, but the pattern is the same.

126. *See id.*

Imaged with the Permission of Yale Law Journal

strategically on the basis of this assumption. (In our example, he would choose 34.) In step 2, he thinks that a fraction of the population has engaged in step 0, and the rest has reached step 1, but none (besides himself) has reached step 2, and he chooses accordingly. This process continues for an arbitrary number of steps.

To predict the distribution of numbers chosen in the beauty contest game, we assume that a portion of the population stops at step 0, another group stops at step 1, another group stops at step 2, and so forth. To make such a prediction, we have to decide how many steps there are, and how the cognitive ability (the number of steps taken) is distributed among the population, but let us suppose that we can make reasonable assumptions about these parameters. Ho and his coauthors show that if we make such assumptions, we will predict something close to the actual distribution—all or most numbers being chosen, with spikes near 34, 23, and so forth, except that there will not be a spike at 0. (If the game is repeated, however, people will learn from their mistakes and eventually nearly everyone will play the Nash equilibrium strategy.)

The model has some attractive features: It captures the importance of the distribution of cognitive capacities that presumably exists in the general population, the effect of limited cognitive capacity on choices, and the role of learning. All of these factors must play a role in the design of contracts and therefore in the proper judicial treatment of them. But the model does not refine the basic intuition with which we started, that contract doctrine might have something to do with mistake, lack of foresight, and similar effects of cognitive limitations.

The role of a model of bounded rationality in normative analysis of contract law is also obscure. If parties cannot foresee certain events, then legal rules will not affect their incentives, and courts can do what they want when those events occur. If parties can foresee the events but fail to think about them fully and accurately, then the possibility of useful judicial intervention remains open. But an accounting of the costs and benefits of this intervention must await a more fully worked out theory of bounded rationality.

E. *A Return to Doctrinalism?*

If interdisciplinary approaches to contract law cannot generate plausible descriptive or normative theories, should legal scholars return to doctrinal analysis? To answer this question, we must first be clear about what doctrinalism means.

Imaged with the Permission of Yale Law Journal

The most ambitious doctrinal scholarship attempts to derive principles from cases. Fuller and Perdue argued that a "reliance principle" explains contract damages.[127] More recently, Eisenberg has argued that a "bargain principle" and a "fairness principle" explain contract-enforcement doctrines,[128] and Farnsworth has proposed "dependence" and "public interest" principles, among many others.[129]

As many have observed, cases will not yield principles that are more general than the case outcomes themselves. The plausibility of the principles that scholars advance always comes from their appeal to moral commitments. The extraction of a fairness principle from the cases, rather than a principle of fairness for litigants who have brown eyes (if such is the case) or for litigants who have some other characteristic coextensive with the cases in which they prevail, is always the result of an implicit appeal to an attractive normative idea—fairness for all, rather than fairness to a morally arbitrary group of people. Ambitious doctrinal scholarship thus converges to a kind of moral philosophy that is especially sensitive to judicial outcomes,[130] and is thus vulnerable to the criticisms of philosophical analysis described in Section V.A.

Examples of authors who write in this vein include Fuller and Perdue, who argue that the reliance principle follows from corrective justice, and Eisenberg, who attempts to tie the bargain principle and fairness principle to policy concerns and moral commitments. Farnsworth avoids philosophizing or engaging in policy analysis by keeping his discussion vague. For example, he does not explain how he resolves conflicts between the many principles that he invokes.[131]

A narrower kind of doctrinal analysis is nothing more than ordinary legal analysis, in which a judge or lawyer explains whether or not a given precedent controls the case under consideration. This kind of doctrinalism is useful, and can be done well or poorly, but a return to this scholarship would have to count as a defeat for the descriptive and normative aspirations of modern legal theory. Doctrinalism does not describe or justify the law; it is simply the use of legal materials and techniques of reasoning to determine the outcome in a given case, or to reconcile or criticize cases that have disparate outcomes.

---

127. Fuller & Perdue, *supra* note 104.

128. Melvin A. Eisenberg, *The Bargain Principle and Its Limits*, 95 HARV. L. REV. 741 (1982).

129. E. ALLAN FARNSWORTH, CHANGING YOUR MIND: THE LAW OF REGRETTED DECISIONS (1998).

130. That is, something like Dworkinian interpretivism.

131. For a critique of Fuller and Perdue's argument, see Richard Craswell, *Against Fuller and Perdue*, 67 U. CHI. L. REV. 99 (2000). For a critique of Farnsworth's argument, see Eric A. Posner, *Law and Regret*, 98 MICH. L. REV. 1468 (2000).

Imaged with the Permission of Yale Law Journal

VI. CONCLUSION

The title of this Essay is a question, not a statement, for two reasons. First, the answer can come only with more experience. As economists and lawyers experiment with new models and variations of old ones, they might find better approaches to understanding contracts and contract law. The answer might also turn out to be "sort of," depending on whether efforts to model bounded rationality, should they succeed, ought to be considered a vindication of economics or psychology.

Second, economics has already accomplished much, just not what its proponents set out as the measure of success. If you look at the best work in contracts scholarship today and compare it with the best contracts scholarship before the 1970s, you will see many differences. One important difference is that earlier work was methodologically sloppy. Much of this work mixed up two separate tasks: excavating the doctrine and evaluating it. Evaluation would often be based on poorly articulated notions of fairness—intuitions that other commentators as well as judges might or might not share—with either no attention to the effects of doctrine on incentives, or casual discussion. The failure to distinguish doctrine and policy often resulted in the displacement of the policy disagreement into rule/standard debates. For example, the voluminous literature on the unconscionability doctrine in the 1960s and 1970s was vague on the incentive effects of the doctrine—with some concern about interfering with freedom of contract, and some concern about unequal bargaining power—and vigorous on the question whether an ambiguous standard like unconscionability could be applied by courts as consistently as they applied similar supposedly rule-like doctrines, such as duress.[132] No one seemed to understand that the rules/standard question presupposed a resolution of the policy question.

The economic literature on the unconscionability doctrine clarified the policy questions at stake and largely displaced the earlier literature. Its main accomplishment was showing that the earlier policy arguments were ill-defined, or made implausible empirical assumptions, or were inconsistent with widely held views or other uncontroversial areas of law and policy. Defenses of the unconscionability doctrine are now more candid and clearer, even if they reject economic premises. The literature as a whole proceeds at a higher level of sophistication.

Economics, then, ushered in a set of scholarly virtues that have improved the literature. These virtues include consistency in the use of terms, clarification of the stakes of the discussion, the distinction between normative and positive analysis, and isolation of different incentives and

132. *See, e.g.*, M.P. Ellinghaus, *In Defense of Unconscionability*, 78 YALE L.J. 757 (1969).

Imaged with the Permission of Yale Law Journal

behaviors. The literature now speaks in an economic idiom, with concepts like transaction cost, risk aversion, default rule, and efficiency substituting for similar but vaguer notions in the earlier writings. These are important accomplishments, and it is hard to imagine serious contract-law research in the future that does not reflect the influence of economics.

But economics fails to explain contract law. It does not explain why expectation damages are the standard remedy, for example, or why liquidated damages are not always enforced. It does not explain the function of the consideration doctrine or promissory estoppel. It does not explain why the law sometimes encourages people to disclose information and at other times does not.

And economics provides little normative guidance for reforming contract law. Models that have been proposed in the literature either focus on small aspects of contractual behavior or make optimal doctrine a function of variables that cannot realistically be observed, measured, or estimated. The models do give a sense of the factors that are at stake when the decisionmaker formulates doctrine, and might give that decisionmaker a sense of the trade-offs involved, but in the absence of information about the magnitudes of these trade-offs—and the literature gives no sense of these magnitudes—the decisionmaker is left with little guidance.

Imaged with the Permission of Yale Law Journal