CONTRACT DESIGN AND THE STRUCTURE OF CONTRACTUAL INTENT

JODY S. KRAUS† & ROBERT E. SCOTT‡

Modern contract law is governed by a two-stage adjudicative regime—an inheritance of the centuries-old conflict between law and equity. Under this regime, formal contract terms are treated as prima facie provisions that courts can override by invoking equitable doctrines so as to substantially "correct" the parties' contract by realigning it with their contractual intent. This ex post judicial determination of the contractual obligation serves as a fallback mechanism for vindicating the parties' contractual intent whenever the formal contract terms fall short of achieving the parties' purposes. Honoring the contractual intent of the parties is thus the central objective of contract law. Yet little scholarly attention has been given to the structure of contractual intent. Courts naturally equate contractual intent with the parties' contractual objectives, which we call the "contractual ends" of their collaboration. But reaching agreement on a shared objective is only the first step to designing an enforceable contract. Thereafter, the parties must create the particular rights and duties that will serve as their "contractual means" for achieving their shared ends. The thesis of this Article argues that the current regime of contract adjudication conflates the parties' contractual means with their contractual ends. In so doing, it reduces the range of contractual arrangements to which contract law gives effect, thereby potentially depriving commercially sophisticated parties of essential tools for contract design. Sophisticated actors engage in ex ante determinations of their means of enforcement, choosing whether enforcement is to be either legal or relational and whether legal enforcement should rely on either rules or standards. Both theory and available evidence suggest that such parties would prefer a default rule that strictly enforces formal contract doctrine unless they have expressly indicated their intent to delegate hindsight authority to a court. By eliminating the risk that courts will erroneously infer the parties' preference for ex post judicial intervention, such a regime increases the reliability of formal contract terms and enhances the parties' control over the content of their contract.

---

† Robert E. Scott Distinguished Professor of Law and Professor of Philosophy, University of Virginia.

‡ Alfred McCormack Professor of Law and Director of the Center for Contract and Economic Organization, Columbia University. We thank Lucian Bebchuk, Richard Craswell, Mel Eisenberg, Victor Goldberg, Avery Katz, Greg Klass, Dan Klerman, Paul Mahoney, Alex Raskolnikov, Alan Schwartz, Rip Verkerke, and participants at workshop presentations at the Northwestern University School of Law, the University of Virginia School of Law, and the American Law and Economics Association for helpful comments on prior drafts.

INTRODUCTION

Modern contract law is governed by a two-stage adjudicative regime—an inheritance of the centuries-old conflict between law and equity. Under this regime, formal contract terms are treated as prima facie provisions that courts can override by invoking equitable doctrines if they believe that doing so is necessary to substantially "correct" the parties' contract by realigning it with their contractual intent. For example, if the parties' objective is to provide the seller a three percent profit, they might specify a term that requires the buyer to pay the seller a price equal to three percent above a published industry price index. If the seller claims it is entitled to more than three percent above the published price index, a court would first apply formal contract doctrine to determine the price term of the agreement.[1] However, if a court believes that the published price index severely underrepresents the seller's actual costs, it can invoke an equitable doctrine to override the formal contract price term and to substitute a price term it believes more accurately represents the seller's true costs.[2] This ex post judicial determination of the contractual obligation serves as a fallback mechanism for vindicating the parties' intent whenever a court determines that the formal contract terms fall seriously short of achieving the parties' purposes. In short, under the current adjudicative regime, every contract by default comes with a judicial insurance policy against formal contract terms that, in the court's view, turn out to have ill-served the parties' intentions.

Honoring the contractual intent of the parties is the central objective of contract law.[3] Yet despite its theoretical and doctrinal cen-

---

[1] In this Article, we characterize as *formal* any American contract doctrine that originated in the legal rules developed and strictly enforced by English common law courts. We will refer to the terms that a court would find by the standard application of formal contract doctrine to the evidence of the parties' agreement as the *formal contract terms*. Thus, the formal contract price term in this example is three percent above the published price index; the parties appear to have chosen this term, rather than a term directly entitling the seller to its actual costs plus three percent, as their contractual means of achieving their contractual end (providing the seller with a three percent profit). *See infra* pp. 119–21 (defining "formal contract law").

[2] We characterize as *equitable* any American contract doctrine that originated from the principles developed and enforced by the English Court of Chancery. The historical purpose of these doctrines was to allow the chancellor to circumvent or override the common law rules whenever he believed doing so was necessary to avoid injustice. In the case of contracts, these doctrines typically treat as unjust any outcome that is dramatically misaligned with the parties' contractual ends. *See infra* pp. 119–22 (defining "equitable contract law").

[3] The search for intention is a key doctrinal element in determining whether the parties have made a binding agreement, the meaning that attaches to the terms of that agree-

trality, little scholarly attention has been given to the structure of contractual intent.[4] Courts naturally equate contractual intent with the parties' contractual objectives, which we will call the *contractual ends* of their collaboration. But reaching agreement on a shared objective is only the first step to designing an enforceable contract. Thereafter, the parties must create the particular rights and duties that will serve as their *contractual means* for achieving their shared ends.[5] The thesis of this Article is that the current regime of contract adjudication conflates the parties' contractual means with their contractual ends. In so doing, it reduces the range of contractual arrangements to which contract law gives effect, thereby potentially depriving commercially sophisticated parties of essential tools for contract design. Sophisticated actors engage in ex ante determinations of their means of enforcement, choosing whether enforcement is to be legal or relational and whether legal enforcement should rely on rules or standards. We argue, therefore, that commercially sophisticated parties would prefer a regime in which courts apply formal doctrine exclusively, unless at the time of formation the parties have expressly indicated their desire for courts to apply equitable doctrine as well.[6]

Our claim that sophisticated parties would prefer courts to honor their formal contract terms even when doing so will frustrate their contractual ends may appear to be unsupportable either in theory or in practice. After all, the contemporary two-stage regime for contract adjudication rests on the natural and powerful intuition that most parties, including commercially sophisticated ones, would prefer courts to take advantage of their hindsight in assisting the parties to achieve

---

ment, the default terms implied in that agreement, and whether the obligations of the parties are impliedly conditioned on unstated assumptions or expressly conditioned on stated events. In addition, the obligation to honor the intention of the parties is the key linkage between efficiency and autonomy, two of the major theoretical approaches to contract. ROBERT E. SCOTT & JODY S. KRAUS, CONTRACT LAW AND THEORY 23–29 (4th ed. 2007).

[4] A recent exception to the lack of sustained analysis of the nature of contractual intent is Gregory Klass, *Intent to Contract*, 95 VA. L. REV. (forthcoming Oct. 2009).

[5] These rights and duties comprise the contract terms that determine how and whether the parties must perform their contractual obligations as well as the consequences of nonperformance.

[6] For a precise definition of "sophisticated economic actors" that limits the category to limited partnerships, professional partnerships, and firms organized in corporate form with five or more employees, see Alan Schwartz & Robert E. Scott, *Contract Theory and the Limits of Contract Law*, 113 YALE L.J. 541, 545 (2003). In principle, our thesis might also be applied to contracts between commercially unsophisticated parties or between commercially sophisticated and unsophisticated parties. But defending an extension of our thesis to such cases would require additional analysis to take into account a range of considerations, such as legal information asymmetry, cognitive error, and nonrepeat play contexts, which we do not undertake here.

their contractual ends.[7] Indeed, had the parties known at the time of formation what the court knows at the time of adjudication, the parties themselves would have crafted different terms. It appears to follow that the parties would prefer courts to do for them in the course of adjudication what the parties would have done for themselves at the time of formation had they known what the court knows. It would seem perverse for a court to insist on holding parties to terms that it knows the parties themselves would have rejected. To hold parties to their formal contract terms when those terms no longer (or never did) constitute reasonable means of achieving the parties' intended contractual ends would exalt formal doctrine over substance.

As compelling as it seems, however, this justification of the two-stage regime of contract adjudication rests on the unsupported premise that contract law should identify the parties' contractual intent with their intended ends rather than their intended means. Sometimes the only way to maintain fidelity to the parties' contractual intent is to enforce the formal contract terms to which they agreed, even when doing so defeats their contractual ends. To return to the example above, by specifying an industry index to determine the seller's costs, the parties might have intended to reject recourse to judicial hindsight in order to increase the certainty of their contract price. The parties might believe that the benefits of judicial hindsight are outweighed by the gains in the reliability of the deal and the decrease in expected litigation costs. When one party claims that the index has not functioned as expected, it invites a court to second-guess the contract's formal price term. But it seems likely that the parties here specified an objective price index precisely to avoid the costs of a judicial inquiry into its accuracy. Therefore, to entertain such an inquiry would perversely defeat the parties' contractual intent by undermining the very means they chose to achieve their contractual ends. Judicial intervention to "correct" this contract by realigning the price term with the seller's actual costs thus would vindicate the parties' contractual ends at the expense of undermining their chosen means. A regime that embraces such intervention impairs the ability

[7] Ex ante, parties have limited knowledge of the present and are uncertain about the future. Therefore, in order to decide on their particular contractual obligations, the parties must make assumptions about their present and future circumstances. Yet they also know that new information that comes to light during the term of their contract may reveal that their assumptions were erroneous. This justification for the two-stage regime thus rests on the premise that, by delegating to courts the historically equitable authority to modify formal contract terms, the parties can both rely on their formal contract terms but also rest safe in the assurance that, if their assumptions turn out to be false, a court will set aside the formal contract terms and craft substitute terms that better serve the parties' initial contractual ends in light of the new information available to the court.

of parties in the future to rely on proxies such as a price index in lieu of an ex post judicial inquiry into actual costs.

The two-stage regime thus ultimately destroys the distinction between the parties' contractual means and ends. It presumes that all parties want courts to advert directly to the parties' contractual ends to determine their contractual rights and duties. But there is good reason to doubt that commercially sophisticated parties typically, let alone always, prefer this method of vindicating their contractual intent. Both theory and available evidence suggest that parties would prefer a regime that strictly enforces formal contract doctrine (unless they have expressly indicated their intent to delegate hindsight authority to a court) over a default rule that automatically subordinates formal contract doctrine to ex post judicial revision.[8] By eliminating the risk that courts will erroneously infer the parties' preference for ex post judicial intervention, such a regime increases the reliability of formal contract terms and enhances the parties' control over the content of their contract. That control, in turn, permits sophisticated commercial parties to implement the most efficient contract design strategies available to them.

In this Article, we develop a theory of contract design that explains why commercially sophisticated parties would prefer a regime that delegates to them complete discretion over the methods for determining the specific content of their contractual obligations. We begin with the premise that in the absence of contracting costs, commercial parties would seek to maximize the expected joint gains from contracting by specifying unambiguously each party's performance obligations in all possible future states of the world. However, parties face two obstacles to writing such "complete contingent contracts": front-end transaction costs and back-end enforcement costs. Collectively, these comprise the costs of contracting. The first axiom of our theory of contract design is that contracting parties are motivated to select those contracting mechanisms that will best achieve their objectives at the least cost. Parties economize on contracting costs in two distinct ways: by choosing between legal and relational means of enforcement[9] and by shifting costs between the front end

---

[8] The core of our argument supporting the claim that commercial parties prefer courts to adhere to formal contract terms absent an express indication otherwise is based on a theory of contract design explicated in Parts II and III, *infra*. Our theory of contract design explains why commercial parties must be able to anticipate and control what portions of their agreements will be legally enforceable and the extent to which their agreements will be governed by rules ex ante or standards ex post. For empirical support for the predictions of the theory, see *infra* notes 148–50, 194–98, 276–78 and accompanying text.

[9] The literature has long recognized two complementary perspectives on enforcement. On the one hand, prospective litigation encourages compliance with those bargains that

and the back end of the contracting process when drafting legally enforceable agreements.[10]

Consider first the tradeoff between legal and relational enforcement. In experimental settings, legal enforcement has been shown to undermine relational norms based on reputation, repeated dealings, and reciprocity. This research has demonstrated that, in many instances, these relational mechanisms operate as substitutes, rather than as complements, for legal enforcement.[11] Moreover, the relational mechanisms that may be displaced by a system of legal enforcement are often less costly and more effective than the legal alternative. In particular, relational mechanisms for motivating performance of contractual obligations are likely to be optimal whenever key information is *observable* but not easily *verifiable*: The parties themselves can obtain the information at reasonable cost, but the costs of proving it to a third-party adjudicator exceed the gains from legal enforcement.[12] To be sure, in some circumstances legal enforcement may be preferable even though it undermines relational norms.

---

satisfy the prerequisites for legal enforcement. On the other hand, parties also face an array of informal or "relational" sanctions if they fail to honor their commitments. In many cases, these relational sanctions may do much of the enforcement work. *See* ROBERT E. SCOTT & PAUL B. STEPHAN, THE LIMITS OF LEVIATHAN: CONTRACT THEORY AND THE ENFORCEMENT OF INTERNATIONAL LAW 84–109 (2006) (discussing choice between formal and informal enforcement); *see generally* Robert E. Scott, *Conflict and Cooperation in Long-Term Contracts*, 75 CAL. L. REV. 2005 (1987) [hereinafter Scott, *Conflict and Cooperation*] (discussing potency of relational sanctions as means of enforcing agreements that are not legally enforceable); Robert E. Scott, *A Theory of Self-Enforcing Indefinite Agreements*, 103 COLUM. L. REV. 1641 (2003) [hereinafter Scott, *Self-Enforcing Agreements*] (same). We discuss the tradeoff between formal and relational enforcement in Part II.B, *infra*.

[10] For discussion, see Robert E. Scott & George G. Triantis, *Anticipating Litigation in Contract Design*, 115 YALE L.J. 814, 822–51 (2006).

[11] *See* Scott, *Self-Enforcing Agreements*, *supra* note 9, at 1670–72 (describing experiments showing that reciprocal fairness rather than explicit sanctions can be used as enforcement device); Robert E. Scott & Paul B. Stephan, *Self-Enforcing International Agreements and the Limits of Coercion*, 2004 WIS. L. REV. 551, 579–80 [hereinafter Scott & Stephan, *Self-Enforcing International Agreements*] (discussing experiments suggesting that coercion undermines reciprocity).

[12] There is no sharp dichotomy between verifiable and nonverifiable conditions. As we discuss in detail in Part II.B.1, *infra*, the question is whether the benefits of using informal norms to enforce a difficult-to-prove condition (such as the level of effort needed to comply with a contractual commitment) are greater than the alternative of verifying compliance with a less accurate but more easily established proxy for the condition in question. The important point, however, is that informal enforcement mechanisms can take into account conditions that are hard to verify even where legal enforcement cannot. For example, parties to an agreement often can observe whether one party has exercised "best efforts" to perform its obligation, but it will be quite costly to marshal the evidence necessary to demonstrate this fact to a disinterested third party. Where this is true, exclusive reliance on legal enforcement can deprive parties of relational mechanisms that can promote better compliance at a lower cost.

The complexity of particular transactions may make it difficult for either the parties or casual observers to determine confidently whether a party has departed from the agreed course of conduct, but the information necessary to make that decision may subsequently be provable to a court at reasonable cost. In such cases, the threat of a legally enforceable sanction might deter opportunistic behavior and clarify the parties' actions and responsibilities. This calculus of relative costs and benefits often motivates parties to partition their agreements into legally enforceable and unenforceable components. A court that breaches the resulting barrier in order to vindicate the parties' contractual ends undermines the efficacy of this first choice of contractual means.

When parties decide to make some or all of their agreement legally enforceable, they face a choice among contractual means that further optimize contracting costs. By choosing between vague or precise contract terms, the parties can thereby shift costs between the back and front end of the contracting process.[13] When the parties agree to vague terms, such as the obligation to use their best efforts, the subsequent adjudication of contractual disputes concerning best efforts will require a court to give concrete and specific meaning to the vague phrase as part of the process of interpretation. By framing their agreement in vague terms, the parties embed their legal obligations in broad standards that delegate discretion to courts ex post and thereby increase back-end enforcement costs. However, this strategy saves the parties the costs of specifying precise performance obligations at the time of formation. Moreover, because the court will have information that was not available to the parties at the time of formation, it potentially allows the parties to benefit from more efficient performance standards than they could have specified at the time of formation.[14]

Alternatively, when the parties agree to precise terms, such as the obligation to pay for a fixed quantity of goods at a price determined by a specified index, the parties reduce the need for a court to interpret the terms of the contract. Precise terms reduce legal obligations to bright-line rules that specify the content of the obligation ex ante. Rules withdraw authority from courts to determine particular performance obligations and instead direct them to enforce the formal obligations that the parties have explicitly specified in advance. Par-

---

[13] We discuss the tradeoffs between ex ante rules and ex post standards in contract design in Part III.B.2, *infra*.

[14] *See* Scott & Triantis, *supra* note 10, at 845 ("Given the court's superior information, the parties can expect that one or both of the proxies will be less noisy [at the time of trial] than the one that the parties would pick ex ante.").

ties might do this because they are better informed than courts about their contractual purposes and have better incentives to pursue them efficiently. This strategy, however, does require the parties to rely on estimates of the likelihood of various future events rather than the actual knowledge of those events that would be available to a court at a later date. If those estimates turn out to be erroneous, a court may be tempted to realign the contract terms with the parties' original contractual ends. But only if courts faithfully enforce the precise terms, notwithstanding subsequent frustration, can parties choose ex ante between precise and vague terms as the best means to achieve their particular contractual ends.[15]

Our theory of contract design thus explains why commercial parties need to anticipate and control what portions of their agreements will be legally enforceable and the extent to which their agreements will be governed by ex ante rules or ex post standards.[16] When faithfully applied, formal contract doctrine provides parties with precisely this control. In the discussion that follows, we consider a number of examples where courts apply formal doctrine in a nonstandard fashion or invoke historically equitable doctrines to override formal doctrine when they perceive that enforcing the parties' formal contract terms would defeat the parties' contractual ends. In particular, we analyze how courts interpret the parol evidence and integration doctrines, the mistake and excuse doctrines, and the law of conditions and waiver in order to avoid adjudicative outcomes that they perceive to be misaligned with the parties' contractual ends. Our aim is not to demonstrate that the court in any particular case necessarily reached an improper result. After all, these contracts were formed against the backdrop of the current two-stage regime of contract adjudication. These parties thus may have designed their contracts on the assumption that a court might advert to equitable doctrine and reasoning. On this question, we claim only that, contrary to the courts' reasoning, the formal terms of these contracts cannot be dismissed on the ground that they are at odds with any rational or reasonable understanding of the parties' contractual intent. Our principal objective, therefore, is to

---

[15] The parties' objective is to select that combination of precise rules and vague standards that optimizes their total costs of contracting. Note that the parties' goal is not simply to minimize contracting costs; parties will incur additional contracting costs so long as those costs result in improvements in the parties' incentives to maximize the expected surplus from the contract. Thus, the parties' objective is to "maximize the incentive bang for the contracting-cost buck." *Id.* at 823; *see also id.* at 840–41 (illustrating concept with hypothetical example).

[16] For a penetrating analysis of the role that rules and standards play in differentiating textualist and intentionalist statutory interpretation, see Caleb Nelson, *What is Textualism?*, 91 VA. L. REV. 347, 398–416 (2005).

use these cases to explain why sophisticated commercial parties, such as the parties to these contracts, might prefer their future contracts to be adjudicated under a regime that applied formal doctrine exclusively, unless the parties indicate otherwise at the time they form their contract.

The Article proceeds as follows. Part I provides a concise overview of the origins and evolution of equitable doctrines in American contract law and claims that the root cause of the problem we identify is the historical English practice of subordinating law to equity. This Part shows how deeply ingrained the equitable mindset is in the history and current practice of adjudication in American contract law. Part II turns to contract design theory, focusing particularly on the strategies that inform the choice between legal and relational means of enforcement. Formal contract doctrines that govern the interpretation of agreements enable parties to partition their agreement into legally enforceable and legally unenforceable components. The faithful application of formal contract doctrine thus permits parties to choose the optimal mix of legal and relational commitments, while ex post equitable intervention impairs this choice. Part III examines the factors informing the choice between precise and vague contract terms as a means of effecting contractual ends. Here the theory of contract design explains why parties might reasonably choose ex ante to forswear doctrines of mistake and excuse even where strict enforcement of formal contract terms may subject them to the risk of severe hardship. Finally, Part IV uses the law of conditions to illustrate the tension inherent in the current two-stage adjudicative regime. We show the deep rift within the common law between the formal contract doctrines requiring strict enforcement of formal terms and equitable doctrines permitting abrogation of those terms when strict enforcement appears to frustrate the parties' contractual ends. The law of conditions shows how the logic of equity invites courts to distort formal doctrine and to interpret express conditions in ways that better align with the parties' contractual ends.

We conclude that the challenges of contract design argue for permitting sophisticated parties the freedom to use combinations of legal and relational norms, as well as contractual rules and standards, to optimize their contract in light of uncertainty about the future and the difficulties of proof. Since design choices are best made by the parties ex ante, both theory and the available empirical evidence suggest that commercial parties would prefer a regime in which equitable override of formal contract doctrine is invoked only if specifically requested at the time the parties form their agreement.

I
Law and Equity in American Contract Law

The current regime of American contract adjudication applies two different sets of doctrines to resolve a contract dispute. The first set consists largely of those doctrines that originated in the first seven centuries of adjudication in King's Bench and Common Pleas, the English courts that produced the corpus of the English common law from the twelfth century until their abolition in the nineteenth century.[17] These doctrines tend to consist of rules that are administered strictly, without exceptions for cases in which the application of a rule appears to defeat its purpose.[18] They also tend to be cast in objective terms that minimize the need for subjective judgment to apply them.[19] As noted above, we refer to contemporary contract doctrines that share these characteristics—and in many cases share their content—as *formal contract law.* We refer to the interpretive stance typical of the English judges who applied common law doctrines from medieval times through the nineteenth century as *formalistic.* The *formal terms* of a contract are those terms that a formalistic judge applying formal contract law would identify and enforce in the course of adjudicating a dispute requiring interpretation and enforcement of the contract.

The second set of doctrines consists largely of equitable principles that originated in the English Court of Chancery, which began to exercise overlapping jurisdiction with the common law courts by the end of the fourteenth century.[20] These doctrines originally consisted of broad principles administered loosely and were designed to provide alternatives and exceptions to the common law. They also tended to

---

[17] J. H. Baker, An Introduction to English Legal History 12, 114 (4th ed. 2002). The third and oldest common law court was the Exchequer of Pleas. *Id.* at 47. Because this court exercised a "relatively minor civil jurisdiction" during the fourteenth, fifteenth, and early sixteenth centuries, it is not essential to understanding the evolution of the divide between law and equity. *Id.* at 48.

[18] *See infra* notes 29–39 and accompanying text (describing approach of King's Bench and Common Pleas); *see also* 2 E. Allan Farnsworth, Farnsworth on Contracts §§ 9.1, 9.5 (2d ed. 1998) (discussing exceptions to general rule that contractual duties are absolute).

[19] *See, e.g.*, David J. Ibbetson, A Historical Introduction to the Law of Obligations 28 (1999) (describing medieval action of debt according to which "a document under seal granting money from one person to another would be strictly enforced").

[20] The Chancery had two kinds of jurisdiction: The first, denominated its "Latin side" because its forms and actions were written in Latin, was to make purely administrative rulings; the second, denominated its "English side" because its bills and pleadings were in English, was to adjudicate the bills of complaint that fell within the original jurisdiction of the King's council. Baker, *supra* note 17, at 100–01. "It seems probable that the English jurisdiction was established in its distinct form during the reign of Richard II . . . . It was perhaps firmly settled while John of Waltham was master of the rolls (1381–86)." *Id.* at 103.

be cast in subjective terms and therefore often required judges to exercise discretion on a case-by-case basis. We refer to contemporary contract doctrines that share these characteristics, and in many cases also share the content of the Chancery courts' equitable principles, as *equitable contract law*. We will use the term *equitable* to refer to the interpretive stance typical of the English chancellor that applied these equitable principles from medieval times until the nineteenth century when law and equity were formally fused and the Chancery courts abolished.[21]

In this Part, we describe the characteristics that distinguish not only the content of formal and equitable contract doctrine but, more importantly, the dramatically opposed interpretive stances that each kind of doctrine invites judges to adopt. Although some of the equitable doctrines incorporated into American contract law are applied routinely by courts adjudicating contract disputes, others are perhaps most notable for how rarely they are successfully invoked. For example, while plaintiffs can in some circumstances enforce oral agreements by asserting the part-performance exception to the Statute of Frauds,[22] the doctrines of mistake and excuse are rarely successful defenses to breach of contract claims.[23] Despite their infrequency, however, these latter doctrines cast a significant shadow of uncertainty across all contracts.[24] By extension, we will argue that equitable reasoning undermines efficient design not only when courts apply equitable doctrines directly, but also when courts use equitable reasoning to guide their application of formal doctrine to the facts of

---

[21] *See id.* at 114 (discussing abolition of Chancery courts and fusion of law and equity).

[22] 2 FARNSWORTH, *supra* note 18, § 6.9.

[23] RESTATEMENT (SECOND) OF CONTRACTS ch. 11, introductory note at 309–10 (1981) ("Contract liability is strict liability. It is an accepted maxim that *pacta sunt servanda*, contracts are to be kept. The obligor is therefore liable in damages for breach of contract even if he is without fault and even if circumstances have made the contract more burdensome or less desirable than he had anticipated. . . . An extraordinary circumstance may make performance so vitally different from what was reasonably to be expected as to alter the essential nature of that performance. In such a case the court must determine whether justice requires a departure from the general rule that the obligor bear the risk that the contract may become more burdensome or less desirable."); RESTATEMENT (SECOND) OF CONTRACTS ch. 6, introductory note at 379 ("The law of contracts supports the finality of transactions lest justifiable expectations be disappointed. This Chapter deals with exceptional situations in which the law departs from this policy favoring finality and allows either avoidance or reformation on the ground of mistake."); *see also* 2 FARNSWORTH, *supra* note 18, § 9.1.

[24] Such uncertainty may be increased by the fact that "[i]n recent years, courts have shown increasing liberality in discharging obligors on the basis of [certain] extraordinary circumstances." RESTATEMENT (SECOND) OF CONTRACTS ch. 11, introductory note at 309–10 (1981). For further discussion, see *infra* Part III and the court's analysis in *Aluminum Co. of America v. Essex Group, Inc. (ALCOA)*, 499 F. Supp. 53 (W.D. Pa. 1980).

contract disputes.[25] The most profound effect of equity on American contract law, therefore, is its gravitational influence on the judicial application of formal doctrine in contract disputes.[26]

The presence of equitable doctrines in American contract law is widely recognized, but the extent of equity's influence has not been fully appreciated. Our objective in this Part, therefore, is to explain the depth and breadth of the influence of equity in American contract law by providing a concise overview of equity's roots in English legal history and its incorporation into the doctrines and reasoning of American contract adjudication. This overview explains when, how, and why the logic of equity took hold in the history of contract adjudication and argues that, despite its eventual merger with formal doctrine, equity remains today a significant force deeply antithetical to the logic of formal doctrine.

A. The Historical Origins of Law and Equity

1. The Creation of Courts of Law and Equity

From their inception, the King's Bench and Common Pleas courts entertained actions only by plaintiffs who purchased an original royal writ, which specified the type of claim that the plaintiff was authorized to bring and the kind of relief to which the plaintiff would be entitled should he prevail.[27] Because the right to bring an action at the common law courts derived from the writs issued by the King's Chancery, the content of judicially cognizable rights were defined by the forms of action authorized in the writs.[28] If parties had complaints that did not fit within the confines of existing forms of action, they could petition the Chancery to issue a new form.[29] When the Chancery's power to issue new writs was curtailed in the thirteenth century,[30] parties began to petition the King himself, since he retained

---

[25] For example, in Part II.A.2, we argue that the court in *Hunt Foods & Industries, Inc. v. Doliner*, 270 N.Y.S.2d 937 (App. Div. 1966), misapplied the parol evidence and integration doctrines because it implicitly embraced equitable reasoning that equates the parties' contractual intent with their intended contractual ends instead of means.

[26] As we explain in Part IV, the gravitational influence of equity on law reaches its zenith in the law of conditions, which explicitly directs courts to disregard formal doctrine in order to avoid the imposition of a forfeiture on one party, even if a neutral application of formal doctrine would otherwise produce a different result.

[27] BAKER, supra note 17, at 53–54.

[28] *Id.* at 54–55.

[29] *Id.* at 55 ("A plantiff . . . had either to find a known formula to fit his case, or apply for a new one to be invented.").

[30] *See id.* at 100 ("A late thirteenth-century writer describes [chancellors] as hearing petitions and complaints, which they determined by issuing writs; though by that time the discretion to invent new remedies was severely curtailed."). As explained, "[i]n the thirteenth century one possible response to a petition had been to allow a new form of original

authority to hear cases in which he believed the common law was "deficient."[31] The King therefore heard cases alleging that the common law courts had failed to do justice because they had either refused to entertain a just claim or had resolved a claim unjustly. In exceptional cases, the King took action by granting a "remedy as of grace."[32]

As these "exceptional" private suits became more common, they were referred to the King's council.[33] The King's council originally received bills alleging that parties or officials were "frustrating the common law."[34] The King's council either resolved such bills itself or passed them on to Parliament or to the chancellor for resolution.[35] Litigants began to anticipate that their bills would be sent to the chancellor and thus eventually addressed their bills directly to him.[36] Under the authority of the council, the chancellor then took responsibility for assigning them to appropriate courts for resolution.[37] Thus, although the Chancery did not originally have the authority to adjudicate claims itself,[38] its power to issue the original writs and to receive bills, coupled with the likelihood that the King's council would ultimately refer extraordinary suits to the chancellor, led parties to petition the chancellor himself to provide relief where the common law courts did not.[39] Eventually, the chancellor began "to issue process and grant decrees in Chancery instead of sending the petition else-

---

writ, and when that power was curtailed the bill might be referred to parliament for a legislative solution . . . ." *Id.* at 102.

31 *Id.* at 98. As one historian has explained, the King "retained an overriding residuary power to administer justice outside the regular system; but the important limitation imposed on that power by the due-process legislation was that it could be invoked only where the common law was deficient, and never in matters of life, limb or property." *Id.*

32 *Id*. ("By the end of the thirteenth century numerous petitions (or 'bills') were being presented to the king, asking for his grace to be shown in respect of some complaint.").

33 *Id.* ("Already in the fourteenth century the petitioning of the king by bill, seeking a remedy as of grace, was so common that such business had to be referred to special sessions of the council or parliament . . . .").

34 *Id*.

35 *Id.*

36 *Id.* at 101.

37 *Id.*

38 *Id.* at 99, 101.

39 *Id.*; *see also id.* at 101–02 ("The Chancery may have been thought an appropriate place to furnish new remedies because of its traditional supervision of the issue of original writs. A plaintiff applying for an original writ was in a sense making a petition in Chancery."). Before the end of the thirteenth century, the Chancery always had the power to create a new writ that would provide a form of action suitable to a plaintiff's complaint. But when the plaintiff's claim was based on idiosyncratic facts rendering existing forms inadequate, rather than a common complaint for which no form of action existed, the Chancery sought an ad hoc solution rather than the creation of a new form of action. *Id.* at 102.

where."[40] By the end of the fourteenth century, the Court of Chancery had separated from the council and exercised independent, "extraordinary" jurisdiction to hear cases when "the ordinary course of law failed to provide justice."[41] By the end of the fifteenth century, "it became a common saying that no deserving plaintiff would be sent out of the Chancery without a remedy."[42]

2. The Contrasting Approaches of Law and Equity

The Chancery's willingness to provide an independent and alternative forum for justice stemmed from its perception that the administration of justice in the common law courts was hamstrung both by the procedural constraints imposed by the forms of action and rules of evidence and by the strict, rule-bound inclination of common law judges to apply the common law rigorously without sympathy for the poor, oppressed, weak, foolish, or careless.[43] In short, King's Bench and Common Pleas courts *defined* justice as the outcome that resulted from the strict application of (their) due process and substantive common law doctrines to the cases that came before them. In this sense, the common law courts provided justice wholesale: Common law judges "preferred to suffer hardship in individual cases than to make exceptions to clear rules."[44] In contrast, "the chancellor's eyes were not covered by the blinkers of due process, and he could go into all the facts to the extent that the available evidence permitted."[45] Because the chancellor viewed the due process and substantive rules of the common law as mere means of achieving justice, he set them aside whenever he believed that they failed to produce a just result. In short, the Chancery provided justice retail, according to the chancellor's own view of what justice required in each individual case.

---

[40] *Id.* at 101.

[41] *Id.* at 117.

[42] *Id.* at 102.

[43] *See id.* at 102–03 (noting that Chancery operated "a court of conscience" that was "free from the rigid procedures" of common law courts).

[44] *Id.* at 102; *see also id.* at 325 ("Under the harsh logic of the common law it was 'better to suffer a mischief to one man than an inconvenience to many, which would subvert the law.'" (quoting Waberly v. Cockerel, (1542) 73 Eng. Rep. 112, 113 (K.B.))).

[45] *Id.* at 104. In its earliest incarnation, the procedure in Chancery was the antithesis of the procedure in common law courts: No writ was necessary, multiple issues could be joined, evidence was taken free of formal rules, decisions were made by a chancellor rather than a jury, the court was always open, and trials could take place anywhere (including the chancellor's home). *Id.* at 103.

Glaston v. Abbot of Crowland[46] provides a classic example of the hard-edged common law reasoning from which the Chancery provided relief. In that case, a promisee sued to enforce a penal obligation on a bond.[47] The promisor had previously paid the debt to the promisee but failed to destroy the deed.[48] The promisee subsequently stole the deed and brought an action of debt against the promisor to secure double payment. Under prevailing common law doctrine, a valid deed provided conclusive proof of the debt,[49] so a promisor could plead the defense of payment only by producing another deed (an acquittance under seal) proving that he had paid the debt. The common law court therefore refused the promisor's motion to admit oral evidence to prove that the promisee had in fact stolen the deed back after the promisor had paid him.[50] The common law reasoning was that it was the promisor's "own folly not to have had [the deed] destroyed or to have obtained an acquittance under seal."[51] The Chancery, however, was "free from the rigid procedures under which such injustices sheltered. His court was a court of conscience, in which defendants could be coerced into doing whatever conscience required in the full circumstances of the case."[52] Thus, if presented with a case such as *Glaston*, the chancellor could ignore the common law rules of evidence, admit the oral testimony, and order the bond cancelled in order to serve justice.[53]

The fundamental opposition of the adjudicative approaches of the common law courts and the Chancery is illustrated by their contrasting views of other cases like *Glaston*. In *Waberley v. Cockerel*, the court summed up the common law's rationale for enforcing its rules without regard to the justice or fairness of the outcome in the case at bar:

> [It is] better to suffer a mischief to one man than an inconvenience to many, which would subvert a law: for if matter in writing may be so easily defeated, and avoided by such surmise and naked breath, a

---

[46] Glaston v. Abbot of Crowland (1330), *reprinted in* SOURCES OF ENGLISH LEGAL HISTORY: PRIVATE LAW TO 1750, at 252 (J.H. Baker & S.F.C. Milsom eds., 1986) [hereinafter SOURCES OF ENGLISH LEGAL HISTORY].

[47] *Id.*

[48] *Id.*

[49] Fishacre v. Kirkham (1289), *reprinted in* 112 SELDEN SOCIETY 322 (1996).

[50] BAKER, *supra* note 17, at 324–25 (citing *Glaston*, *reprinted in* 98 SELDEN SOCIETY 665 (1982), *supra* note 46, at 252; Donne v. Cornwall, Y.B. 1 Hen. 7, fol. 14v, Pasch, pl. 2 (1486), *reprinted in* SOURCES OF ENGLISH LEGAL HISTORY, *supra* note 46, at 255).

[51] BAKER, *supra* note 17, at 325.

[52] *Id.* at 103 (footnote omitted).

[53] *See id.* at 104 (explaining that chancellor "could order bonds and other writings to be cancelled where they would only serve unjust ends").

> matter in writing would be of no greater authority than a matter of fact.[54]

The *Waberley* court thus justified the rigidity of the common law rule on the ground of its prospective effect. Rephrased in contemporary economic terms, the court's rationale is that the writing requirement for proof of payment lowers the expected costs of enforcing bonds and thereby decreases the costs of lending and borrowing money. Injustices will find shelter under such a rule only if borrowers either fail to destroy the deed upon payment or to secure a written acquittance, both of which are far less costly than the expected costs of adjudicating oral payment defenses. When such injustices occur, they reveal the failure of borrowers either to understand the legal rules governing their transactions or to take measures to protect their interests in light of those rules. Thus, the common law judges treated doctrines as devices for *prospective regulation*. To create long-term prospective benefits for the entire population, the common law rules had to be known in advance and the costs of complying with them reasonable. Thus, the courts could not adjust their holdings according to the impact of the rules on the parties to an individual case at the time of adjudication.

In contrast, the Chancery's sole focus was on just and fair *dispute resolution*. Its concern was with the equities of the case at bar, not the prospective effects of its ruling. Indeed, for many years the Chancery's decrees had no formal precedential effect,[55] which initially freed the Chancery from any concern that its rulings could undermine the consistency and predictability of adjudication. Thus, even though the Chancery sometimes reversed or avoided outcomes decided by common law courts,[56] it understood these actions to be necessary in order to vindicate, rather than undermine, the common law.[57] While the common law courts were responsible for creating and administering a set of rules to guide individuals in the future, the Chancery's role was to stand behind those courts and provide ad hoc remedies for

---

[54] Waberley v. Cockerel, (1542) 73 Eng. Rep. 112, 113 (K.B.).

[55] *See* BAKER, *supra* note 17, at 104 ("In Chancery each case turned on its own facts, and the chancellor did not interfere with the general rules observed in courts of law. The decrees operated *in personam*; they were binding on the parties in the cause, but were not judgments of record binding anyone else."); *see also id.* at 202 ("So long as chancellors were seen as providing ad hoc remedies in individual cases, there was no question of their jurisdiction bringing about legal change or making law.").

[56] *See id.* at 202 ("In Chancery the just remedy was provided not by changing the law but by avoiding its effect in the special circumstances of particular cases.").

[57] Councillors and chancellors viewed themselves as "reinforcing the law by making sure that justice was done in cases where shortcomings in the regular procedure, or human failings, were hindering its attainment by due process. They came not to destroy the law, but to fulfil it." *Id.* at 102.

perceived injustices that, in its view, slipped through the cracks of the common law as applied in individual cases.[58] By circumventing the writing requirement in cases such as *Glaston*, the Chancery viewed itself not as subverting the common law but rather as preventing the common law from being subverted by the lender. The premise of the Court of Chancery was that the common law courts were often powerless to prevent litigants from invoking black-letter doctrine for unjust purposes but the Court of Chancery was not.[59] By its own lights, the Chancery existed to work the common law pure.[60]

Fundamentally, however, the institutions of the common law and the Chancery were at cross purposes. The common law's prospective regulatory enterprise viewed the adjudication of cases primarily as a means of creating and sustaining a system of rules justified by its aggregate effect over time. In contrast, the Chancery's dispute resolution enterprise viewed the adjudication of cases as an end in itself, in which the sole objective was to do justice between the parties. The result was two competing systems of adjudication, often with incompatible procedural and substantive doctrines, yet overlapping in jurisdiction.[61] For a brief period, the two systems were directly opposed, with equity literally undoing the final judgments of common law cases.[62] Moreover, the subjectivity of the Chancery's practice of deciding cases according to the chancellor's conscience led to inconsistent, and therefore unpredictable, decisions within the Chancery.[63] As the number and importance of the cases before it grew dramati-

---

[58] *See id.* at 103 (observing that Chancery was "free from the rigid procedures" of common law courts and could "coerce[ ]" defendants "into doing whatever conscience required in the full circumstances of the case").

[59] The common law courts were not wholly insensitive to the illegitimate use of legal doctrines, however. For example, the common law of contracts refused to enforce contracts supported by illegal or immoral consideration. IBBETSON, *supra* note 19, at 211.

[60] In *Jurisdictional Competition and the Evolution of the Common Law*, 74 U. CHI. L. REV. 1179 (2007), Daniel Klerman argues that competition between common law courts led to simple, pro-plaintiff rules because judges were paid, in part, by court fees, and plaintiffs chose the forum in which to bring their claims. Chancery therefore attracted disgruntled defendants who petitioned it to enjoin or otherwise interfere with common law suits. Because Chancery faced no real competition, it had no incentive to reduce costs through simplification of doctrine. Klerman argues that all the equitable doctrines of American contract law, other than those granting injunctive relief, resulted from Chancery's exercise of original jurisdiction in order to compete against common law courts for business.

[61] *See* Robert E. Scott, *Chaos Theory and the Justice Paradox*, 35 WM. & MARY L. REV. 329, 334–36 (1993) (noting that common law "dilemma" led to "formal separation of courts of law and equity" and discussing overlapping jurisdictions of these courts).

[62] In 1616, James I issued a decree confirming the chancellor's jurisdiction to entertain suits in equity even after judgments in law had been handed down on the same matters. James I's decree remained in effect until it was declared illegal in 1670. BAKER, *supra* note 17, at 108–09.

[63] As Baker explains,

cally and the practice of reporting its reasoning and decisions became regularized, the Chancery eventually created general principles of equity to constrain discretion and enhance predictability. Many equitable principles matured into a system of rules as clear as law[64] and with equally binding precedential effect.[65] Yet equitable doctrine was still "more flexible than the common law, because it [took] greater account of individual circumstances."[66]

B. The Legal and Equitable Roots of American Contract Law

The system of equitable principles created by the Chancery left an indelible impression on the contemporary common law of the United States. Historically, the division between the Kings Bench and Common Pleas courts and the court of Chancery acted as a blood-brain barrier between the matter of the common law and the anti-matter of equity. In the nineteenth century, the Chancery was eliminated and law and equity were merged in both England and the United States.[67] The result was an incoherent combination of legal and equitable doctrines, officially interwoven into the tangled skein comprising American common law. To this day, American common law is torn between the prospective regulatory bent of historically legal doctrines and the retrospective, dispute-resolution focus of historically equitable doctrines. Conflicts among these doctrines were inevitable. When England officially considered the question of how

---

> The essence of equity as a corrective to the rigour of law was that it should not be tied to rules. If, on the other hand, no consistent principles whatever were observed, parties in like cases would not be treated alike; and equality was a requisite of equity. As John Selden quipped in the mid-seventeenth century, if the measure of equity was the chancellor's own conscience, one might as well make the standard measure of one foot the chancellor's foot.

Id. at 109.

64 During the seventeenth century, equitable doctrines began to congeal into formal rules, although their objective was still to provide alternatives and exceptions to the common law. "By 1676 a chancellor could repudiate the idea that equity had any dependence on his own inner conscience: '[T]he conscience by which I am to proceed is merely *civilis et politica*, and tied to certain measures.' Thus equity hardened into law." *Id.* at 110 (footnote omitted); *accord* Scott, *supra* note 61, at 336.

65 Although equity still operated to blunt and sometimes circumvent the common law, within its own rules "even the Chancery would sooner suffer a hardship than a departure from known rules." BAKER, *supra* note 17, at 110.

66 *Id.* at 110–11.

67 Ironically, by the nineteenth century, the Chancery had developed a set of procedures more arcane and burdensome than the common law procedures it originally sought to mitigate. The resulting administrative delay, combined with corruption born of the Chancery's practice of paying clerks on a fee basis rather than a salary, ultimately led to the Chancery's demise. *See id.* at 111–12 (explaining major practical defects that developed in Chancery). Soon thereafter law and equity were merged. *Id.* at 114.

such conflicts should be resolved, it answered clearly that "equity should prevail."[68] American courts and commentators followed suit.

The American common law of contracts is thus suffused with historically legal doctrines, providing relatively clear and objective rules, combined with historically equitable doctrines directing courts to circumvent or override these rules whenever a judge believes that application of the legal doctrine would produce a result that is contrary to the doctrine's own purpose or otherwise unjust. American contract doctrines originating in the English common law courts include the doctrines of offer and acceptance,[69] consideration (including the illegality and immorality doctrines),[70] capacity,[71] duress,[72] warranties and conditions,[73] impossibility,[74] fraud-in-the-execution,[75] expectation damages,[76] foreseeability,[77] and avoidability,[78] as well as the plain meaning rule[79] and the parol evidence rule.[80] Along with these historically legal contract doctrines, American contract law also absorbed and developed doctrines originally developed in Chancery "to mitigate the rigours of the Common law."[81] Such doctrines include fraud-in-the-inducement and intentional misrepresentation,[82] negligent and

---

[68] Judicature Act, 1873, 36 & 37 Vict., c. 66, § .25(11) (Eng.), *reprinted in* CHALONER W. CHUTE, EQUITY UNDER THE JUDICATURE ACT 209 (1874) ("[I]n all matters . . . in which there is any conflict or variance between the Rules of Equity and the Rules of the Common Law with reference to the same matter, the Rules of Equity shall prevail.").

[69] IBBETSON, *supra* note 19, at 222.

[70] *Id.* at 203–04.

[71] *Id.* at 208–09.

[72] *Id.* at 71–72.

[73] *Id.* at 223–25.

[74] *Id.* at 228; *see also* Taylor v. Caldwell, (1863) 122 Eng. Rep. 309, 314 (K.B.) (holding that "a condition is implied that the impossibility of performance arising from the perishing of the person or thing shall excuse the performance").

[75] The English common law precursor to the American fraud-in-the-execution doctrine prohibited enforcement of agreements procured by one party misleading the other as to the contents of a writing under seal. *See* IBBETSON, *supra* note 19, at 72.

[76] *Id.* at 87–90.

[77] *Id.* at 229–32.

[78] *See* 2 FARNSWORTH, *supra* note 18, § 12.12 (discussing Vertue v. Bird, (1677) 84 Eng. Rep. 1000, 86 Eng. Rep. 200 (K.B.), to illustrate avoidability doctrine).

[79] The common law applied "to documents a rule of construction that the words had to be given their ordinary meaning." *Id.* at 226.

[80] BAKER, *supra* note 17, at 324–25.

[81] IBBETSON, *supra* note 19, at 203. In general, equity evolved contract doctrines designed to provide far broader protection against perceived fraud than the common law provided. In particular, the core equitable contract doctrines provided relief where an agreement was not fully voluntary or informed. *Id.* at 208.

[82] This doctrine first required that a defrauded party bring a deceit action in trespass, then allowed the action to be brought as trespass on the case, and finally permitted it to be brought in assumpsit. *See* BAKER, *supra* note 17, at 329–39. As an action outside of contract, it required the defrauded party to allege that his contractual partner intentionally breached a confidence or intentionally misrepresented present facts in order to induce the

innocent misrepresentation,[83] fraudulent nondisclosure,[84] unilateral and mutual mistake,[85] specific performance and other injunctive relief,[86] expanded versions of the common law doctrines of capacity and duress,[87] and illegal and immoral consideration.[88] American contract law also adopted equitable doctrines specifically designed to vitiate clear common law rules: the penalty doctrine,[89] the forfeiture

---

defrauded party to enter into the agreement (the equivalent of the modern defense of fraud-in-the-inducement). By the time the action could be brought in assumpsit, however, its intentional component had been discarded, thus laying the foundation for the contemporary doctrine of warranty liability. *See id.* at 331–33, 337.

[83] The equitable defenses of negligent or innocent misrepresentation were the precursors to the contemporary doctrines of fraudulent and material misrepresentation. *See* IBBETSON, *supra* note 19, at 208; *see also* RESTATEMENT (SECOND) OF CONTRACTS §§ 162, 164 (1981). Strictly speaking, neither contemporary contract nor tort imposes liability or provides relief for negligent misrepresentation, save in rare cases in which promissory estoppel is invoked to serve that purpose. Robert E. Scott, Hoffman v. Red Owl Stores *and the Myth of Precontractual Reliance*, 68 OHIO ST. L.J. 71, 92 (2007). But the contemporary doctrines of fraudulent and material misrepresentation in the Restatement conceivably could be construed to void contracts induced by negligent misrepresentation. RESTATEMENT (SECOND) OF CONTRACTS § 162(1)(b), (c) (1981). Originally, the equitable anti-fraud doctrines operated to bar relief for promisees but did not affect the promisor's right to sue at law. *See* IBBETSON, *supra* note 19, at 209.

[84] The equity defense of wrongful silence was the precursor to the contemporary nondisclosure doctrine. For examples of cases discussing the equity defense of wrongful silence, see *Broderick v. Broderick*, (1713) 24 Eng. Rep. 369 (Ch.), and *Chesterfield v. Janssen*, (1751) 28 Eng. Rep. 82 (Ch.). *See* IBBETSON, *supra* note 19, at 208 ("Central to the Chancery's intervention was the power to relieve against fraud . . . . The most obvious case of this was where one person deliberately made a false statement in order to lure another into a bargain, but it could extend much further than this: a representation by conduct might equally constitute fraud, as might a representation by silence."); *see also* RESTATEMENT (SECOND) OF CONTRACTS §§ 161, 164 (1981).

[85] IBBETSON, *supra* note 19, at 210.

[86] *Id.* at 206, 213; *see also* BAKER, *supra* note 17, at 320 ("The scope of specific performance was unclear, and plaintiffs seeking such a remedy were drawn into the Chancery.").

[87] The common law capacity doctrine was confined to prohibiting enforcement of agreements made by women and infants. Equity provided a more expansive capacity defense based on impaired mental capacity, including drunkenness in certain circumstances. The common law duress doctrine was limited to voiding agreements procured by serious violence or imprisonment. Equity expanded duress to include undue influence. IBBETSON, *supra* note 19, at 209.

[88] *See id.* at 210–11 (explaining that Chancery and common law courts were concerned with one party "taking unfair advantage" of another and "were quick to sniff out any whiff of illegality"). Although both the common law and equity refused to enforce illegal contracts, equity went further to refuse enforcement of contracts it deemed to be against public policy. *See id.* at 211–12.

[89] The penalty doctrine voids any contract clause providing for liquidated damages in excess of the parties' actual or expected compensatory damages. "In the sixteenth century the Chancery began to mitigate this by issuing injunctions against the enforcement of penalties in an initially limited range of situations . . . ." IBBETSON, *supra* note 19, at 150.

> By the seventeenth century liability in contract was seen as absolute, in the sense that, once the parties had reached an agreement, they would in principle

doctrine,[90] the equitable exceptions to the parol evidence rule,[91] and the part-performance exception to the Statute of Frauds.[92]

The English Chancery is therefore the original source of the two-tiered system of American contract doctrine. The English common law controlled the disposition of a case before the common law courts, unless the chancellor exercised his discretion to invoke overriding principles of equity to reverse or provide an alternative to that disposition. Similarly, American contract law typically is controlled only prima facie by common law doctrines. Equitable doctrines abound, inviting courts to avoid the strictures of these rules if doing so would lead to a more fair, just, or equitable result in any given case. Although courts typically resist the urge to invoke equitable doctrines in routine cases, the likelihood of equitable intervention in cases with unusual circumstances is significant.[93] And even in less exceptional cases, the presence of equitable doctrines throughout the common law has exerted pressures that subtly affect the interpretive stance of judges applying the strict common law doctrines of contract law. Thus, rather than simply choosing the most natural of two plausible interpretations of contract language, judges will often choose the

---

> be held to it unless the defendant could point to duress, fraud, or some other vitiating factor. Consistently with this position, the courts' remedies would normally give effect to the agreement . . . . This principle was subject to the important qualification that the courts would not enforce penalties. The origins of this lie in the Chancery practice of the late fifteenth century . . . .

Id. at 213. Penalties were also forbidden by the statutes of 1669 and 1705. BAKER, *supra* note 17, at 325. For discussion of the evolution of the contemporary penalty doctrine, see Robert E. Scott & George G. Triantis, *Embedded Options and the Case Against Compensation in Contract Law*, 104 COLUM. L. REV. 1428, 1440–45 (2004).

[90] BAKER, *supra* note 17, at 202–03. The forfeiture doctrine sets aside implied and express conditions. *See* RESTATEMENT (SECOND) OF CONTRACTS § 229 (1981) ("Excuse of a Condition to Avoid Forfeiture: To the extent that the non-occurrence of a condition would cause disproportionate forfeiture, a court may excuse the non-occurrence of that condition unless its occurrence was a material part of the agreed exchange."); *see also id.* § 225 cmt. a ("Where discharge would produce harsh results, this . . . effect may be avoided by rules of interpretation or of excuse of conditions." (citation omitted)); *id.* § 227, cmt. b–c. We discuss the inherent tension between formal law and equity embedded in the forfeiture doctrine in Part IV.B, *infra*.

[91] *See, e.g.*, JOSEPH STORY, COMMENTARIES ON EQUITY JURISPRUDENCE, §§ 153–157 (W.E. Grigsby ed., 1884) (explaining that courts of equity allow parol evidence to modify contracts upon clear evidence of mistake).

[92] Even before the Statute of Frauds was passed, equity had created the part performance exception to the common law writing requirement. IBBETSON, *supra* note 19, at 203 n.4. The equitable doctrine of part performance continued to apply after the statute took effect. BAKER, *supra* note 17, at 350.

[93] *See, e.g.*, Aluminum Co. of Am. v. Essex Group, Inc. (*ALCOA*), 499 F. Supp. 53 (W.D. Pa. 1980) (finding that unpredictable increase of electricity costs entitled seller to reformation of long-term contract when seller stood to lose amount in excess of sixty million dollars).

interpretation that produces the most congenial result, even if that interpretation is the less natural one.[94]

Indeed, the modern law of excuse, frustration, and impracticability follows the nineteenth-century English rationalization of the equitable doctrine of mistake. That doctrine, which explicitly informs the American mistake doctrine, holds that, as a matter of interpretation, it is presumed that the parties to an agreement are not mistaken as to any basic and material assumption about a present fact at the time of formation.[95] The excuse, frustration, and impracticability doctrines simply extend this presumption to mistakes about future facts. These doctrines thus formalize the interpretive stance that requires judges to interpret express contract language, or imply contract terms, in whatever manner proves necessary to avoid an adjudicative result that strikes the court as unfair, unjust, or inequitable in light of facts known to the parties at the time of adjudication.[96] In most cases, courts treat as unjust and unfair any result that diverges significantly from the parties' presumed intended contractual ends.

The combination of law and equity in modern American contract law is widely viewed as unproblematic.[97] This view is colored by the unexamined premise that courts use these doctrines in a straightforward manner to vindicate contractual intent. But once the concept of intent is itself partitioned into means and ends, the rift between law and equity becomes apparent. In the remainder of this Article, we illustrate how both equitable reasoning and equitable doctrines continue to exert a significant influence on the resolution of contract disputes. Our aim is to show how judicial efforts to vindicate contractual intent by "doing equity" can perversely undermine that intent and thereby undermine the ability of sophisticated commercial parties to design efficient contracts through the use of formal contract doctrine.

---

[94] *See, e.g.*, Hunt Foods & Indus., Inc. v. Doliner, 270 N.Y.S.2d 937 (App. Div. 1966) (looking to parol evidence rather than written contract for evidence of parties' intent); Corthell v. Summit Thread Co., 167 A. 79 (Me. 1933) (holding that although company had sole discretion to determine amount of payment for inventions per contract terms, company owed reasonable compensation for inventions). For further discussion, see *infra* Part II.A.2 and *infra* Part IV.C.1.

[95] RESTATEMENT (SECOND) OF CONTRACTS §§ 152, 153 (1981).

[96] See *infra* Part III.A for a discussion of the equitable doctrines of mistake, excuse, and frustration in the specific context of *ALCOA*.

[97] Indeed, some scholars applaud the role of equity in contract adjudication. *See, e.g.*, Melvin A. Eisenberg, *The Responsive Model of Contract Law*, 36 STAN. L. REV. 1107 (1984) (rejecting formalism in favor of "responsive" contract law that mirrors individualized and subjective approach of equity).

II
Legal and Relational Means of Effecting Contractual Ends

In this Part, we analyze the formal rules of interpretation, in particular the parol evidence rule and integration doctrines. We use a well-known contracts case to illustrate the ways in which courts use equitable reasoning to effectively set aside the parties' written contract terms by applying formal doctrines in a nonconventional fashion. In this class of cases, courts conclude that the parties could not reasonably have believed that the formal contract terms to which they agreed constituted a rational means of pursuing their contractual ends; judicial intervention is thus necessary to vindicate the parties' "true" contractual intent. Contract design theory shows, to the contrary, that the parties' formal contract terms may well have constituted a rational means of pursuing the parties' objectives. By refusing to enforce the parties' formal contract terms, courts prevent parties from implementing those contract designs that require strict adherence to formal contract doctrine.

A. Contractual Interpretation and Intent

1. The Function of Doctrines of Interpretation

The formal doctrines of contract interpretation, such as the parol evidence and plain meaning rules, are widely understood to provide parties with the means to express their contractual intent. The sole purpose of both doctrines is to give parties some control over the process courts will use to interpret their intent.[98] The parol evidence rule enables parties to control the admissibility of certain kinds of evidence in any future adjudication of disputes over their agreement. The rule holds that when parties choose to fully integrate a final written agreement, they forfeit the right in subsequent litigation to prove understandings they declined to include in their integrated writing.[99] There are a number of reasons why parties may wish to exclude evidence of certain understandings reached during the contracting process: Tentative agreements may be abandoned, agents may misrepresent their principal's commitments, or, importantly for our purposes, the parties may deliberately choose to exclude a portion of their agreement from legal enforcement. They may do this in order to rely on self-enforcing mechanisms, such as reputation, the prospect of future dealings, and

[98] Scott & Kraus, *supra* note 3, at 541.

[99] *See id.* (noting that parol evidence rule allows parties to agree not to introduce "extrinsic evidence" in court).

the propensity to reciprocate.[100] The common law thus has consistently treated the decision to integrate an agreement entirely as a matter of party discretion. Courts have no presumption for or against a finding that a written agreement is integrated. Instead, they have devised various neutral tests for determining whether parties intended to integrate part or all of their agreement into a final, legally enforceable writing.[101]

Similarly, the plain meaning rule is best understood as a device for preserving a reservoir of terms with clear meanings that cannot be contradicted in adjudication by contextual evidence supporting a different meaning.[102] The best rationale for the plain meaning rule is that it makes available a public fund of terms with judicially "proofed" or "protected" meanings on which contractual parties can rely to effectively communicate their commitments to each other and to courts.[103] The parol evidence and plain meaning rules therefore can be viewed as tools parties can use to either restrict or provide the evidence that courts will use to interpret the portion of the agreement that the parties intend to make legally enforceable.

This straightforward account of the interpretation doctrines as tools for parties to express their contractual intent is relatively uncontroversial. No one would be surprised to learn that the central interpretation doctrines of contract law can be understood as devices for effectuating contractual intent. But courts often perceive a conflict between the rules governing contractual interpretation and the principle of honoring the expressed intention of the parties.[104] Courts therefore sometimes behave as if they face a dilemma: either maintain fidelity to the language and purpose of the various interpretive doctrines, even though this leads to an outcome that conflicts with the parties' expressed intention, or maintain fidelity to the parties' inten-

---

100 *See infra* text accompanying notes 129–33 (discussing possible motivations of parties in *Hunt Foods*).

101 *See* SCOTT & KRAUS, *supra* note 3, at 542–43 (reviewing how common law and Uniform Commercial Code identify terms of agreement in contract).

102 *See* 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, 273 (1985) (explaining that plain meaning rule prohibits external evidence from being used to interpret terms of contract if terms appear unambiguous in document); Schwartz & Scott, *supra* note 6, at 584–90.

103 *See* Goetz & Scott, *supra* note 102, at 316–17. ("[S]killful use of the [plain-meaning] presumption by courts will, over time, increase the supply of officially recognized invocations and other express conventions.").

104 *See infra* text accompanying notes 109–21 (discussing one court's analysis of whether terms "consistent" and "contradict" in Uniform Commercial Code's parol evidence rule should be treated as synonymous).

tion even though this requires an application of the interpretive doctrines that conflicts with the doctrines' language and purpose.

Courts confronted with this apparent dilemma are predisposed to deploy their equitable powers to avoid the application of formal contract doctrine that appears to achieve an unfair or unjust result. But the dilemma is false. It conflates contractual intent with the parties' intended contractual ends. By doing so, courts effectively reduce the range of available contractual means for best achieving contractual ends. To the extent that a more constrained choice of means forecloses optimal design strategies, this judicial predisposition also reduces the expected gains from contracting. Since sophisticated parties wish to maximize the expected gains from contracting, they would prefer courts to maintain fidelity to those formal doctrines that facilitate optimal contract design without regard to their effect on the parties' intended contractual ends in any particular contract. We can best support this claim—and in particular illustrate the value of retaining the choice between legal and relational means of enforcement—by focusing on a paradigmatic case.

2. Hunt Foods v. Doliner*: The Parol Evidence Rule and Relational Norms*

Consider the case of *Hunt Foods & Industries, Inc. v. Doliner.*[105] Hunt Foods was interested in acquiring the assets of Eastern Can Company.[106] To that end, it entered into extensive negotiations with George Doliner and his family, who collectively owned 73% of Eastern Can's stock. The parties reached agreement on the price to be paid for the stock, but before they could agree on the remaining issues, including the form of the acquisition, they decided to recess the negotiations for several weeks. Fearing that Doliner would use the break to shop the offer to other bidders, Hunt Foods requested an option to buy the Doliner stock. The parties executed a written agreement for the purchase of an option that made no mention of any conditions on the exercise of the option.[107]

Negotiations resumed before the option's expiration date, and the parties could not reach agreement on the remaining issues. Following this impasse, Hunt Foods exercised the option according to its written terms. Doliner refused to tender the stock, and Hunt Foods moved for summary judgment on its suit for specific performance. Doliner opposed the motion on the grounds that there was an oral

---

[105] 270 N.Y.S.2d 937 (App. Div. 1966).
[106] *Id.* at 939.
[107] *Id.*

understanding between the parties that the option, though unconditional on its face, was only to be used in the event that Doliner solicited an outside offer for the stock. Hunt Foods contended that there was no such condition and that, in any event, evidence of such an oral understanding was barred by the parol evidence rule. Because the underlying transaction was a purported contract for the sale of securities, Article 8 of the Uniform Commercial Code (U.C.C.) applied and, by extension, so did the parol evidence rule as embodied in U.C.C. § 2-202.[108]

a. The Court's Analysis

The court denied Hunt Foods's motion for summary judgment on the grounds that the alleged oral understanding was not barred by the U.C.C.'s parol evidence rule.[109] The court cited to U.C.C. § 2-202, which provided that terms in a writing intended by both parties as final "may not be *contradicted* by evidence of any prior agreement or of a contemporaneous oral agreement but may be explained or supplemented . . . by evidence of *consistent* additional terms . . . ."[110] Evidence of such additional terms is admissible unless the court finds that the writing is fully integrated (perhaps by a merger clause) and thus is not only final as to the agreed terms but exclusive as well.[111]

The court's analysis focused on the question of whether the proffered evidence was of a "consistent additional term."[112] The proffered condition was "clearly [an] 'additional'" term because it was not set out in the writing.[113] The remaining question was whether the condi-

---

[108] *Id.* at 939–40 & n.1 ("While article 2 of the Uniform Commercial Code which contains this section does not deal with the sale of securities, this section applies to article 8, dealing with securities. . . . All parties and Special Term so regarded it." (applying 1966 version of U.C.C.)).

[109] *Id.* at 940.

[110] *Id.* (emphasis added); *see also* U.C.C. § 2-202 (amended 2003) (stating essentially same rule as that articulated in *Hunt Foods*).

[111] *Hunt Foods*, 270 N.Y.S.2d at 940.

[112] *Id.* The question of whether the proffered evidence is consistent with the writing does not arise until a court has first determined that the agreement is at least partially integrated. The court in *Hunt Foods* never expressly found the agreement to be partially integrated. Instead, the court claimed that, under then comment 3 to U.C.C. § 2-202, the parties would not have necessarily included the condition in their writing and the court appeared to conclude implicitly that the writing was therefore not fully integrated with respect to the condition. *Id.*; *see also* U.C.C. § 2-202 cmt. 3 (amended 2003) (containing current version of comment relied on in *Hunt Foods*). Ideally, the *Hunt Foods* court would first have found expressly that the parties intended the written option term as their final expression of that term before turning to the question of whether the condition contradicts that term. If the court did not believe the agreement was partially integrated, then there would be no need to exclude the evidence of the condition even if it did contradict the written option term.

[113] *Hunt Foods*, 270 N.Y.S.2d at 940.

tion "contradicted" a term in the written option and whether the condition was "consistent" with the writing.[114] Answering these questions requires a court to interpret the meaning of the statutory language. Is a term "consistent" as that word is used in § 2-202(b) as long as it does not "contradict" a term in the writing?[115] Or is the requirement that the term be "consistent" an *independent and additional* requirement for admissibility? In other words, are the terms "may not . . . contradict[ ]" and "consistent" synonymous or do they impose two independent tests for admissibility?[116] The latter interpretation—that an admissible additional term must both not contradict another term in the writing *and* be consistent with the terms in the writing—is suggested by the use of different words (with presumably different meanings) in § 2-202. This interpretation would suggest that a term is "inconsistent" where it is not in "reasonable harmony" with a term of the writing even though it did not directly contradict any of the written terms.[117]

The court chose instead to interpret "may not . . . contradict[ ]" and "consistent" as imposing only a single requirement for admitting the proffered additional term into evidence. To justify this conclusion, the court introduced a novel interpretive principle: "To be inconsistent the term must contradict or negate a term of the writing. A term or condition which has a lesser effect is provable."[118] In other words, even if the written option was properly interpreted as unconditional, evidence that the parties agreed that the option could be exercised only on condition that Doliner shopped Hunt Foods's offer would not "contradict" the written option because it merely "lessened" the effect of the option. Presumably, the *Hunt Foods* court would have barred evidence of a term only if it literally negated the written option (i.e., evidence that the parties agreed that Hunt Foods would have no option to purchase the Doliner stock).

The court never addressed explicitly the question of whether the option was an exclusive statement of the terms of the agreement and thus whether evidence of even consistent additional terms would be barred. But, at least implicitly, the court concluded that the writing was only a partial but not a complete integration. The court cited the comment to § 2-202, which sets out a test for complete integration: "If

---

114 *Id.*

115 U.C.C. § 2-202(b) (amended 2003).

116 *Id.*

117 This "two step" interpretation of the U.C.C.'s test for admissibility under § 2-202 was adopted by the court in *Snyder v. Herbert Greenbaum & Associates*, 380 A.2d 618, 623 (Md. Ct. Spec. App. 1977).

118 *Hunt Foods*, 270 N.Y.S.2d at 940.

the additional terms are such that, if agreed upon, they would certainly have been included in the document . . . then evidence of their alleged making must be kept from the trier of fact."[119] The court interpreted this standard as barring oral evidence of a consistent additional term only if the court finds that it is "impossible" that the parties agreed to that term given that they did not include it in their writing.[120] Finding that it was at least *possible* that the parties agreed to the condition, even though they did not include the condition in their written agreement, the court found the agreement was not fully integrated and admitted the oral evidence of the condition.[121]

b. Contractual Ends Trump the Rules of Interpretation

The interpretive gloss that *Hunt Foods* adds to the U.C.C.'s parol evidence rule threatens to undermine the purposes of the doctrine. If the parties intended the written option to be the final expression of the option term, and if it is clear that the proper interpretation of the option as written is that it is unconditional,[122] then allowing evidence

---

[119] *Id.* The Code now embodies this in comment 3. U.C.C. § 2-202 cmt. 3 (amended 2003).

[120] *Hunt Foods*, 270 N.Y.S.2d at 940.

[121] *Id*. The opinion also provides the barest suggestion that the court might have believed that the oral statement was admissible as evidence that the written agreement, even if fully integrated, was subject to an oral condition precedent. Under the rule of *Pym v. Campbell*, (1856) 119 Eng. Rep. 903 (K.B.), courts have allowed oral evidence offered to prove that a fully integrated agreement never came into existence because it was subject to an oral condition precedent. *Id.* at 905. The Restatement admits such evidence on the ground that it proves the agreement was not fully integrated. *See* RESTATEMENT (SECOND) OF CONTRACTS § 217 cmt. b (1981). *But see* 2 FARNSWORTH, *supra* note 18, § 7.4, at 231 n.8 (arguing Restatement explanation is inferior to account that holds agreement is valid but admits oral statement as evidence that agreement did not take effect). The *Hunt Foods* court concluded that "the alleged oral condition precedent cannot be precluded as a matter of law or as factually impossible." 270 N.Y.S.2d at 940. Some commentators have stated that the court's holding therefore should be interpreted as treating the alleged oral statement as evidence that the agreement did not take effect because it was subject to an oral condition that never occurred. *See, e.g.*, 2 FARNSWORTH, *supra* note 18, § 7.4, at 231 & n.9 (noting that qualification of parol evidence rule has been "read into UCC 2-202 even though that section contains no specific language on the point."). Even on this view, however, evidence of the oral condition precedent is admissible only if it does not contradict an express term of the written agreement. *See, e.g.*, Intercontinental Monetary Corp. v. Performance Guars., Inc., 705 F. Supp. 144, 149 (S.D.N.Y. 1989) ("[T]he parol evidence rule does not bar proof of every orally established condition precedent, but only of those which in a real sense contradict the terms of the written agreement." (quoting Hicks v. Bush, 225 N.Y.S.2d 34, 37 (1962))); *see also* 2 FARNSWORTH, *supra* note 18, § 7.4, at 232 (stating that extrinsic evidence is inadmissible if it contradicts written term).

[122] By making the argument that the alleged oral condition would merely lessen the effect of the written option, the court implicitly held that the written option standing alone should properly be interpreted as creating an unconditional option. Other courts have held that a written option that fails to state expressly whether it is conditional may be subject to an interpretive default rule that treats it as unconditional absent evidence that

that the option was conditional undermines the parties' agreement that the written option term is a final expression. Similarly, the court's interpretation of the U.C.C.'s test for integration[123] eviscerates the notion that the terms can be both final and exclusive in the absence of a carefully drafted merger clause.[124] In order to exclude evidence of an additional term on grounds that the agreement was fully integrated with respect to that term, the *Hunt Foods* court required a finding that it would have been *impossible*, not merely *implausible*, for the parties to have intended that term to be part of their agreement but failed to put it in their writing.[125] Such a loose interpretive standard makes it significantly more difficult for parties to protect an agreement from ex post changes to the allocation of contractual benefits and burdens based on alleged additional terms.

Why did the *Hunt Foods* court embrace this controversial analysis of the U.C.C.'s parol evidence rule? The best inference is that the court believed that the parties' contractual ends were inconsistent with the claim that Doliner agreed to grant Hunt Foods an unconditional option. In short, the court seemed to believe that a straightforward application of parol evidence doctrine would frustrate the parties' intent: There was no rational reason why the parties would have intended to exclude the condition from their legally enforceable agreement. After all, Doliner wished to adjourn negotiations without consummating the agreement. Granting an unconditional option would allow Hunt Foods to decide unilaterally to consummate the agreement. The court, however, also noted that Doliner asked his attorney why the written option agreement failed to state the condition.[126] His attorney responded that Hunt Foods "insisted" that the written option not state any conditions, but that the attorney had

---

the parties intended it to be conditional. *See, e.g.*, Masterson v. Sine, 436 P.2d 561, 565 (Cal. 1968) ("The fact that there is a written memorandum [creating an option] . . . does not necessarily preclude parol evidence rebutting a term that the law would otherwise presume."). Such a court would not have reason to treat oral evidence that such a written option is conditional as inconsistent with the written option term. Instead, the oral evidence would be admissible as evidence that the parties opted out of the interpretive default rule. *See id.*

123 This test was found in comment 3 to § 2-202 prior to the 2003 Amendments to Article 2. The pre-Amendment version remains current law in all jurisdictions that have adopted the U.C.C. In the 2003 amended version of Article 2, the same test for integration is found in comment 3. U.C.C. § 2-202 cmt. 3 (amended 2003).

124 Courts have long recognized that a writing can be found to be a total integration even in the absence of a merger clause. *See, e.g.*, Mitchill v. Lath, 160 N.E. 646, 648 (N.Y. 1928) (excluding evidence of separate oral agreement and adding to terms of deed that appeared complete on its face despite absence of merger clause); *see also* SCOTT & KRAUS, *supra* note 3, at 541–43 (discussing tests for total integration in absence of merger clause).

125 *Hunt Foods*, 270 N.Y.S.2d at 940.

126 *Id.* at 939.

"obtained an understanding" that the option was to be used only if Doliner solicited an outside offer.[127] The court never reconciled its implicit conclusion that Doliner would have been willing to grant Hunt Foods only a conditional option with the fact that Doliner acquiesced in Hunt Foods's demand that he sign a written agreement that deliberately omitted an express condition on the option.

The only reasonable conclusion Doliner could have drawn from Hunt Foods's refusal to accept an expressly conditional option is that Hunt Foods was not willing to accept any legally enforceable condition on the option. One might suppose that Doliner was willing to sign the writing without the condition only because he believed that the oral understanding his attorney obtained would be legally enforceable. But if Doliner in fact believed that both he and Hunt Foods intended the oral condition to be enforceable, then he would have been utterly mystified by Hunt Foods's insistence that the condition not be included in the written option. Doliner could not reasonably have believed both that Hunt Foods was content to receive a legally conditional option and that Hunt Foods had good reason to insist that the condition not be stated in the writing.[128] The puzzle, then, is to explain why Doliner might have knowingly agreed to grant Hunt Foods a legally unconditional option even though such a grant seems inconsistent with his desire to preserve his own option to refuse Hunt Foods's current offer. The most plausible explanation requires an understanding of the role that relational norms can play in the efficient design of contracts.

B. A Contract Design Explanation of the Hunt Foods-Doliner Contract

1. Reconstructing the Parties' Motivations

Understanding the relationship between legal and relational enforcement provides a key to resolving the puzzle posed above concerning the *Hunt Foods* contract: Why did Doliner agree, on the advice of counsel, to sign an unconditional option? One can imagine two scenarios in which Doliner valued a break in the negotiations. In the good faith scenario, he was not yet satisfied with all the elements of the agreement, anticipated he would be able to work them out in negotiations, but had a pressing demand that required adjourning

---

[127] *Id*.

[128] If Doliner had concluded that Hunt Foods had agreed to a legally enforceable oral condition but had insisted on excluding the condition from the writing in order to preserve the possibility of denying the condition later, then Doliner obviously would not have acceded to Hunt Foods's demand.

negotiations for several weeks. In the bad faith scenario, Doliner wanted the adjournment because he planned to shop Hunt Foods's bid. Either way, Hunt Foods would understandably be worried that Doliner might shop its bid, so it insisted on an option to protect itself in the event Doliner did so. But, importantly, Hunt Foods understood that an option conditional on proving that Doliner shopped its bid would be of little value in court because it would have great difficulty carrying its burden of proof. Doliner could simply terminate negotiations on independent grounds and then consummate the deal with another bidder to whom he had secretly shopped Hunt Foods's bid during the adjournment. Even if Hunt Foods's officers believed that they could *observe* Doliner's bad faith (e.g., by finding out through sources if he shopped the bid), they could not readily *verify* his bid-shopping to a court.

In short, the costs of verification would likely have precluded the choice of including an explicit bid-shopping condition in Hunt Foods's written option. But the parties had an alternative: Hunt Foods's commitment not to exercise the option unless Doliner solicited another bid was credible because it was subject to powerful informal sanctions. In many instances, an agreement between two such commercial parties will be *self-enforcing* because both parties want to earn and preserve a good reputation in order to gain esteem and future business dealings.[129] Social esteem and a reputation for keeping one's word are powerful motivations whenever other potential trading partners can easily learn why a party's deal broke down.[130] Even in the

---

[129] *See* SCOTT & STEPHAN, *supra* note 9, at 88–94; Schwartz & Scott, *supra* note 6, at 557; Scott, *Conflict and Cooperation*, *supra* note 9, at 2039–42; Scott, *Self-Enforcing Agreements*, *supra* note 9, at 1680–82.

[130] *See* Avner Greif, *Informal Contract Enforcement: Lessons from Medieval Trade*, *in* 2 THE NEW PALGRAVE DICTIONARY OF ECONOMICS AND THE LAW 287, 287 (Peter Newman ed., 1998) ("[I]nformal contract enforcement institutions . . . are a product of the larger economic, cultural, social and political processes of which they are an integral part."); Janet T. Landa, *A Theory of the Ethnically Homogeneous Middleman Group: An Institutional Alternative to Contract Law*, 10 J. LEGAL STUD. 349, 350–52 (1981) (finding that "invisible codes of ethics, embedded in the personalized exchange relations among the members of the [ethnically homogenous middleman group] function as constraints against breach of contract" and that "a rational trader will enter into particularistic exchange relations with traders . . . whom he knows to be trustworthy and reliable in honoring contracts"). As Alan Schwartz and Robert E. Scott have noted:

> [R]eputations work best in small trading communities, especially those with ethnically homogenous members [or other cooperation-inducing structures], where everything that happens soon becomes common knowledge, and boycotts of bad actors are easy to enforce. Reputational sanctions also can be effective in industries that can establish trade associations; the associations become a form of collective memory regarding the contracting behavior of their members.

absence of conditions for establishing reputations, agreements will be self-enforcing to the extent that the parties anticipate the prospect of future dealings: Neither party will breach an early contract if the gains from that breach are lower than the expected profits from future contracts that a breach would eliminate.[131]

To be sure, both of these incentives have natural limits.[132] But these limitations do not justify the conclusion that parties such as Doliner and Hunt Foods would rely exclusively on the legal enforcement of their understanding. There is strong empirical support for the claim that powerful norms of reciprocity enhance and extend the reach of reputation and repeat dealings as means of self-enforcement.[133] In sum, contracting parties frequently can (and do) turn to

---

Schwartz & Scott, *supra* note 6, at 557. For a discussion of formalist contract enforcement by a private system of merchant courts established by the National Grain and Feed Association, resulting in part from a recognition of the role of extralegal, reputational sanctions, see generally Lisa Bernstein, *Merchant Law in a Merchant Court: Rethinking the Code's Search for Immanent Business Norms*, 144 U. PA. L. REV. 1765 (1996) [hereinafter Bernstein, *Merchant Law*]. *See also* Lisa Bernstein, *Private Commercial Law in the Cotton Industry: Creating Cooperation Through Rules, Norms, and Institutions*, 99 MICH. L. REV. 1724 (2001) [hereinafter Bernstein, *Private Commercial Law*] (discussing importance of reputation-based nonlegal sanctions in cotton industry).

131 *See* Schwartz & Scott, *supra* note 6, at 557 ("Reputation . . . will induce performance when a single contract partner's boycott would not.").

132 Reputations are difficult to establish in large, heterogeneous economies in which particular contracting parties are anonymous and ongoing relationships inevitably come to an end. When parties come to realize that the relationship is soon to terminate (say, when the promisor contemplates retirement or otherwise withdraws from the trading community), the threat that the other party will no longer deal with the promisor is insufficient in and of itself to induce performance. *See* Benjamin Klein, *Why Hold-Ups Occur: The Self-Enforcing Range of Contractual Relationships*, 34 ECON. INQUIRY 444, 447–50 (1996) (discussing conditions for self-enforcing contracts); Scott, *Conflict and Cooperation*, *supra* note 9, at 2039–49 (same).

133 Experimental evidence shows that a preference for reciprocity—the willingness to reward cooperation and to punish selfishness—can motivate cooperation even in arms-length interactions between complete strangers. *See generally* Ernst Fehr & Klaus M. Schmidt, *A Theory of Fairness, Competition, and Cooperation*, 114 Q.J. ECON. 817, 818 (1999) [hereinafter Fehr & Schmidt, *A Theory of Fairness*] ("[I]n addition to purely self-interested people, there are a *fraction* of people who are also motivated by fairness considerations."); Ernst Fehr, Simon Gächter & Georg Kirchsteiger, *Reciprocity as a Contract Enforcement Device: Experimental Evidence*, 65 ECONOMETRICA 833, 833 (1997) ("[R]eciprocal behavior may cause an increase in the set of enforceable contracts and may thus allow the achievement of nonnegligible efficiency gains."); David K. Levine, *Modeling Altruism and Spitefulness in Experiments*, 1 REV. ECON. DYNAMICS 593, 594 (1998) (examining quantitative implications of "the theory" that "fairness plays a role in individual decision making"); Matthew Rabin, *Incorporating Fairness into Game Theory and Economics*, 83 AM. ECON. REV. 1281, 1282 (1993) (discussing "a game-theoretic framework for incorporating [reciprocity] into a broad range of economic models"). For a review of literature discussing how fairness considerations "shape the behavior of people in important economic domains," see Ernst Fehr & Klaus M. Schmidt, *Theories of Fairness and Reciprocity—Evidence and Economic Applications* 3–4 (Ctr. for Econ. Studies & Ifo Inst. for Econ. Research, Working Paper No. 403, 2000; Univ. of Zurich, Inst. for Empirical

relational enforcement based on trust, or the desire to maintain a good reputation in the relevant community, or on the prospect of profitable future dealings. In a context such as the negotiations between Doliner and Hunt Foods, these norms may provide the best available means of regulating the contractual relationship and of enforcing promissory commitments. Not only would have the reputation of Hunt Foods been jeopardized if it had behaved opportunistically in the acquisition negotiations, but, more importantly, counsel for both sides could have used their reputations to lend support to the understanding. Since attorneys in merger and acquisition deals are drawn from homogeneous communities, their respective commitments are credible and they can "lend" their reputations to enhance the credibility of their principals' commitments.

What role does legal enforcement play in contexts where relational enforcement is pervasive and robust? The experimental evidence suggests that relational enforcement, when it is effective, is both cheaper and better than legal enforcement.[134] Relational enforcement is cheaper because a party only needs to expend costs to observe the other's behavior, while legal enforcement requires the parties to expend additional resources (e.g., attorneys' fees, court costs, etc.) in verifying that behavior to a court. Less obvious but perhaps even more significant is the fact that relational enforcement is better. It permits parties to make credible promises regarding observable but nonverifiable measures of performance, thus achieving contractual objectives that may not be possible with legal enforcement.[135] In short, there are good reasons why both Doliner and Hunt Foods might have been motivated to exclude the alleged condition from the legally enforceable portion of their agreement.

The preceding reconstruction of motivations is consistent with the facts of *Hunt Foods*. One of two scenarios occurred. Either Hunt Foods never agreed to a condition and Doliner's testimony was untruthful; or, as he testified, the parties agreed *informally* not to exercise their option unless Doliner shopped its bid, but Hunt Foods insisted *formally* on a legally unrestricted option. Doliner would agree to these terms if either he had no intention of shopping Hunt Foods's bid and trusted Hunt Foods to abide by its nonlegally enforceable agreement, or he planned to shop the bid and then render the

---

Research in Econ., Working Paper No. 75, 2001), *available at* http://ssrn.com/abstract=255223. For discussion of how reputation and reciprocity serve as "effective means of self enforcement" of agreements, see Scott & Stephan, *Self-Enforcing International Agreements*, *supra* note 11, at 565–66.

134 SCOTT & STEPHAN, *supra* note 9, at 22–23.

135 Scott, *Self-Enforcing Agreements*, *supra* note 9, at 1667–72.

option useless by claiming that the option was conditional and requiring Hunt Foods to prove the condition was met. Hunt Foods then exercised the option either because Doliner did, in fact, shop the bid or because its officers acted in bad faith by taking advantage of the legally unrestricted option and breaching the informal agreement.

2. *The Costs of Conflating Contractual Intent with the Parties' Contractual Ends*

The court in *Hunt Foods* apparently believed that the parties had no reason to exclude from legal enforcement an agreed-upon condition on the exercise of the option. Assuming the condition was part of the agreement, a strict application of the Code's parol evidence rule would have frustrated the parties' objective of suspending negotiations while protecting Hunt Foods from the risk of bid-shopping. Thus, the most plausible explanation for the unconventional interpretation of the parol evidence rule in *Hunt Foods* is that the court chose to apply the formal doctrine in a nonstandard fashion in order to vindicate Doliner's assumption that the oral understanding was part of the parties' agreement. The court appears to have concluded that a strict enforcement of the formal contract terms would have defeated the parties' contractual intent. But manipulating the doctrine in this manner has the perverse effect of impairing the contractual means that parties use to effect their ends. This is so regardless of whether the parties in *Hunt Foods* deliberately chose relational enforcement or Doliner was simply mistaken in his understanding of the effect of the agreement.

The costs of the two-tiered system of adjudication that implicitly authorizes a court to deploy its equitable powers and manipulate formal contract doctrine are prospective. Without a fully functioning parol evidence rule, parties will have difficulty partitioning their agreement into separate legal and relational components. Instead, the entire agreement will be subject to legal enforcement. The problem with this outcome is that the empirical evidence suggests that legal enforcement is often imperialistic: An effort to superimpose legal enforcement on a regime of relational enforcement is likely to displace the nonlegal mechanisms.[136] For example, experimental evidence shows that reciprocity, operating alone, generates high levels of cooperative behavior between contracting parties.[137] But once the

---

[136] Scott, *Self-Enforcing Agreements*, *supra* note 9, at 1688–92.

[137] Fehr & Schmidt, *A Theory of Fairness*, *supra* note 133, at 817–18; *see also* Fehr, Gächter & Kirchsteiger, *supra* note 133, at 833 ("[R]eciprocal behavior may cause an increase in the set of enforceable contracts and may thus allow the achievement of nonnegligible efficiency gains."); Ernst Fehr & Armin Falk, *Psychological Foundations of Incen-*

entire relationship, including its nonlegal aspects, is subject to legal enforcement, voluntary reciprocity declines along with the overall level of cooperation.[138] These experimental results suggest that legal sanctions and relational sanctions may well conflict with one another.[139] In other words, legal enforcement may "crowd out" behavior based on reciprocity.[140]

The experimental evidence of crowding out supports the claim that commercial parties would prefer formal contract doctrines that permit them to partition their agreements into legally enforceable and

---

tives (Inst. for the Study of Labor, Working Paper No. 507, 2002), *available at* http://ssrn.com/abstract_id=294287 ("[R]eciprocity induces agents to cooperate voluntarily with the principal if the principal treats them kindly."). *See generally* Rabin, *supra* note 133 (incorporating individual's feelings about whether individual is being treated fairly into game theory); Martin Dufwenberg & Georg Kirchsteiger, *A Theory of Sequential Reciprocity*, 47 GAMES & ECON. BEHAV. 268 (2003) (developing model to describe reciprocity in context of sequential games); Fehr & Schmidt, *A Theory of Fairness*, *supra* note 133, at 818 (analyzing conflicting evidence and developing model to explain why "[s]ome pieces of evidence suggest that many people are driven by fairness considerations, other pieces indicate that virtually all people behave as if completely selfish, and still other types of evidence suggest that cooperation motives are crucial"); David K. Levine, *Modeling Altruism and Spitefulness in Experiments*, 1 REV. ECON. DYNAMICS 593 (1998) (proposing simpler model than Rabin's based on altruism); Armin Falk & Urs Fischbacher, *A Theory of Reciprocity* (Ctr. for Econ. Studies & Ifo Inst. for Econ. Research, Working Paper No. 457, 2001), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=203115 (presenting formal theory of reciprocity).

138 Scott, *Self-Enforcing Agreements*, *supra* note 9, at 1688–92.

139 *Id.*

140 To be sure, the crowding-out phenomenon is complex. A number of studies have confirmed the crowding-out hypothesis in single iteration experiments, where the parties must choose either nonlegal or legal enforcement. In these studies, the choice of legal enforcement uniformly suppresses the evidence of reciprocity that is found in the alternative scenario of no legal enforcement. *See generally* Iris Bohnet, Bruno S. Frey & Steffen Huck, *More Order with Less Law: On Contract Enforcement, Trust, and Crowding*, 95 AM. POL. SCI. REV. 131, 131, 141 (2001) ("Individuals perform a contract when enforcement is strong or weak but not with medium enforcement probabilities: Trustworthiness is 'crowded in' with weak and 'crowded out' with medium enforcement."); Fehr, Gächter & Kirchsteiger, *supra* note 133; Bruno S. Frey & Reto Jegen, *Motivation Crowding Theory*, 15 J. ECON. SURVEYS 589, 596–606 (2001) (surveying experimental evidence suggesting that introduction of monetary penalties and other financial incentives can decrease individuals' willingness to work together and behave reciprocally); Bruno S. Frey & Matthias Benz, Motivation Transfer Effect (June 12, 2001) (unpublished manuscript, *available at* http://www.iew.unizh.ch/home/benz/downloads/MTE.pdf) (finding that economic incentives reduce intrinsic motivation, including desire to cooperate and follow social norms, even with respect to conduct not subject to incentive mechanism). But recent experiments also show that, where there is some probability that the same buyers and sellers will continue transacting in the next period, legal enforcement that is limited to the verifiable dimensions of the agreement actually enhances cooperation in those dimensions of the agreement that are nonverifiable. *See* Sergio G. Lazzarini, Gary J. Miller & Todd R. Zenger, *Order with Some Law: Complementarity Versus Substitution of Formal and Informal Arrangements*, 20 J.L. ECON. & ORG. 261, 261 (2004) ("[B]y enforcing contractible exchange dimensions, contracts facilitate the self-enforcement of noncontractible dimensions.").

legally unenforceable components. The evidence suggests that an attempt to extend legal enforcement to difficult-to-verify contract terms—such as an understanding that an option will only be exercised if the counterparty is bid-shopping—is likely to impair the efficacy of those means of enforcement that rely on relational norms. As we noted in Part I above, formal common law contract doctrine has resisted this invitation to imply broad standards of fairness or equitable adjustment. Formal contract law uses rules of interpretation that afford the parties wide discretion to determine the scope of legal enforcement precisely. The availability of nonlegal means of enforcing commitments that courts cannot easily verify supports this approach. Therefore, the more general lesson for courts is that the equitable instinct to judicialize preferences for fairness and reciprocity may well destroy the effectiveness of these nonlegal mechanisms.[141]

Regardless of which scenario better conforms to the underlying facts, therefore, commercial parties intent on maximizing the expected gains from contracting would prefer that a court faced with the issues in *Hunt Foods* would rule inadmissible the evidence of an oral condition on an option. A court could reach this result either (a) by faithfully applying the "certain inclusion" test to find that such an agreement was fully integrated, thereby barring evidence of even a consistent additional term, or (b) if the agreement was only partially integrated, by ruling the evidence of the condition inadmissible because it was being offered to prove a term that was inconsistent with the express condition in the writing.[142] A court could even reconcile enforcement of the unrestricted written option with the belief that the parties' contractual ends were to grant only a conditional option. As long as the parties understood that the condition was to be enforced informally, then legal enforcement should not be available as a backstop. After all, self-enforcing mechanisms occasionally break down, and, in the event they do, parties who negotiated legally nonenforceable mechanisms should accept their failure and not seek the legal enforcement they eschewed in the original agreement. Though the mechanisms may break down, there is strong evidence that commercial parties prefer to retain the option to enforce their commit-

---

[141] It is important to keep in mind that legal disputes only arise when the informal modes of enforcement have broken down. Litigated cases, therefore, offer little evidence of the power of reciprocity, reputation, and other informal mechanisms in enforcing the agreements between commercial parties that never reach litigation.

[142] Alternatively, if the *Hunt Foods* court were to be interpreted as having treated the oral statement as evidence of a condition precedent to the agreement's taking effect, *see supra* note 121, the court should have held that the condition was not admissible because it contradicted the express option in the agreement.

ments through two sets of rules: an explicit (and formal) set of rules for those parts of their relationship that require legal enforcement and an implicit (and flexible) set of rules for those aspects that respond best to self-enforcement.[143] By correctly identifying the parties' contractual ends but failing to understand that the parties might reasonably have chosen to achieve those ends by using both legal and nonlegal means, the court in *Hunt Foods* undermined the ability of parties in the future to regulate their relationships with a combination of legal and nonlegal norms.

How representative is the decision in *Hunt Foods*? All the available evidence suggests that courts in the United States are sharply divided between jurisdictions that retain a "hard" parol evidence jurisprudence that hews closely to formal contract doctrine and those courts that have relaxed the formal doctrine in favor of a "soft" parol evidence rule that invites courts to vindicate contractual ends ex post.[144] Perhaps the clearest examples of this trend are the differing approaches to contractual interpretation employed by courts in New York and California. New York courts continue to adhere to formal doctrine in cases that fall outside the more liberal purview of the U.C.C.[145] Thus, New York courts continue to follow the traditional

---

143 *See* Edward A. Bernstein, *Law & Economics and the Structure of Value Adding Contracts: A Contract Lawyer's View of the Law & Economics Literature*, 74 OR. L. REV. 189, 227–28 (1995) (discussing tradeoffs between legal and relational means of enforcement); Bernstein, *Merchant Law*, *supra* note 130, at 1788–89 (same); David Charny, *Nonlegal Sanctions in Commercial Relationships*, 104 HARV. L. REV. 373, 391–426 (1990) (same); Stewart Macaulay, *Non-Contractual Relations in Business: A Preliminary Study*, 28 AM. SOC. REV. 55, 60–65 (1963) (discussing pervasive use of relational enforcement in sample of business contracts); Ronald J. Mann, *Verification Institutions in Financing Transactions*, 87 GEO. L.J. 2225, 2249–52 (1999) (discussing tradeoffs between legal and relational means of enforcement); John McMillan & Christopher Woodruff, *Private Order Under Dysfunctional Public Order*, 98 MICH. L. REV. 2421, 2458 (2000) (same).

144 *See, e.g.*, Eric A. Posner, *The Parol Evidence Rule, the Plain Meaning Rule, and the Principles of Contractual Interpretation*, 146 U. PA. L. REV. 533, 534 (1998) (distinguishing two polar positions: "hard-PER," where "the court generally excludes extrinsic evidence and relies entirely on the writing," and "soft-PER," where "the court gives weight both to the writing and to the extrinsic evidence"). *See generally* Robert A. Hillman, *The "New Conservatism" in Contract Law and the Process of Legal Change*, 40 B.C. L. REV. 879, 881–88 (1999) (identifying shift in contract law in favor of written contracts over "alleged oral, less formal representations or agreements" and predicting shift in judicial interpretation away from "new conservative" outlook toward more liberal approach to contract interpretation); Ralph James Mooney, *The New Conceptualism in Contract Law*, 74 OR. L. REV. 1131, 1132–35 (1995) (identifying difference between hard and soft interpretive regimes and arguing that hard regimes tend to benefit already powerful and wealthy class). For a discussion of approaches to choosing a hard or soft interpretive regime, see SCOTT & KRAUS, *supra* note 3, at 596–97.

145 *See, e.g.*, Morgan Stanley High Yield Sec., Inc. v. Seven Circle Gaming Corp., 269 F. Supp. 2d 206, 220 (S.D.N.Y. 2003) (denying defendant's claim of unsatisfied oral condition precedent and enforcing written contract when contract said "this Agreement is [the

"four corners of the contract" presumption by which the finality of a writing is determined in the first instance from an examination of the writing itself.[146] In California, by contrast, the courts have for many years invited collateral attacks on contractual writings if necessary to better determine the "true intent of the parties."[147]

Our claim that sophisticated parties prefer a "hard" parol evidence jurisprudence is supported by recent work by Theodore Eisenberg and Geoffrey Miller studying choice of law and choice of

---

signing party's] legal, valid and binding obligation enforceable against it in accordance with its terms" (alteration in original)); Intershoe, Inc. v. Bankers Trust Co., 571 N.E.2d 641, 642, 644 (N.Y. 1991) (excluding parol evidence where writing appeared to embody parties' final agreement unambiguously); Mitchill v. Lath, 160 N.E. 646, 647–48 (N.Y. 1928) (upholding four-corners presumption and excluding evidence of collateral agreement to land-sale contract).

146 Geoffrey P. Miller, *Bargaining on the Red-Eye: New Light on Contract Theory* 40 (N.Y. Univ. Law & Econ. Working Papers, Paper No. 131, 2008), *available at* http://lsr.nellco.org/nyu/lewp/papers/131; *see, e.g.*, *Morgan Stanley*, 269 F. Supp. 2d at 214 ("Under New York law a contract which appears complete on its face is an integrated agreement as a matter of law."). In addition, merger clauses are given virtually conclusive effect in New York. *See* Tempo Shain Corp. v. Bertek, Inc., 120 F.3d 16, 21 (2d Cir. 1997) ("Ordinarily, a merger clause provision indicates that the subject agreement is completely integrated, and parol evidence is precluded from altering or interpreting the agreement."); Jarecki v. Louie, 745 N.E.2d 1006, 1009 (N.Y. 2001) ("The purpose of a merger clause is to require the full application of the parol evidence rule in order to bar the introduction of extrinsic evidence . . . . The merger clause accomplishes this purpose by evincing the parties' intent that the agreement 'is to be considered a completely integrated writing.'"); Norman Bobrow & Co. v. Loft Realty Co., 577 N.Y.S.2d 36, 36 (App. Div. 1991) ("Parol evidence is not admissible to vary the terms of a written contract containing a merger clause.").

147 Masterson v. Sine, 436 P.2d 561, 564 (Cal. 1968) (admitting parol evidence to vary terms of deed on ground that "[e]vidence of oral collateral agreements should be excluded only when the fact finder is likely to be misled"); Pac. Gas & Elec. Co. v. G.W. Thomas Drayage & Rigging Co., 442 P.2d 641, 645 (Cal. 1968) ("[R]ational interpretation requires at least a preliminary consideration of all credible evidence offered to prove the intention of the parties."); *see also* Susan J. Martin-Davidson, *Yes, Judge Kozinski, There* Is *a Parol Evidence Rule in California—The Lessons of a Pyrrhic Victory*, 25 Sw. U. L. Rev. 1 (1995) (examining California cases raising issue of admissibility of extrinsic evidence and concluding that parol evidence jurisprudence represents one of most confused and incoherent areas of law in California); Miller, *supra* note 146, at 41–42 ("If . . . the extrinsic evidence reveals ambiguity, then the court may consider all such extrinsic evidence as may be relevant to interpreting the contract."). Martin-Davidson found the following after examining California cases:

> The cases in the study did not confirm the accusation that California has abandoned the parol evidence rule. On the other hand, neither did they suggest that the rule is alive and well. Instead, they supported a different accusation: that the parol evidence rule persists in California like a neomort maintained on permanent life-support as a ready source of transplant organs. The many standard but incompatible formulations of the parol evidence rule are ready as needed in this on-going and inconclusive battle. One faction is fighting in defense of the written word while its many enemies insist that a written agreement is not "all they wrote."

Martin-Davidson, *supra*, at 9.

forum clauses in a data set of 2865 contracts.[148] Their study shows that parties chose New York law in 46% of the contracts and New York as the forum state in 41% of the contracts.[149] California, on the other hand, was chosen for its contract law in less than 8% of the contracts even though its commercial activity, as measured by the place of business of the contracting parties, was second only to New York.[150] The Eisenberg and Miller study illustrates the strong preferences of commercial parties for the formal contract law of New York in lieu of the frequent exercise of equitable overrides by courts in California.

III
Rules and Standards as Means to Effect Contractual Ends

In the preceding Part, we examined the tradeoffs between legal and nonlegal enforcement that parties confront when designing their contracts. In this Part, we focus on the second design challenge that arises once parties have determined to enter into a legally enforceable agreement: How should the parties choose between precise, bright-line contract terms (or *rules*) and broad, vaguely defined terms (or *standards*)? We begin our analysis of the tradeoffs in the design of legally enforceable contracts by focusing on the contract dispute in a well-known case, *Aluminum Co. of America v. Essex Group, Inc.* (*ALCOA*).[151] *ALCOA* is a paradigmatic illustration of cases in which courts conclude that the parties would have intended judicial intervention under the post-formation circumstances that have materialized. While the reformation remedy adopted by the court in *ALCOA* has not been followed elsewhere (and, as a consequence, claims of excuse are only rarely granted by courts),[152] the case nevertheless provides a vivid illustration of how equitable reasoning can be used to vitiate formal doctrine.

---

148 Theodore Eisenberg & Geoffrey P. Miller, *The Flight to New York: An Empirical Study of Choice of Law and Choice of Forum Clauses in Publicly-Held Companies' Contracts* 4–5 (N.Y. Univ. Law & Econ. Working Papers, Paper No. 124, 2008), *available at* http://lsr.nellco.org/nyu/lewp/papers/124.

149 *Id.* at 19, 34. Delaware was a distant second to New York, with about 15% of the parties choosing its law. No other state accounted for even 10% of the choices of law. *Id.* at 19.

150 *Id.* at 19 tbl.2, 23 tbl.5.

151 499 F. Supp. 53 (W.D. Pa. 1980).

152 *See* Alan Schwartz, *Relational Contracts in the Courts: An Analysis of Incomplete Agreements and Judicial Strategies*, 21 J. Legal Stud. 271, 271–72 (1992) (describing passivity of courts in contract-excuse cases). For a case rejecting *ALCOA*'s reasoning, see *Atlas Corp. v. United States*, 895 F.2d 745, 750 (Fed. Cir. 1990).

It is tempting to object that *ALCOA*'s application of the mistake doctrine is the exception that proves the rule: Because few courts have followed *ALCOA*'s reasoning, the vast majority of long-term contracts between sophisticated commercial parties are unlikely to be set aside or reformed by application of the mistake doctrine.[153] Despite their infrequency, we believe that cases like *ALCOA* cast a significant shadow of uncertainty across all contracts. More importantly, however, we believe that the impact of equitable doctrines such as the mistake doctrine is significant because of the far more numerous cases in which equitable reasoning plays a significant role in shaping the outcome. Thus, we use *ALCOA* to illustrate the *kind* of reasoning that underpins all equitable doctrines. That reasoning, explicit in *ALCOA*, invites courts to vindicate the parties' intended contractual ends, either directly by using equitable doctrines to override formal doctrine or indirectly by applying formal doctrine in a nonstandard fashion. When a court concludes that the ex post circumstances are historically exceptional, it reasons that the parties would have wanted the court to intervene to realign their contract terms with their contractual ends. Contract design theory shows, however, that this inference is not justified.

A. Contractual Intent and the Law of Mistake and Excuse

1. The ALCOA-Essex Contract

ALCOA and Essex entered into a long-term tolling contract whereby ALCOA undertook to convert alumina supplied by Essex into aluminum.[154] The agreement was to last for sixteen years, with an option by Essex to renew for five additional years, and specified a fixed quantity of seventy-five million pounds of aluminum to be delivered to Essex per year. Along with its obligation to take seventy-five million pounds per year, Essex had the concomitant obligation to deliver sufficient alumina for ALCOA to convert into aluminum, which Essex could then dispose of as it pleased.

The contract contained a detailed price indexing provision. The initial contract price was 15 cents per pound, to be adjusted according to a complex formula: 5 cents of the price was designated as a "demand charge." This reflected ALCOA's capital costs for the smelting capacity at its Warrick, Indiana plant. In addition, Essex agreed to pay a 10 cent production charge, of which 4 cents were

---

[153] See *infra* text accompanying notes 160–71 for a discussion of the doctrine of mistake and the reasoning of the court in *ALCOA*.

[154] The contract between ALCOA and Essex is analyzed in detail in VICTOR GOLDBERG, FRAMING CONTRACT LAW: AN ECONOMIC PERSPECTIVE 348–69 (2006).

fixed. Of the remaining variable costs, 3 cents corresponded to the labor cost and 3 cents reflected the nonlabor production cost. The former was indexed by a labor index; the latter was indexed by the Wholesale Price Index, Industrial Commodities (WPI-IC). At Essex's insistence, the parties agreed that the contract price would be capped at 65% of the price of a standard grade of aluminum ingot as published in American Metal Market, a trade publication. Although ALCOA did not choose to protect itself similarly against the possibility that the WPI-IC would fail to track costs accurately, the contract did provide for the risk that the indexes would no longer be available.[155] In short, at Essex's insistence, the index related to ALCOA's costs included a circuit breaker if the index rose too fast relative to the underlying market price, but ALCOA did not require a corresponding "booster" if the index moved too slowly. Thus, the index had a ceiling but not a floor.

The contractual purposes for the index were clear from ALCOA's perspective: The 5 cent demand charge represented 4 cents for financial costs (depreciation, interest, and so on) and 1 cent "guaranteed profit." The fixed portion of the production charge included 3 cents profit and 1 cent to cover general administrative and selling expenses. Of the variable costs, 3 cents represented labor costs and 3 cents reflected material costs, about 75% of which was for power. As Victor Goldberg suggests, the contract was thus designed to "replicate" a situation wherein "Essex invested in a smelter with a capital cost of 40 cents per pound with a 10% (4 cent) return. The demand charge would be paid by Essex regardless of whether it took any aluminum, so in that sense it was taking the risk of ownership."[156]

2. *The Litigation in* ALCOA

Unfortunately for ALCOA, the index moved too slowly relative to the actual market, owing in part to the underrepresentation of energy costs in the basket of inputs that comprise the WPI-IC relative to the costs of converting alumina into aluminum. Three things happened that caused the contract price to be out of line with both ALCOA's costs and the price of aluminum. The OPEC oil embargo

---

[155] Clause 28 of the ALCOA-Essex contract provides:

> In the event that any index referred to in paragraph 8 and 9 hereof is discontinued, no longer published in the sources indicated or become [sic] unavailable, the parties shall agree on a comparable substitute index and if they are unable to agree [on] the selection of a substitute index, the selection of a comparable substitute index shall be submitted to arbitration . . . .

Id. at 354.

[156] *Id.* at 353.

caused coal prices to rise far more rapidly than the WPI-IC, so the nonlabor variable production cost component failed to track the actual costs: ALCOA's nonlabor production costs rose from 5.8 cents in 1973 to 22.7 cents in 1978, while the WPI-IC rose by less than two-fold of the original.[157] In addition, there were extraordinary rates of inflation in the 1970s, such that the nominal price for aluminum rose significantly. Finally, there was an increased demand for aluminum and thus the underlying or "real" price of aluminum rose as well.[158]

By 1979, the per-pound market price of aluminum ingot was around 73 cents, while ALCOA's costs were around 35 cents and the indexed contract price was 25 cents. Since Essex's costs for the alumina and for transportation to the Warrick smelter were about 11 cents, it could capture a profit of 37 cents per pound if it chose to resell the ingot on the market. And in fact, ALCOA claimed that Essex had exploited this favorable structure and had resold almost 23% of the aluminum on the open market.[159]

ALCOA sought to renegotiate the contract, but Essex refused. ALCOA then brought suit asking for equitable reformation of the contract. The crucial issue for the court was the rapid increase in the cost of power at the Warrick smelter and the failure of the price index to track those costs. The court granted ALCOA relief on the grounds of mutual mistake, commercial impracticability, and frustration of purpose.[160] Rather than excuse ALCOA, however, the court chose to reform the contract by rewriting the price term. The new contract price would be the lesser of the 65% cap and either the price as defined in the contract or ALCOA's costs plus 1 cent, whichever was greater.[161]

In its mutual mistake analysis, the court concluded that there was a mistake as to a fact—the belief that the index would work as the parties expected. It also found that this mistake was mutual because

---

157 Aluminum Co. of Am. v. Essex Group, Inc. (*ALCOA*), 499 F. Supp. 53, 59 (W.D. Pa. 1980).

158 The price term also failed to properly index the fixed-demand charge representing ALCOA's capital investment in the smelter. GOLDBERG, *supra* note 154, at 356.

159 *ALCOA*, 499 F. Supp. at 59; GOLDBERG, *supra* note 154, at 355.

160 *ALCOA*, 499 F. Supp. at 60–78. The court based its analysis of the three doctrines on the relevant provisions in the Uniform Commercial Code and the Restatement (Second) of Contracts. *See, e.g.*, U.C.C. § 2-615 (amended 2003) (containing current version of excuse by failure of presupposed conditions doctrine similar to that relied on in *ALCOA*); RESTATEMENT (SECOND) OF CONTRACTS §§ 151–154 (1981) (mistake); *id.* § 261 (discharge by supervening impracticability); *id.* § 265 (discharge by supervening frustration).

161 Ultimately, the new contract price "would have meant that in 1979 ALCOA would have received about 11 cents per pound more than it would have absent the modification." GOLDBERG, *supra* note 154, at 357.

mutuality is a question of understanding, not motivation, and thus rejected Essex's argument that the mistake was not mutual because Essex was not concerned about keeping ALCOA's costs in check. The court emphasized the strict test required to establish excuse on the basis of mutual mistake: The mistake doctrine applies to a mistaken belief about a particular fact only if, as Corbin asserted, a court decides "the parties made a definite assumption that it existed and made their agreement in the belief that *there was no risk with respect to it*."[162] According to the court, the alleged mutual mistake in *ALCOA* was the parties' belief in "the suitability of the WPI-IC as an index to accomplish the purposes of the parties."[163] The court found that "each [party] assumed the Index was adequate to fulfill its purpose. This mistaken assumption was essentially a present actuarial error."[164]

Because the court found the doctrine of mutual mistake applicable to the case, it would seem to follow that it believed the parties "made their agreement in the belief that there was no risk"[165] that the price index would fail to reflect ALCOA's actual costs. This is precisely the claim Essex denied.[166] The court's response was to demonstrate that the parties "plainly sought to limit the risks of their undertaking."[167] Chief among the pieces of evidence cited to demonstrate the parties' efforts to limit risk under their agreement was the fact that ALCOA hired Dr. Alan Greenspan to advise it on the drafting of the objective price index.[168] As to the question of why

---

[162] *ALCOA*, 499 F. Supp. at 60 (emphasis added) (quoting 3 ARTHUR LINTON CORBIN, CORBIN ON CONTRACTS § 605, at 643 (1960)).

[163] *Id.* at 61.

[164] *Id.* at 63. The requirement that the mistaken belief must concern a fact that exists at the time of the agreement serves to distinguish the law of mistake from the law of excuse. While the mistake doctrine voids agreements based on mistaken assumptions about matters of fact that exist at the time of formation, it is designed not to apply to cases in which parties make mistaken predictions about the future, such as erroneous assumptions about future market conditions or their future financial situations. RESTATEMENT (SECOND) OF CONTRACTS § 152 cmt. b (1981). The *ALCOA* court therefore could only use the mistake doctrine as a ground for reformation by characterizing the parties' mistake about the future functioning of the price index as a mistake concerning a fact in existence at the time of the contract's formation.

[165] *ALCOA*, 499 F. Supp. at 60.

[166] *Id.* at 68. As the court noted, "Essex first asserts that ALCOA expressly or implicitly assumed the risk that the WPI-IC would not track ALCOA's non-labor production costs. Essex asserts that ALCOA drafted the index provision . . . and that ALCOA's officials knew of the inherent risk that the index would not reflect cost changes." *Id.*

[167] *Id.*

[168] Alan Greenspan's actual role in developing the price term is a matter of some dispute. In its brief to the Third Circuit, Essex claimed that "[c]ontrary to the trial court's finding, George [Alcoa's key negotiator] testified that Alan Greenspan . . . was not consulted by Alcoa in connection with [the] Contract." GOLDBERG, *supra* note 154, at 361

ALCOA did not include a price floor, whereas Essex included a price ceiling, the court responded: "[T]he absence of an express floor limitation can only be understood to imply that the parties deemed the risk too remote and their meaning too clear to trifle with additional negotiation and drafting."[169] The court argued that:

> The proper question is not simply whether the parties to a contract were conscious of uncertainty with respect to a vital fact, but whether they believed that uncertainty was effectively limited within a designated range so that they would deem outcomes beyond that range to be highly unlikely. . . . Both consciously undertook a closely calculated risk rather than a limitless one.[170]

Thus, the court moved from Corbin's requirement that the parties believe there is *no risk* that they are wrong about their shared factual belief, to a more liberal requirement that the parties believe it is *highly unlikely* that their shared factual belief is wrong.

But when the parties are conscious of a particular risk and yet form a belief that the risk has been reduced or eliminated, it is no longer plausible to argue that it did not occur to them that their belief might be wrong. In such a case, it is far more difficult to conclude that the parties did not intend to allocate this risk. Nevertheless, the court argued that the parties' demonstrable efforts to reduce the risk showed that they intended not to allocate the residual risk. In its view, the parties considered that risk so remote that it was not worth allocating. This explains and justifies ALCOA's failure to bother with drafting a floor to protect itself against the remote risk that the contract's price index would not function as ALCOA anticipated. Thus, the court concluded, ALCOA did not assume the risk through any of the four ways by which a party can do so.[171]

(quoting Brief of Appellant at 16 n.9, Aluminum Co. of Am. v. Essex Group, Inc., No. 80-1604 (3d Cir. 1980)).

169 *ALCOA*, 499 F. Supp. at 69.

170 *Id.* at 70.

171 *Id.* at 67–70 (explaining that ALCOA did not assume risk expressly, through common understanding or trade usage, according to general policies, or through conscious ignorance). *See generally* RESTATEMENT (SECOND) OF CONTRACTS § 154 (1981) (stating that party assumes risk of mistake by agreement of parties, conscious treatment of limited knowledge with respect to mistake-related facts as sufficient, or reasonable allocation by court). In applying this provision to the facts of the case, the court first rejected the argument that, by not including a floor in the price term, ALCOA expressly or implicitly agreed to bear the risk that the escalator might rise too slowly. 499 F. Supp. at 68. The court found that ALCOA considered the possibility that a floor might be necessary too remote to put into the contract. The court also rejected Essex's argument that the contract should be interpreted against its drafter, ALCOA, because according to the court, that interpretive maxim is only appropriate when there is ambiguity or a policy concern, neither of which were present in the case. *Id*. at 69. As to whether ALCOA assumed the risk by proceeding in the face of conscious ignorance of the risk, the court said that the test is not

The court's conclusion that the risk was not allocated to ALCOA in the contract is certainly possible, but it is highly improbable. Even if both parties—perhaps because of Greenspan's role in drafting the index—believed the probability that the price index would malfunction was extremely low, it does not follow that ALCOA, let alone Essex, understood the contract to be subject to the condition that the index function as the parties anticipated it would. Indeed, if both parties believed the contract was, in effect, conditional on the price index functioning to limit the range of price variation to no more than 3 cents per pound, it is difficult to explain why Essex insisted on a price ceiling on the index. Insisting on a price ceiling reflects the realization that the index would not work as anticipated. Since Essex did include a price ceiling, it is difficult to explain how ALCOA could reasonably conclude that Essex believed the contract was conditional on the price index working as predicted.[172]

3. The Ex Post Justification: Honoring the Parties' Intent

The court's focus on contractual intention is critical to its analysis of the doctrinal grounds it used to justify reformation. The price index, the court reasoned, was designed to reduce, if not eliminate, the deviation between contract price and market price. Thus, this contract was unlike a standard fixed price contract where the risk of price increases is allocated to the seller and the risk of price declines is assigned to the buyer. Here the objective of the parties was to "avoid the full risk of future economic changes" by using the complex price indexing mechanism.[173] In this way, the intent was to limit each partner's risk in a particular way: to provide Essex "an objective pricing formula" and to give ALCOA "a formula which would cover

---

whether they were aware at some level that uncertainty existed, but rather whether they believed the uncertainty was limited (which ALCOA did). *Id.* at 69, 70. Finally, the court noted that there was no reason to allocate the risk to ALCOA, nor did customary dealing call for ALCOA to assume the risk. *Id.* at 67.

[172] The court also considered impracticability and frustration, which have the same basic doctrinal requirements as mistake, but focus on the hardship imposed on the plaintiff. *ALCOA*, 499 F. Supp. at 70–71 ("In broad outline the doctrines of impracticability and of frustration of purpose resemble the doctrine of mistake."). The court found for ALCOA on both of these grounds based on its $60 million expected loss resulting from the failure of the index. *Id.* at 70–78. The court applied its findings on the mistake issue to the impracticability issue, finding that the nonoccurrence of the event (and the nonexistence of the fact) that caused the impracticability was a basic assumption on which the contract was made. The difficulty envisioned in the impracticability doctrine, according to the court, must be extreme and unreasonable: In this case, the loss was so significant in absolute size and proportion to the value of the contract that it altered the essential nature of performance. *Id.* at 72–74.

[173] *Id.* at 63.

its out-of-pocket costs over the years and yield a return of around four cents a pound."[174] As a consequence, the failure of the price index to achieve the ends sought by this contract justified invocation of each of the three doctrines the court used to relieve ALCOA from full liability under the contract. In short, under the Restatement, the failure of the price index to achieve the parties' ends was (a) "a mistake of both parties . . . as to a basic assumption" of the contract justifying rescission on the grounds of mutual mistake,[175] (b) the occurrence of an event whose nonoccurrence was a basic assumption of the contract justifying an excuse on the grounds of commercial impracticability,[176] and (c) the occurrence of an event whose nonoccurrence was a basic assumption that frustrated ALCOA's principal purposes under the contract.[177]

Viewed ex post, the decision in *ALCOA* may seem justifiable. Indeed, Victor Goldberg, who sharply criticizes the transactional lawyers who wrote the ALCOA-Essex contract and the litigators who argued the case, nonetheless concludes that "the judge imposed a cost-plus gloss on the contract which was not explicit in the contract but most likely comported with the parties' intentions."[178] But this understanding of the parties' intentions mistakenly equates contractual intent with the parties' intended contractual ends. While the court may well have correctly recognized the parties' contractual ends, it failed to take account of the contractual mechanisms that serve as essential means of achieving those ends. This point can best be appreciated by turning the focus of analysis away from the ex post litigation perspective assumed by the court to the ex ante perspective of the parties charged with designing the contract. From the vantage point of contract design, the *ALCOA* decision takes on a much different cast.

B. A Contract Design Explanation of the ALCOA-Essex Contract

1. Reframing the Question of Intent

In a case like *ALCOA*, where both parties were clearly aware of the risk that the price index might not achieve ALCOA's profit objective, the fundamental question is whether the parties conditioned the contract on the risk not materializing or whether they allocated the risk between them. The court found that the parties' significant

---

[174] *Id.*
[175] RESTATEMENT (SECOND) OF CONTRACTS § 152 (1981).
[176] *Id.* § 261.
[177] *Id.* § 265.
[178] GOLDBERG, *supra* note 154, at 369.

efforts to reduce the risk were evidence that they conditioned their agreement on the risk not materializing. The court relied on these efforts to demonstrate that the parties were motivated to reduce the risk that the contract price and the market price of aluminum might deviate significantly in the future, thus imposing large losses on one party or the other. But the issue is not whether the parties intended to reduce this risk; clearly they did. That they intended to devise an indexing methodology that reduced the risk of deviation between market price and contract price does not, however, provide any evidence of their intent to condition their agreement on the residual risk that the index might malfunction.

To see this point, it is helpful to return to the distinction we have drawn between two kinds of contractual intent: the parties' intended contractual means and their intended contractual ends. The evidence reflects that the parties intended to create a pricing mechanism that tracked ALCOA's costs (and thus implicitly tracked the market price for the smelting services subject to a built-in price discount). One alternative would have been to agree to a cost-plus contract, one in which ALCOA could recover its costs of performance plus an agreed profit. Cost-plus contracts have many problems, however. They reduce the seller's incentives to economize on costs, they are hard to monitor, and they require revelation and verification of confidential information.[179] So the parties may have been motivated instead to create a verifiable proxy for a cost-plus contract that avoided these difficulties. From that perspective, the agreed-upon price index was intended to allow ALCOA to recover its capital costs in the Warrick plant plus a return on its investment and to permit Essex to obtain a favorable price for the smelting services over the life of the contract. These were the parties' intended contractual ends. But the question before the court concerned the parties' intended contractual means: What instruments did the parties select in order to achieve their goal of having the contract price track the market price of smelting services?

2. The Choice of Contractual Means: Rules Versus Standards

Parties shift costs between the front and back end of the contracting process depending on exogenous factors such as the degree of uncertainty and the relative costs of writing the contract. They do this by changing the character of the terms in the contract.[180] To see how,

---

[179] *See id.* at 363 (describing problems in designing cost-based contracts).

[180] The following analysis of the tradeoff between rules and standards in contract design draws on the discussion in Scott & Triantis, *supra* note 10, at 822–56 (discussing first, in

recall that courts do not directly observe the effects of a contingency—such as the oil shock in *ALCOA*—on which a contract is implicitly conditioned. Rather, they rely on evidence or *proxies*. Consider a simple example of how these evidentiary proxies work. Units of evidence (e.g., the oil shock's causing carbon-based energy costs to increase by 40%) prove directly whether a precise contract term (or *rule*) is satisfied (e.g., "the contract price will be adjusted by x% if ALCOA's energy costs increase by more than 35%"). Now consider a vague contract term (or *standard*), such as "the contract price is subject to adjustment if there are extraordinary increases in production costs caused by extrinsic factors beyond ALCOA's control." Vague terms are one step further removed from the evidentiary units: The evidence may establish that the oil shock caused the cost of coal power to increase by 40%, but then the court must determine how much importance to give to that evidence compared to alternative factors (e.g., labor disputes, management's decision to shift from hydroelectric power to coal-based power, etc.).[181]

A precise rule is "noisy." It is inevitably under- and over-inclusive relative to the underlying contractual purpose it is designed to serve.[182] But it also restricts the court's discretion more severely and thus reduces ex post enforcement costs. With a precise rule, the parties are increasing their investment in front-end transaction costs, but by reducing the role of the court in litigation, back-end enforcement costs will decline. At the same time, the parties are exploiting their informational advantage (they know their contractual ends and have the right incentives to choose the best means to achieve them), but they are sacrificing the hindsight advantage that a court might have. A vague standard, on the other hand, has less noise because it can be custom-tailored ex post to promote the parties' contractual ends. Since the parties cannot foresee all contingencies, they can use vague contract terms to delegate to a court the task of completing the contract ex post. But vague terms increase enforcement costs because the parties will be able to introduce further evidence to support their com-

---

Part I, differences in costs between front and back ends of contracting, and second, in Part II, how parties decide between precise and vague contract terms).

[181] According to Goldberg,

> Warrick was the only Alcoa smelter that relied on coal for its electric power. The remainder used hydroelectric power, power provided by [the Tennessee Valley Authority], or natural gas. The costs of all of Alcoa's other smelters did not rise with the price of oil. So, while Warrick was Alcoa's low cost plant when it was built, the changing fuel prices made it far and away the highest cost plant, post-1973.

GOLDBERG, *supra* note 154, at 356.

[182] SCOTT & KRAUS, *supra* note 3, at 390.

peting positions on whether the standard is satisfied. It follows that parties will choose a precise term when they believe that tailoring to their private information is more important than hedging against the possible negative effects of future contingencies on contract performance. Alternatively, when contractual obligations are highly contingent on the realized state of the world and less dependent on the parties' private information, parties will be more inclined to use vague standards.

There are thus two basic approaches to designing a price term in long-term contracts such as the ALCOA-Essex deal. The parties could specify the contract price in terms of a broad standard: "The price shall be a reasonable price that guarantees ALCOA a profit of not less than $.01 nor more than $.07 per pound." In the extreme, the parties could omit any express price term at all, thereby delegating to a court the task of imputing a reasonable price at the time and place of delivery.[183] By choosing a standard, the parties shift contracting costs from the front end to the back end. To optimize the total costs of contracting, however, the parties must compare the back-end costs (and benefits) of standards to the front-end costs (and benefits) of a precise and conclusive price term. This alternative strategy requires parties to incur ex ante the costs of specifying a rule: a precise pricing mechanism that achieves their contractual ends subject to an acceptable risk of deviation from those ends.[184] Here, the parties economize on back-end enforcement costs and also exploit their informational advantage in understanding their contractual ends. But unlike the standards approach, when the parties reduce the price term to a rule, they anticipate no role for a court ex post in selecting a proxy in the event the precise term does not function as predicted.

In short, parties face the familiar tradeoff of standards versus rules. Under either approach, the parties might fail fully to achieve their contractual ends. Courts might misinterpret a standard and select a poor proxy, thereby undermining the parties' ends. Alterna-

---

[183] *See, e.g.*, U.C.C. § 2-305(1) (amended 2003) ("The parties if they so intend may conclude a contract for sale even if the price is not settled. In such a case the price is a reasonable price at the time for delivery if: (a) nothing is said as to the price; (b) the price is left to be agreed by the parties and they fail to agree; or (c) the price is to be fixed in terms of some agreed market or other standard as set or recorded by a third person or agency and it is not so set or recorded.").

[184] The advantage of contractual rules is that the parties themselves know better than the courts what their objectives are and presumably can specify a proxy that best enhances contractual incentives. Moreover, by making proof of the price term trivial, a rule reduces the expected costs of adjudication. The chief disadvantage of a rule is that it might not achieve the objective the parties intended. Like all rules, a precise price term, no matter how sophisticated, is likely to be over- or under-inclusive over a large range of possible future circumstances.

tively, a price term framed as a rule might fail to function as predicted once the future state of the world is realized. The tradeoff, as noted above, is between the parties' informational advantage in knowing their contractual ends and the court's informational advantage in knowing what future states have materialized. Contract design theory argues that the parties themselves are best able to fully evaluate this tradeoff and select the particular contractual means that maximizes the expected value of their contract.

3. Determining the Chosen Contractual Means

Returning to *ALCOA*, we can now focus on the central question in the case: What were the parties' intended contractual means for achieving their objective? The key point is that *both the rules-based and the standards-based approaches are consistent with the parties' intended contractual ends of having the contract price track ALCOA's costs.* The question is which strategy the parties adopted as the means of achieving those ends. In *ALCOA*, the evidence strongly suggests that ALCOA and Essex both understood that they were using the rule-based approach. The parties incurred substantial front-end drafting costs. These costs are rational to incur only if they exploit the parties' informational advantage and/or create offsetting savings in expected enforcement costs (including the strategic consequences attending adjudication). If the parties preferred an ex post determination of the contract price, they could have forgone the substantial investment in drafting a price index designed to replicate a cost-plus contract.[185] If the goal was to exploit a court's hindsight advantage, the parties could have drafted a vague standard that gave the court discretion to determine the appropriate proxy. Instead, ALCOA's drafting strategy seemed plainly designed to achieve the benefits, and incur the costs, of a rule-based approach.

By choosing a precisely defined price index, sophisticated parties implicitly allocate to the seller the risk that the index might malfunction and increase too slowly, and allocate to the buyer the risk that the

---

185 The contract does not explicitly state that it was intended to mimic a cost-plus contract while avoiding familiar difficulties with cost-plus agreements; "[n]onetheless, the structure of the agreement makes it quite clear that the intent was to make the contract cost-based." GOLDBERG, *supra* note 154, at 363. To be sure, the agreed-upon indexing formula in the contract had many other problems, in addition to the failure of the WPI-IC to track carbon-based fuel costs. *See id.* at 363–65. But the mere fact that commercial parties design an ex ante rule badly does not justify judicial action to rewrite a precise contract term ex post. Indeed, one of the major reasons for inserting caps and floors in such indices is to guard against other risks, in addition to the risk of low-probability/high-impact states of the world. Among those risks is the risk of formulation error by the parties themselves.

index might malfunction and increase too rapidly. The court claimed to be maintaining fidelity to the parties' contractual intent by reforming the price index when it did not function as the parties anticipated it would. Ironically, however, the court's decision undermines commercial parties' ability to design their contracts optimally. Even if the court correctly interpreted the parties' intended ends in this case, the prospect of judicial intervention under these circumstances nonetheless impairs the ability of future commercial parties to choose the contractual means that best achieve their contractual ends.

But does this analysis imply that a claim of mutual mistake or excuse *necessarily* should be rejected whenever parties choose a contractual rule rather than a standard? To answer this question, let us contrast *ALCOA* with *Eastern Air Lines v. Gulf Oil Corp.*,[186] another case involving a malfunctioning price index. Here the parties entered into a long term contract for the sale of jet fuel at designated locations.[187] Their objective was to set a price for the jet fuel in such a way as "to allocate the risk of exogenous changes in the input price of crude oil to Eastern Air Lines and the risk of fluctuations in production costs to Gulf."[188] They settled on a precise term that adjusted the contract price according to an easily verifiable indicator of crude oil price—West Texas Sour crude "as listed . . . in Platts Oilgram Service."[189] Subsequently, as a result of governmental deregulation following the oil crisis in the 1970s, Platts Oilgram seriously understated the current market price of crude oil; the court declined to grant Gulf relief on the grounds of excuse,[190] perhaps because the parties could have anticipated the failure of the contract indicator by stating in the contract that the price would be tied to that of Platts Oilgram or "any other appropriate index."

The question for a court in a case such as *Eastern Air Lines*, therefore, is whether to select a substitute indicator when the parties fail to delegate expressly to the court the choice of a replacement if the designated index should fail. Clearly, this specific risk was too remote to occur to the parties: Neither party considered the possibility that oil would be deregulated *and* that Platts Oilgram would not report that unregulated price. Because the price index was neither complex nor expensive to create, there is less reason than in *ALCOA* to believe the parties' intended contractual means for replicating input prices was a conclusive price term. On the other hand, that fact also

---

186 415 F. Supp. 429 (S.D. Fla. 1975).
187 *Id.* at 432.
188 Scott & Triantis, *supra* note 10, at 843.
189 *Eastern Air Lines*, 415 F. Supp. at 433 (quoting contract).
190 *Id.* at 440.

suggests that it would have been relatively simple for the parties to have specified a replacement index.[191]

Eastern Air Lines therefore presents a test case for the general question of whether courts should ever intervene to modify clear, express contract terms in light of changed circumstances in the absence of an express request for ex post intervention in the contract. When courts do so, they use the doctrines of excuse, mistake, frustration, or commercial impracticability either to void the contract and return the parties to their precontract positions or to change the contract's terms to better align them with the parties' contractual ends given the changed circumstances. The *ALCOA* decision presses these doctrines beyond their traditional domain by using them to reallocate a loss resulting from a risk that the parties quite clearly contemplated.[192] Nevertheless, at bottom, the decision in *ALCOA* implicitly assumes that the actual events that transpired in that case were no different in kind from the failure of Platts Oilgram to report the deregulated price of crude oil in *Eastern Air Lines*. In both cases, the parties never contemplated the highly unlikely events that materialized and thus neither contract can properly be interpreted as having allocated the risks of these events. In short, the *ALCOA* court's analysis of the excuse doctrines is premised on a conceptual argument: Contractual liability is confined to losses caused by events whose risks were allocated by contract, but contracts cannot allocate risks that the parties did not consider. Therefore, losses caused by events that the parties did not consider cannot have been allocated by contract.

The problem with this argument is that it contains a false, suppressed premise—namely, that some risks are not contemplated by contracting parties. At some level of generality, *all* parties consider *every* risk because they assume the general risk that something might go wrong. Any specific risk the parties did not consciously consider (e.g., the risk of an embargo combined with inflation and increased demand) can be recast as an instance of a more general risk type they certainly did consider (e.g., the risk that economic and political factors might negatively affect the performance of their contract).[193] Simi-

---

[191] Note that the parties in *ALCOA* included such a clause in their contract. *See supra* note 155.

[192] By their own terms, these doctrines apply only to losses resulting from a "basic assumption" the parties shared that turns out to be wrong because of "circumstances not within the contemplation of the parties at the time of contracting." U.C.C. § 2-615 cmt. 1 (amended 2003). Given that ALCOA and Essex negotiated over and agreed to an express price ceiling, their belief that the price index would not malfunction cannot qualify as a basic assumption under the mistake or excuse doctrines.

[193] *See generally* George G. Triantis, *Contractual Allocations of Unknown Risks: A Critique of the Doctrine of Commercial Impracticability*, 42 U. TORONTO L.J. 450 (1992) (chal-

larly, any specific risk that the parties did consider (e.g., that events in the Middle East might adversely affect energy costs) can always be recast under a more specific definition that the parties did not consider (e.g., that Saudi Arabia would lead an OPEC embargo that would combine with inflation caused by a decrease in aggregate supply in the economy and with an increase in the world-wide demand for aluminum). Given that any event description admits of infinite levels of specificity or generality, the excuse doctrines require some nonarbitrary and operational method of settling on the proper level of description of events that should be used to determine the doctrines' applicability in any given case. No such method exists in principle, let alone in practice.

In consequence, the excuse doctrines in contract law lack a satisfactory explanatory or normative foundation. Especially given the availability of express force majeure clauses (which set out detailed, prescribed grounds for excusing contract performance), it is hard to defend a set of interpretive default rules that license or require courts to hold that promises in contracts between commercially sophisticated parties are always implicitly conditional on the nonoccurrence of events not contemplated (under some level of description) by the parties. Yet this is precisely what the excuse doctrines do. In contrast, contract design theory argues that unless express contract terms provide otherwise, courts should resist the temptation to use the excuse doctrines to imply conditions into contracts between commercially sophisticated parties, even if the events that occur were not, under a description at some salient level of specificity, contemplated by the parties.

Our claim that commercial parties have strong reasons, at least in theory, to prefer courts not to apply doctrines of excuse or mistake unless specifically requested to do so is also supported by the available empirical evidence. Geoffrey Miller's recent study comparing

---

lenging assumption that contracting parties are unable to rationally manage and allocate risks of unanticipated events). Triantis argues:

> While an unknown risk cannot be priced and allocated specifically, it can be priced and allocated as part of the package of a more broadly framed risk. For example, consider a party who agrees to transport a shipment of goods for a fixed fee. The risk of a nuclear accident in the Middle East that causes a dramatic decrease in the production of oil and a consequent increase in its price might not be foreseen. As a result, this risk cannot be allocated explicitly in the contract. However, the broader risk of a large increase in the price of oil for any reason can be. Therefore, there is no gap to be filled by the doctrine of impracticability: [T]he risk of nuclear accident, though unforeseen, is allocated implicitly. Instead, the doctrine alters the contractual allocation of the risk and its proponents must advance a rationale for the reallocation.

Id. at 452.

California and New York contract doctrines seeks to explain why commercial parties exhibit such a strong preference for New York contract law over that of other jurisdictions, especially California.[194] Miller's analysis shows that New York is less willing to relieve parties of their commitments on the basis of mistake or excuse.[195] "New York law allows rescission or reformation on the basis of mutual mistake of law only in limited circumstances," which is "[c]onsistent with its general philosophy of holding parties to their bargains."[196] California, on the other hand, is more receptive to pleas for relief from mistake and excuse: "Instead of emphasizing the importance of respecting the parties' written agreement, California courts focus on the requirement of mutual consent."[197] Moreover, because of California's liberal rules for admitting extrinsic evidence, either party has ample opportunity to invite the California courts to privilege contractual ends over their chosen means.[198] Though courts may believe

---

194 Miller, *supra* note 146; *see also* Eisenberg & Miller, *supra* note 148, at 21, 27 (providing empirical evidence showing parties choose New York law in about 46% of contracts studied while choosing California law in less than 8% of contracts even though parties choose California as place of business more frequently than New York).

195 Miller, *supra* note 146, at 37–39.

196 *Id.* at 37; *see also* Nash v. Kornblum, 186 N.E.2d 551, 553 (N.Y. 1962) (finding that mistake must be shown by clear and convincing evidence); Lacoparra v. Bellino, 745 N.Y.S.2d 693, 694 (App. Div. 2002) (same); Jossel v. Meyers, 629 N.Y.S.2d 9, 10–11 (App. Div. 1995) ("Where the parties have made an instrument as they intended it should be, and the instrument expresses the transaction as it was understood and designed to be made, then the party who had an opportunity to know the contents of the instrument cannot obtain cancellation or reformation because he misunderstood the legal effect . . . .").

197 Miller, *supra* note 146, at 38; *see also* CAL. CIV. CODE §§ 1550, 1565, 1580 (West 1982) ("Consent is not mutual, unless the parties agree upon the same thing in the same sense."); Weddington Prods., Inc. v. Flick, 71 Cal. Rptr. 2d 265, 277 (Ct. App. 1998) (noting importance of "parties' outward manifestations" of mutual consent). Miller writes:

> The necessity that the parties agree to the same thing at the same time makes it relatively easy for parties to frame claims for relief under this theory. California courts do not emphasize the restrictions that limit the availability of mutual mistake under New York law, such as the lack of fault on the part of the party seeking relief or the need for clear and convincing evidence to establish the claim.

Miller, *supra* note 146, at 38–39.

198 *See* CAL. CIV. PROC. CODE, § 1856(e), (g) (West 2007) (setting out exceptions to rule excluding parole evidence); Casa Herrera, Inc. v. Beydoun, 83 P.3d 497, 502 (Cal. 2004) (finding that parol evidence rule "does not . . . prohibit the introduction of extrinsic evidence 'to explain the meaning of a written contract . . . [if] the meaning urged is one to which the written contract terms are reasonably susceptible'" (quoting BMW of N. Am., Inc. v. New Motor Vehicle Bd., 209 Cal. Rptr. 50, 57 n.4 (Ct. App. 1984))). Note also that "[r]eformation is liberally available in California if mistake is shown, provided that the changed terms do not affect substantial rights of third parties." Miller, *supra* note 146, at 39. The California Civil Code states that,

> When, through fraud or a mutual mistake of the parties, or a mistake of one party, which the other at the time knew or suspected, a written contract does not truly express the intention of the parties, it may be revised on the applica-

doing so leads to the most just result, such privileging of ends over means can significantly impair the ability of commercially sophisticated parties to form optimal contracts.

IV
Conditions and the Interpretation of Express Terms

The preceding discussion has outlined a theory of contractual design that focuses on two salient strategies available to contracting parties intent on maximizing expected surplus. First, the parties must decide which portions of their agreement are to be legally enforceable and which portions should be enforced relationally. To do this, they partition their agreement into legally enforceable and unenforceable components; the latter are not included in the written document. Second, the parties must then address the design challenge of formulating the terms of their legally enforceable contract. Here, they choose between precise and vague terms as the means of shifting contracting costs between the front and back ends of the contracting process. Both of these strategies rely on the ability of parties to use formal contract doctrine in a predictable fashion. But, as we discussed in Part I, formal contract doctrine may be overridden by courts employing equitable doctrines to achieve ex post outcomes that comport with the parties' contractual ends. In this Part, we explore in depth the sharp rift between these two stages of the contemporary adjudicatory regime. We expose a deep-seated conflict in the law of conditions between formal contract doctrines requiring strict enforcement of express terms and equitable doctrines permitting abrogation of those terms when strict enforcement appears to frustrate the parties' contractual ends. We argue that courts undermine the reliability of basic doctrinal tools essential for contract design when they set aside clear express terms ex post in order to vindicate contractual ends.

A. The Tension over the Interpretation of Express Terms

The formal contract interpretation doctrines direct courts to respect the parties' express terms. Express terms can specify both primary terms governing the parties' performance obligations and secondary, or meta, terms governing the interpretation of their agreement.

> tion of a party aggrieved, so as to express that intention, so far as it can be done without prejudice to rights acquired by third persons, in good faith and for value.

Cal. Civ. Code § 3399 (West 1997).

For example, as discussed in Part II.A, express written terms can constrain a court's interpretive discretion by directing the court under the parol evidence doctrine not to admit prior evidence of implied terms.[199] Express terms therefore provide the most powerful tool available to parties for selecting the contractual means that best promote their contractual objectives: In principle, express terms not only allow the parties to communicate to each other and to courts the precise content of the terms they wish to include in their agreement, but they also allow the parties to control the extent to which courts may imply additional terms into their agreement.[200] However, because express terms themselves require interpretation, a court's tendency to equate contractual intent with the parties' intended contractual ends can lead it to impress novel meanings on express terms or imply additional terms that alter the content of the parties' agreement.

The doctrines governing the interpretation of express terms thus present yet another occasion on which courts can undermine the parties' intended contractual means in order to promote their intended ends. For example, courts have long been divided on the question of whether express terms should be given a so-called contextual or a plain meaning interpretation.[201] Under the contextual view, the interpretation of express terms inevitably requires judicial speculation about what meaning the parties were likely to have attached to the express terms they used. Thus, courts that identify contractual intent with the parties' contractual ends will assign to express terms a meaning that, at the time of adjudication, best promotes those ends instead of seeking to identify the meaning that reflects the contractual means the parties most likely intended at the time of formation.[202]

---

[199] *See, e.g.*, U.C.C. § 2-202 (amended 2003) (evidence not admissible to prove additional terms consistent with express terms of fully integrated writing); *id.* § 2-208(2) (express terms control course of performance, course of dealing, and usage of trade); RESTATEMENT (SECOND) OF CONTRACTS § 215 (1981) (evidence of prior or contemporaneous agreements or negotiations not admissible to contradict term of writing); *id.* § 216(1) (evidence of consistent additional term not admissible to supplement fully integrated agreement).

[200] *See* Goetz & Scott, *supra* note 102, at 281–83 (arguing that express terms are signals that enable parties to opt out of implied default terms and to supplement defaults with additional customized terms); Schwartz & Scott, *supra* note 6, at 584–89 (asserting that maximizing party control over express terms promotes efficient contracting).

[201] *See* SCOTT & KRAUS, *supra* note 3, at 543–45, 578–602 (discussing differences between "plain meaning" and "contextual" modes of interpretation); Schwartz & Scott, *supra* note 6, at 584 (arguing that most parties would prefer plain meaning interpretation because it would "(1) reduce contracting costs; (2) minimize the opportunities for strategic behavior; (3) reduce the risk of judicial error; and (4) expand the set of efficient contracts parties could write"). For discussion, see cases cited *supra* notes 145–47.

[202] *See, e.g.*, Pac. Gas & Elec. Co. v. G.W. Thomas Drayage & Rigging Co., 442 P.2d 641, 646 & n.9 (Cal. 1968) (finding context evidence admissible to show that promise to

Under the plain meaning view, however, courts ultimately base contract interpretation on their common knowledge of word meaning rather than context-based speculation about intent.[203] Yet even under a plain meaning regime, the interpretation of express terms can be significantly influenced by a court's desire to vindicate contractual ends instead of means. The plain meaning rule does not apply if a court finds an express term to be ambiguous; a court's threshold determination of whether an express term is ambiguous can be influenced by its perception of the effect various interpretations would have on the parties' contractual ends.[204] This deep tension between the formal obligation to respect the parties' ex ante specifications through express terms and the desire to vindicate the parties' contractual objectives ex post is most vividly illustrated in the law of conditions, to which we now turn.

B. The Schizophrenic Law of Conditions

Since parties incur duties in contracts by making promises, a party who makes an event a condition of its promise is under a duty to perform that promise only if the event occurs.[205] A common example

---

indemnify plaintiff against "all loss, damage, expense and liability" does not require defendant to indemnify plaintiff for injuries to plaintiff's property); *In re* Soper's Estate, 264 N.W. 427, 431–33 (Minn. 1935) (finding context evidence admissible to show that reference to "wife" in will does not mean person to whom testator was legally married).

[203] For example, in *Cernohorsky v. Northern Liquid Gas Co.*, the court said:

> The language of a contract must be understood to mean what it clearly expresses. A court may not depart from the plain meaning of a contract where it is free from ambiguity. In construing the terms of a contract, where the terms are plain and unambiguous, it is the duty of the court to construe it as it stands, even though the parties may have placed a different construction on it.

68 N.W.2d 429, 433 (Wis. 1955) (citation omitted). For more cases exemplifying the plain meaning view, see 11 SAMUEL WILLISTON, A TREATISE ON THE LAW OF CONTRACTS § 30:6, at 82–83 n.38 (4th ed. 1999).

[204] *Pac. Gas & Elec. Co.*, 442 P.2d at 646 n.8 ("Extrinsic evidence has often been admitted . . . on the stated ground that the contract was ambiguous. This statement of the rule is harmless if it is kept in mind that the ambiguity may be exposed by extrinsic evidence that reveals more than one possible meaning." (citation omitted)). Even if a court concludes that an express term has a plain meaning, it can narrow or expand that meaning in order to vindicate the parties' contractual ends. *See, e.g.*, Nanakuli Paving & Rock Co. v. Shell Oil Co., 664 F.2d 772, 780 (9th Cir. 1981) (holding that trade usage and course of performance evidence could vary normal meaning of "price protection" clause); Brunswick Box Co. v. Coutinho, Caro & Co., 617 F.2d 355, 359 (4th Cir. 1980) (holding that evidence of course of performance could vary normal meaning of "F.A.S." term in written contract).

[205] RESTATEMENT (SECOND) OF CONTRACTS § 224 (1981). The event on which the promise is conditioned is "largely within the control of the obligor (the homeowner's honest satisfaction with the paint job), the obligee (the insured's furnishing proof of loss), or a third person (the bank's approval of the mortgage application), or is largely beyond the control of anyone (damage as a result of fire)." 2 FARNSWORTH, *supra* note 18, § 8.2, at 395.

is an insurance contract that imposes on the insurer a duty to pay if the insured brings a claim within a specified time period after the insured suffers a covered loss. The insurer's duty to pay arises when the insured suffers a covered loss, but that duty is discharged if the insured fails to bring the claim within the specified time period. The law of conditions explicitly endorses the principle of freedom of contract by committing to the strict enforcement of all express conditions.[206] Yet, it is also home to the hoary equitable maxim that "the law abhors forfeitures."[207] The antiforfeiture norm suffuses the law of conditions, which therefore reads like a schizophrenic text, in one sentence insisting on the sanctity of strict construction and enforcement of conditions in spite of forfeiture,[208] while in the next admonishing courts, whenever interpretation allows, to avoid the conclusion that

---

[206] Renovest Co. v. Hodges Dev. Corp., 600 A.2d 448, 452–53 (N.H. 1991) ("[W]hen the parties expressly condition their performance upon the occurrence or non-occurrence of an event, rather than simply including the event as one of the general terms of the contract, the parties' bargained-for expectation of strict compliance should be given effect."); *see also* Nielsen v. Provident Sav. Life Assurance Soc'y, 66 P. 663, 665 (Cal. 1901) ("[C]onditions[,] . . . when made, must be construed and enforced . . . according to the expressed understanding of the parties making them. It is not for the courts to dispense with such limitations and conditions, nor by judicial legislation to insert a different contract from that deliberately made by the parties.").

[207] JOHN EDWARD MURRAY, JR., MURRAY ON CONTRACTS § 102(A), at 641 (4th ed. 2001); *see also, e.g.*, Naftalin v. John Wood Co., 116 N.W.2d 91, 100 (Minn. 1962) ("[It is a] well-recognized principle that forfeitures are not favored either in law or equity. . . . One claiming forfeiture carries a heavy burden of establishing his right thereto by clear and unmistakable proof."); Stevenson v. Parker, 608 P.2d 1263, 1267–68 (Wash. Ct. App. 1980) ("This court has held the general doctrine that forfeitures are not favored in the law, and that courts should promptly seize upon any circumstance . . . that would indicate an election or an agreement to waive the harsh, and at times unjust, remedy of forfeiture . . . ."). The Restatement defines "forfeiture" as the denial of compensation for losses the promisee incurs when "[t]he non-occurrence of a condition of [the promisor]'s duty . . . cause[s] the [promisee] to lose his right to the agreed exchange after he has relied substantially on the expectation of that exchange . . . ." RESTATEMENT (SECOND) OF CONTRACTS § 227 cmt. b (1981).

[208] RESTATEMENT (SECOND) OF CONTRACTS § 227 cmt. b (1981) ("The policy favoring freedom of contract requires that, within broad limits . . . the agreement of the parties should be honored even though forfeiture results."); *id.* § 226 cmt. c ("[T]o the extent that the parties have, by a term of their agreement, clearly made an event a condition, they can be confident that a court will ordinarily feel constrained strictly to apply that term . . . ."); *id.* § 229 cmt. a ("[I]f the term that requires the occurrence of the event as a condition is expressed in unmistakable language, the possibility of forfeiture will not affect the interpretation of that language."); *see also* Gillman v. Bally Mfg. Corp., 670 A.2d 19, 21 (N.J. Super. Ct. App. Div. 1996) ("'[E]quity's jurisdiction in relieving against a forfeiture is to be exercised with caution . . . [A] court of equity will not interfere to substitute a different and more liberal agreement' than that which existed between the parties." (quoting Dunkin' Donuts of Am. v. Middletown Donut Corp., 495 A.2d 66, 74 (N.J. 1985) and, Fox v. Haddon Twp., 45 A.2d 193, 196 (N.J. Ch. 1945))).

the promisor's obligation is subject to an enforceable condition if enforcement of the condition would raise the specter of forfeiture.[209]

The antiforfeiture norm takes both an ex ante and ex post form. In its ex ante form, the antiforfeiture norm controls the interpretation of whether an obligation is subject to a condition.[210] The ex ante version of the antiforfeiture norm applies when the language of an agreement fails to make clear whether it is intended to subject the promisor's duty to a condition, to impose a duty on the promisee, or both. For example, an employment contract might provide that the employer shall not terminate the employee without just cause, but also provide a clause stating that "the employee will, within thirty days of termination, give written notice to the employer of any claim of wrongful termination and will not take any legal action based on the claim within six months of such notice."[211] Suppose that after being terminated without cause the employee violates this provision by failing to provide written notice of his claim. A court could interpret the clause as a promise by the employee to provide written notice. On this interpretation, the employer remains subject to suit for wrongful discharge but would be entitled to recover damages from the employee for breach of her promise. A court might instead interpret this language as creating a condition precedent to the employee's right to sue for breach of the employer's promise not to terminate without just cause. On this view, the employee's failure to satisfy the condition would relieve the employer of any liability for wrongful dis-

---

[209] *See, e.g.*, Bornholdt v. S. Pac. Co., 327 F.2d 18, 20 (9th Cir. 1964) ("A condition involving a forfeiture must be strictly interpreted against the party for whose benefit it is created. . . . Where there are two possible constructions, one of which leads to a forfeiture and the other avoids it, the rule of law is well settled . . . that the construction which avoids forfeiture must be made if it is at all possible." (citing CAL. CIV. CODE § 1442 (West 2007))); Kalina v. Eckert, 497 A.2d 1384, 1385 (Pa. Super. Ct. 1985) ("[A] provision will not be construed to result in a forfeiture unless no other reasonable construction is possible."); Leitner v. Lonabaugh, 402 P.2d 713, 720 (Wyo. 1965) ("Even as a general rule conditions, if possible, are to be construed 'so as not to operate as a forfeiture of the rights of the parties.'" (quoting Pac.-Wyo. Oil Co. v. Carter Oil Co., 226 P. 193, 198 (Wyo. 1924))).

[210] Courts use general principles of interpretation to identify and distinguish conditions and duties. RESTATEMENT (SECOND) OF CONTRACTS § 226 cmt. a (1981). In addition, courts have developed specialized principles for resolving interpretive questions regarding conditions. *See id.* ("There are also some special standards of preference that are of particular applicability to conditions, and these are set out in § 227."). The ex ante version of the antiforfeiture norm falls into this second category.

[211] A contractual provision similar to this appears in *Inman v. Clyde Hall Drilling Co.*, 369 P.2d 498, 499 n.2, 500 (Alaska 1962) (finding that provision is not "unfair or unreasonable"). Employment contracts often contain such provisions to give the employer the opportunity to settle disputes and take measures to prevent ongoing or future violations of contractual or statutory provisions.

charge.[212] Following the antiforfeiture maxim, contract law regards the employee's detrimental reliance on the just cause restriction on termination—in which the employee forfeits any claim for wrongful discharge by failing to notify the employer—as a forfeiture to be avoided, whenever the contract's express language and circumstances allow, by favoring an interpretation of the term as creating a promissory duty rather than a condition.[213]

The rationale underlying these interpretive rules is not that the law should intervene to prevent forfeitures that have materialized under an agreement but rather that the promisee is unlikely to have agreed to a risk of forfeiture at the time of formation.[214] Thus, if finding that a promise is subject to a condition would impose a forfeiture on the promisee but the parties would not have understood the condition to create such a risk at the time of formation, then the ex ante antiforfeiture norm allows the courts to interpret the term as a promise rather than a condition.

Even if the parties succeed in writing an express term that unequivocally creates a condition, the ex post form of the antiforfeiture norm strongly encourages courts to exercise their discretion to excuse the condition whenever its enforcement would create a forfeiture and the court deems the condition not to have been a material part of the agreement at the time of formation.[215] In addition, even if

---

[212] A court might also interpret the language as creating both a promise and a condition, in which case the employee's failure to satisfy the contract terms not only relieves the employer of liability but also subjects the employee to liability for breach of her promise.

[213] For example, in *United-Buckingham Freight Lines v. Riss & Co.*, the court held that

> Any duty . . . to cooperate imposed upon plaintiff by . . . the contract is obviously in the nature of . . . a promise rather than a condition. To permit [defendant's] construction of the covenant would be to construe it as a condition precedent, the breach of which would result in a forfeiture [for the plaintiff]. . . . It is axiomatic that the law does not favor forfeitures. Therefore, any failure to cooperate . . . on the part of the plaintiff clearly could do no more than breach an independent promise.

241 F. Supp. 861, 863 (D. Colo. 1965). For an example in which the court construes a "pay if paid" clause to be a promise rather than a condition because "it is a rule of construction that a forfeiture, by finding a condition precedent, is to be avoided when another reasonable reading of the contract exists," see *Sheldon L. Pollack Corp. v. Falcon Indus., Inc.*, 794 S.W.2d 380, 383 (Tex. App. 1990).

[214] The Restatement explains:

> Since the intentions of the parties must be taken as of the time the contract was made, the test is whether a particular interpretation would have avoided the risk of forfeiture viewed as of that time, not whether it will avoid actual forfeiture in the resolution of a dispute that has arisen later.

RESTATEMENT (SECOND) OF CONTRACTS § 227 cmt. b (1981).

[215] *See id.* § 229 ("To the extent that the non-occurrence of a condition would cause disproportionate forfeiture, a court may excuse the non-occurrence of that condition unless its occurrence was a material part of the agreed exchange.").

a court agrees that a contract contains a material, express condition, the ex post norm encourages the court to find that the promisor has implicitly waived the condition, either retrospectively or prospectively, whenever enforcement of the condition would create a forfeiture.[216]

In sum, the law of conditions explicitly stacks the deck heavily against the finding and enforcement of conditions on the ground that the law abhors a forfeiture. In the discussion that follows, we argue that the antiforfeiture norm is based on two false assumptions. The first is that parties are unlikely to select terms that create the risk of forfeiture. This assumption underlies the doctrine directing courts to avoid finding a condition absent express language that unmistakably creates it. The second assumption is that express conditions are sometimes not material at the time of formation. This assumption underlies the doctrine directing courts to avoid enforcing even clear, express conditions.

We have argued that considerations of contract design often favor the selection of precise terms that create rule-like obligations that are easy for the parties to observe and to enforce in court. Express conditions serve just this purpose: They afford a promisor protection from certain risks, in lieu of having to prove losses that may be difficult to verify in a suit for damages. When sophisticated commercial parties clearly agree to express conditions, there is no systematic reason to doubt that the promisee understood the risk of forfeiture and bargained for compensating contractual benefits from the promisor. On this view, conditions are always material from the ex ante perspective because they allocate risks between the parties, the contract compensates each party for bearing those risks, and the parties inevitably rely on that allocation of risks. Since materiality is determined by the parties' intent at the time of formation, conditions will always be material.

---

216 *See, e.g.*, Knarston v. Manhattan Life Ins. Co., 73 P. 740, 741 (Cal. 1903) ("[T]he right to declare a forfeiture, being a matter entirely for the benefit of a lessor or vendor, can be, even by parol, effectually waived by either."); Bielski v. Wolverine Ins. Co., 150 N.W.2d 788, 790 (Mich. 1967) ("[W]aivers [of contract clause requiring arbitration as condition precedent to suit] need not be expressed in terms, but may be implied by the acts, omissions, or conduct of the insurer or its agents authorized in such respect."); Cochran v. Grebe, 578 S.W.2d 351, 354 (Mo. Ct. App. 1979) ("Forfeitures are highly disfavored by the law and the courts are therefore quick to find a waiver or estoppel in a case [producing hardship]."); Miraldi v. Life Ins. Co., 356 N.E.2d 1234, 1236 (Ohio Ct. App. 1971) (stating that law does not favor forfeiture and that waiver will be inferred whenever it reasonably can be from facts); Brown v. Powell, 648 N.W.2d 329, 333 (S.D. 2002) ("Because forfeitures of land sale contracts are highly disfavored by the law, courts are generally quick to find a waiver of conditions alleged as a basis for a claim of breach.").

C. Express Conditions of Satisfaction: Corthell v. Summit Thread Co.

We have seen how the antiforfeiture norm gives rise to interpretive rules favoring promises over conditions. This norm also influences the interpretation of contracts that subject a promise to the express condition of the promisor's satisfaction with the promisee's performance. If the express condition does not specify whether the promise is subject to the promisor's objectively reasonable satisfaction or subjective but honest satisfaction, the law directs courts to prefer the former unless it is not practicable to determine "objective reasonableness."[217] The rationale of this interpretive rule is that a promisee ordinarily is unlikely to agree to subject himself to the idiosyncratic tastes of the promisor in circumstances where objectively reasonable satisfaction would be practical to ascertain.[218] By its own terms, however, this rule applies only in the absence of clear language creating a condition of honest, subjective satisfaction.[219] Yet the equitable instinct to avoid perceived forfeitures is so strong that it often overrides the doctrinal caveats directing courts to enforce such express conditions even when they do create a forfeiture. A court is especially likely to misapply formal doctrine when the formal terms of a contract would impose a forfeiture on a promisee and the court fails to understand why it might have been rational for the promisee to agree to subject himself to this risk. To illustrate this tendency, we turn to a case in which the court needed to decide the effect to give to a condition that appeared to make the promisor's obligation illusory.

1. The Corthell-Summit Thread Contract

In *Corthell v. Summit Thread Co.*,[220] Robert Corthell, an employee of the Summit Thread Company, promised to turn over patents for three of his existing inventions, as well as all of his future

---

[217] *See* RESTATEMENT (SECOND) OF CONTRACTS § 228 (1981) ("When it is a condition of an obligor's duty that he be satisfied with respect to the obligee's performance or with respect to something else, and it is practicable to determine whether a reasonable person in the position of the obligor would be satisfied, an interpretation is preferred under which the condition occurs if such a reasonable person in the position of the obligor would be satisfied.").

[218] *See id.* § 228 cmt. b ("When, as is often the case, the preferred interpretation [i.e. relying on an objective standard] will reduce the obligee's risk of forfeiture, . . . there is an additional argument in its favor.").

[219] *See id.* § 228 cmt. a ("If the agreement leaves no doubt that it is only honest satisfaction that is meant and no more, it will be so interpreted, and the condition does not occur if the obligor is honestly, even though unreasonably, dissatisfied.").

[220] 167 A. 79 (Me. 1933).

inventions, for a period of five years.[221] In return for this, Summit Thread promised to increase Corthell's annual salary by $620 for five years, make Corthell a onetime payment of $3500, and provide Corthell additional "reasonable recognition" for his future inventions, "the basis and amount of recognition to rest entirely with the Summit Thread Company at all times."[222] The written agreement between the parties additionally specified that the terms of the contract were "to be interpreted in good faith on the basis of what is reasonable and intended, and not technically."[223] During the term of the contract, Corthell turned over four additional inventions to Summit Thread and requested, but never received, additional payment for them as "reasonable recognition" from Summit Thread.[224] Corthell brought suit seeking payment for the four inventions. Summit Thread argued that "the vagueness and uncertainty of [the reasonable recognition] provisions relating to the price to be paid render[ed] the contract unenforceable."[225]

In the first part of the agreement, Summit Thread promised Corthell reasonable recognition for his inventions.[226] Had the contract provision stopped there, standard interpretive canons should have led a court to enforce it by implying a term requiring Summit Thread to pay Corthell the value of his inventions according to an objective standard. But the agreement further stipulated that the amount of recognition rested "entirely" within Summit Thread's discretion.[227] Summit Thread's promise of reasonable recognition thus expressly granted Summit Thread unfettered discretion to determine the amount of recognition, if any, that was reasonable. Under formal contract doctrine, a promise is illusory if it is subject to a condition that is entirely within the promisor's control and the promisor incurs no detriment by deciding not to satisfy that condition.[228] Summit

---

221 *Id.* at 80.

222 *Id.* By "reasonable recognition," the parties apparently meant reasonable compensation. *Id.* at 81 ("No contention is made that the term 'reasonable recognition,' as used in the contract under consideration, means other than reasonable compensation . . . .").

223 *Id.* at 82.

224 *See id.* at 80–81.

225 *Id.* at 81.

226 *Id.* at 80.

227 *Id.*

228 *See* RESTATEMENT (SECOND) OF CONTRACTS § 77 cmt. a (1981) ("Words of promise which by their terms make performance entirely optional with the 'promisor' do not constitute a promise. . . . Where the apparent assurance of performance is illusory, it is not consideration for a return promise."); 1 SAMUEL WILLISTON, THE LAW OF CONTRACTS § 43 (1921) ("One of the commonest kind of promises too indefinite for legal enforcement is where the promisor retains an unlimited right to decide later the nature or extent of his performance. This unlimited choice in effect destroys the promise and makes it merely illusory.").

Thread's promise thus appears to qualify as a classic illusory promise. Historically, courts have refused to enforce illusory promises on the ground that they are not supported by consideration.[229] If the court in *Corthell* had followed suit, it would have declared Summit Thread's promise unenforceable. If Corthell had not yet performed, he also would have been free from any legal obligation to turn over his inventions. But because Corthell had already performed by turning over his inventions, he would have made a gratuitous transfer and therefore would have been denied recovery in contract.[230]

2. The Court's Analysis

The court in *Corthell* instead implied a term imposing objective duties on a promisor whose indefinite promise otherwise would be unenforceable. Although the court noted that reasonable recognition was "coupled with the reservation that the 'basis and amount of recognition (was) to rest entirely with' the company 'at all times,'" it noted that "[n]evertheless, the contract was 'to be interpreted in good faith on the basis of what is reasonable and intended, and not technically.'"[231] On the basis of these provisions, the court concluded:

> [T]he parties continued to exhibit a contractual intent and a contemplation of the payment of reasonable compensation to the plaintiff for his inventions. The company was not free to do exactly as it chose. Its promise was not purely illusory. It was bound in good faith to determine and pay the plaintiff the reasonable value of what it accepted from him.[232]

The court's justification for its interpretation appears to be that the clause granting Summit Thread "entire" discretion to set the amount of compensation is, on its face, inconsistent with the contract's express

---

229 Paul v. Rosen, 122 N.E.2d 603 (Ill. App. Ct. 1954); Wickham & Burton Coal Co. v. Framers' Lumber Co., 179 N.W. 417, 419–20 (Iowa 1920); Strong v. Sheffield, 39 N.E. 330 (N.Y. 1895).

230 Corthell could have sued in restitution for a *quantum valebant* recovery of the value of the inventions he turned over to Summit Thread. But in our view, a court that properly ruled against recovery in contract would for the same reasons deny recovery in restitution on the ground that Corthell accepted the risk that Summit Thread would retain his inventions without paying him for them.

231 *Corthell*, 167 A. at 82.

232 *Id.* The court offered no reason to believe that the literal interpretation of the word "entirely" (as meaning entirely) was a technical definition, rather than the simple plain meaning of the language the parties used. As defendant argued in its brief, "if the basis and amount of recognition lies entirely with the defendant at all times, it lies with the defendant now, not with the plaintiff and not with any third party or parties, not even with a court of law." Brief for the Defendant at 15, *Corthell*, 167 A. 79 (docket no. not available) (on file with the *New York University Law Review*).

language requiring that the contract be interpreted in good faith.[233] Presented with two clauses that it believed were inconsistent, the court gave priority to the good faith term and effectively set aside the clause granting Summit Thread unfettered discretion.[234]

But even when contracts contain no express language requiring parties to act in good faith, contract law implies such a duty into all contracts.[235] *Corthell* then might be read to suggest that promises can never be illusory, for the implied duty of good faith would always cure an otherwise illusory promise by imposing enforceable duties on the promisor. Yet this is clearly not the case. The illusory promise doctrine still operates to render a promise unenforceable when its clear language unequivocally subjects the promise to a condition that is entirely within the promisor's control such that the promisor can avoid satisfying the promise without incurring any detriment.[236] Instead, courts use doctrines such as the implied duty of good faith to impose duties on the promisor only if they find the promisee reasonably believed the promisor's discretion was implicitly constrained.[237] Thus, the implicit premise of the *Corthell* decision is that it was reasonable for Corthell to believe that Summit Thread was undertaking an obligation to pay an objectively reasonable amount in compensation for his inventions, and therefore intended to be legally bound by its promise, despite the express contractual language stating that the amount of compensation to be paid to Corthell rested "entirely" with Summit Thread.[238]

The *Corthell* court, however, provides no plausible account of why the parties expressly stipulated that the amount of recognition was to rest entirely with Summit Thread. As an interpretation of the parties' agreement, *Corthell* fails the basic requirement of providing at least a minimally adequate account of the express language in the par-

---

233 *Corthell*, 167 A. at 80 (noting that contract includes following clause: "All of the above is to be interpreted in good faith on the basis of what is reasonable . . . ").

234 *Id.* at 82.

235 U.C.C. § 1-304 (amended 2003); RESTATEMENT (SECOND) OF CONTRACTS § 205 (1981).

236 *See* RESTATEMENT (SECOND) OF CONTRACTS § 76 cmt. d (1981) ("Words of promise do not constitute a promise if they make performance entirely optional with the purported promisor. . . . [T]here may be consideration if forbearance from causing the condition to occur would itself have been consideration if it alone had been bargained for.").

237 *See, e.g.*, Mattei v. Hopper, 330 P.2d 625, 627–28 (Cal. 1958) (finding that "promisor's duty to exercise his judgment in good faith [was] adequate consideration to support the contract" and that it prevented contract from "nullifying the consideration otherwise present in the promises exchanged"); Seymour Grean & Co. v. Grean, 82 N.Y.S.2d 787, 788–89 (App. Div. 1948) (per curiam) (finding that employment contract was not illusory because promisor was required to render substantial services and to act in good faith).

238 *Corthell*, 167 A. at 80, 82.

ties' writing. The court simply treats the writing as if it did not contain the word "entirely." Why would the court in *Corthell* insist on an interpretation seemingly contradicted by the writing's express language? The answer, we believe, is that it could not otherwise explain why Corthell would have been willing to promise to transfer his future inventions to Summit Thread. If the parties intended Summit Thread to have unfettered discretion to decide the amount, if any, of compensation to be paid to Corthell after he turned over an invention, then Corthell would be subjecting himself to the risk of giving something in return for nothing. Since the court was unable to explain why a minimally rational party would promise to turn over potentially valuable inventions without requiring that the beneficiary pay for any benefit it received, the court simply dismissed the possibility that the parties could have meant the word "entirely" to be taken seriously. The express clause directing the court to interpret the contract language in good faith provided convenient linguistic cover to justify an interpretation of the contract that eviscerated the plain meaning of the word "entirely." The better understanding of that clause was not to undermine Summit Thread's complete discretion to make an honest judgment about the value of the invention but rather to prevent Summit Thread from dishonestly claiming that no compensation was due for inventions that had objectively verifiable value.

3. A Contract Design Explanation: Conditions and Relational Enforcement

The *Corthell* court's apparent conclusion that no rational person would have accepted the risk that Summit Thread would pay him nothing in return for a valuable invention likely results from the court's failure to appreciate the role that asymmetric information and relational enforcement play in designing contracts.[239] Contract theory teaches us that the more difficult information is to verify, the less likely it is that parties will subject their contractual obligations to conditions whose satisfaction can be proved only by verifying that information.[240] This basic principle suggests a perfectly reasonable story

---

[239] The explanation for the court's apparent failure to appreciate why reasonable commercial parties might rationally choose relational enforcement over legal enforcement mirrors our analysis in Part II.B.2, *supra*, of the court's decision in *Hunt Foods*.

[240] Recent scholarship argues against the standard assumption underlying traditional contract theory literature that information is either verifiable or not. In fact, information falls along a continuum of verifiability; where a piece of information falls on that continuum is a function of both its inherent character and the contingent rules of contract interpretation and evidence law. *See* Scott & Triantis, *supra* note 10, at 825–26 (noting that verifiability of contractual obligations at trial are highly context-specific and endogenous to judicial process).

explaining the plain language of the contract in *Corthell*. Summit Thread would not be willing to condition its performance obligation on the objective value of Corthell's inventions unless that value was likely to be verifiable. But that value was not likely to be easily verifiable. First, as a general matter, the economic value of intellectual property is often difficult to determine. In this case, the value of Corthell's inventions is especially difficult to determine: Like the previous inventions Corthell sold to Summit Thread,[241] the inventions at issue were relatively minor variations on the design of thread-handling devices, such as spools, bobbins, and shuttles, that Summit Thread and its customers used to store and sew the thread it manufactured. Summit Thread manufactured and used spools, bobbins, and shuttles, but it did not sell them.[242] Thus, the value, if any, of these inventions to Summit Thread consisted exclusively in the value placed on these inventions by Summit Thread's customers. That value consisted either in the increased ease of using the spools on which Summit Thread's thread was wound when delivered or in the increased effectiveness of the sewing machine parts that Summit Thread supplied free of charge to its customers. Presumably, the easier it was for Summit Thread's customers to use Summit Thread's thread and the more valuable Summit Thread's free machine parts were to its customers, the more likely they would be to buy their thread from Summit Thread and to pay a higher price for it.

---

[241] Corthell's early inventions were "bobbin control adjuncts" and "guarding attachments for thread caps." Brief of the Plaintiff at 4, *Corthell*, 167 A. 79 (docket no. not available) (on file with the *New York University Law Review*).

[242] Corthell sought compensation for four inventions. His first invention was a minor variation on the "King Spool" on which Summit Thread spooled the thread it sold. Brief for the Defendant, *supra* note 232, at 16. Corthell maintained that the principal value of this invention was that it effected "the continuation of th[e] King Spool patent." Brief of the Plaintiff, *supra* note 241, at 20. He did not claim that the value of this invention was attributable to any increase in the spool's practical value. Indeed, Corthell was prepared to concede that the invention had never been used, let alone sold. The second invention consisted of "an adjunct to sewing machine shuttles to be attached to bobbins." Brief for the Defendant, *supra* note 232, at 20. However, "[Summit Thread] never sold bobbins." *Id.* at 21. The third invention consisted "of celluloid discs . . . to be attached to certain sewing machine shuttles." *Id.* at 6. According to Summit Thread's brief, "[i]nsofar as the discs were distributed to the public it was a gratuitous distribution." *Id.* at 30. And the fourth invention consisted of "paper or celluloid discs to be attached to all bobbins used by [Summit Thread]." *Id.* at 31. Corthell maintained that the value of this invention was its unrealized potential to allow Summit Thread to attach a disc to the bobbins it manufactured and distributed to its customers without violating a patent by one of Summit Thread's competitors. Brief of the Plaintiff, *supra* note 241, at 22–23. The invention was never patented and neither Summit Thread nor any other company manufactured or used it, apart from one made for an experimental trial, which failed. Brief for the Defendant, *supra* note 232, at 32.

Clearly, Summit Thread believed that Corthell's previous inventions were sufficiently valuable to warrant paying him $3500 and negotiating a contract renewal that included a clause entitling Summit Thread to Corthell's future inventions during the contract term.[243] Even if Corthell could observe the inventions' gross market value, Corthell would not know Summit Thread's costs and revenues and thus would have great difficulty verifying the net market value of the inventions to a court. Since neither party expected that Summit Thread would sell Corthell's inventions (and in fact none of them were ever sold), no objective measure of their value would be available. Instead, their value to Summit Thread's revenue stream would have to be distinguished from the many other inputs that contribute to Summit Thread's sales revenues. The complexity and subjectivity of such an evaluation would invite litigation over value as well as strategic attempts by both parties to either inflate or deflate the value assessment. Contract design theory thus provides a plausible explanation for why a commercial party such as Summit Thread,[244] in this situation, would prefer not to condition performance under such a contract on difficult-to-verify private information.

Defending the literal interpretation of the word "entirely," however, also requires a demonstration that a sophisticated person in Corthell's position plausibly could have given the word "entirely" a literal interpretation as well. There are several plausible reasons why Corthell might have agreed to turn over his future inventions without requiring Summit Thread to make a legally enforceable promise to pay for them. First, Summit Thread had already demonstrated its willingness to pay him for such inventions in the past. In the same agreement in which he promised to turn over his future inventions, Corthell received a payment of $3500 for inventions already submitted to Summit Thread. This earlier interaction revealed Summit Thread's propensity to reciprocate and created an element of trust in the relationship.[245] Second, the agreement between Corthell and Summit

---

243 *Corthell*, 167 A. at 80.

244 *See infra* note 250 (explaining why Summit Thread is sophisticated commercial party).

245 *See* Scott, *Self-Enforcing Agreements*, *supra* note 9, at 1682–83 (discussing apparently widespread prevalence of "comfort agreements" in which parties use legally unenforceable agreements to learn about each side's propensity to reciprocate). Recall that the experimental evidence shows that a preference for reciprocity can motivate cooperation even in arm's length interactions. This evidence shows that many people behave in a reciprocal manner by responding cooperatively to generous acts, and, conversely, punishing noncooperative behavior. The observed preference for reciprocity is heterogeneous. Scott & Stephan, *Self-Enforcing International Agreements*, *supra* note 11, at 565–66 (discussing reciprocity theory and concluding that "this is a heterogeneous world where some people exhibit reciprocal fairness and others are selfish" and that "[t]aking all the experiments

Thread constituted a repeat-play game in which Summit Thread's defection from its nonlegal commitment to pay Corthell the reasonable value of his inventions could be punished by Corthell's refusal to produce more inventions during the remainder of the contract period.[246] Because the contract placed him under no obligation to create inventions, Corthell could protect himself by electing to withdraw from the relationship.[247] In many contexts, the norm of reciprocity and the prospect of repeat dealings are the best available means of regulating the contractual relationship.[248] In short, as a general matter, Corthell had good reason to believe that his contract with Summit Thread was self-enforcing and that Summit Thread would honor its commitment even if not legally required to do so.

In sum, it is plausible that Summit Thread and Corthell might have concluded that the difficulty in verifying the value of the inventions made the option of legal enforcement more costly than the alternative of informal (or relational) enforcement.[249] While it is clear that the shared contractual end of the parties was for Summit Thread to pay Corthell the reasonable value of the inventions he created during the contract period, it seems equally clear that the parties had good reason to choose nonlegal means to achieve those ends. Regardless of whether they did so, the important point is that, for the reasons discussed above, sophisticated commercial parties would prefer that courts preserve the choice between informal and formal enforcement, a choice that requires courts both to give full effect to express terms and to adhere to the nonenforcement doctrines of indefiniteness and

---

together . . . the fraction of fair subjects ranges from forty to sixty percent as does the fraction of subjects who are selfish"); *see also* Fehr, Gachter & Kirchsteiger, *supra* note 133, at 850 (noting that "reciprocal motivations have important implications for the enforcement of contracts" and that "[i]n view of the powerful behavioral impact and efficiency consequences of reciprocity . . . it seems doubtful that one can design optimal incentive contracts on the basis of a neglect of reciprocal motivations"); Fehr & Schmidt, *supra* note 133, at 818 ("Some pieces of evidence suggest that many people are driven by fairness considerations[;] other pieces indicate that virtually all people behave as if completely selfish . . . .").

246 When parties contemplate making a series of contracts, neither party will breach an early contract if the gains from that breach are lower than the expected profits from future contracts that a breach would eliminate. *See supra* note 131 and accompanying text.

247 Moreover, if Summit Thread's promise is interpreted as illusory and therefore unenforceable, any return promise made by Corthell would likewise be unenforceable. *See* RESTATEMENT (SECOND) OF CONTRACTS §§ 17, 18, 71 (1981) (establishing key doctrine that promise unsupported by return promise or performance is unenforceable).

248 *See supra* Part II.B.1 (analyzing parties' motivations in Hunt Foods v. Doliner, 270 N.Y.S.2d 937 (App. Div. 1966)).

249 Moreover, Corthell could limit the risk of Summit Thread's breach of their informal agreement by reducing his efforts to develop future inventions until Summit Thread paid him the reasonable value of inventions as he produced them *seriatim*.

illusory promises.[250] This prediction from contract design theory is supported by the available evidence. Once again, New York and California are exemplars of the differences in the evolution of contract common law in the various states. The strong preference of sophisticated parties for New York contract law is consistent with our claim that these parties prefer an adjudication system that faithfully and consistently applies formal doctrine, absent the parties' express indication otherwise at the time of formation.[251] New York generally adheres to formal contract doctrine and, in particular, adheres to the plain meaning rule of interpretation.[252] On the other hand, California law, notably disfavored by commercial parties, rejects the common law plain meaning rule in lieu of a rule of interpretation that relies on contextual evidence to ascribe meaning to express terms.[253] This clear

---

[250] According to our working definition of a "sophisticated commercial party," *see supra* note 6, Summit Thread is a sophisticated commercial party but Corthell is not. For purposes of illustration, however, we presume that Corthell was represented by counsel, and thus his contract does not fall outside the scope of our thesis. Our claim that sophisticated commercial parties would prefer formal doctrine to apply to agreements like the one in *Corthell* is buttressed by substantial evidence that many parties to long-term, collaborative supply contracts choose to create agreements based on illusory promises that are legally unenforceable. *See, e.g.*, Stanadyne Corp., Deere & Co. and Stanadyne Corp. Long Term Agreement (Form 10-K, Exhibit 10.15), at § IV.F (Mar. 28, 2002), *available at* http://www.secinfo.com/dRsjx.33q.c.htm. The contract between Deere and Stanadyne is not legally enforceable because it does not actually require that the parties do anything. Although the contract does refer to anticipated levels of Deere purchases, Stanadyne does not have to produce any parts, and if it does produce them, Deere is under no obligation to take them. Thus, the promises are illusory. The parties in the Deere-Stanadyne contract could easily have written a legally enforceable supply contract. The parsimonious conclusion is that they chose to avoid legally enforceable commitments and instead chose to rely largely on relational enforcement:

> [T]he Deere-Stanadyne contract resembles a more famous contract that, over the years, has been the focus of a great deal of academic attention: the General Motors-Fisher Body supply contract for the supply of auto bodies to GM in the 1920s. As Victor Goldberg has recently shown, the General Motors-Fisher Body supply contract was, in truth, legally unenforceable [because GM's promise was illusory].

Ronald J. Gilson, Charles F. Sabel & Robert E. Scott, *Contracting for Innovation: Vertical Disintegration and Interfirm Collaboration*, 109 COLUM. L. REV. 431, 477 (2009) (citing Victor P. Goldberg, *Lawyers Asleep at the Wheel? The GM-Fisher Body Contract*, 17 INDUS. & CORP. CHANGE 1071, 1076 (2008)). GM and Fisher Body apparently chose to rely on a variety of relational mechanisms instead of legal enforcement.

[251] *See supra* notes 148–50 and accompanying text.

[252] *See, e.g.*, Vt. Teddy Bear Co. v. 538 Madison Realty Co., 807 N.E.2d 876, 879 (N.Y. 2004) ("When interpreting contracts, we have repeatedly applied the 'familiar and eminently sensible proposition of law [ ] that, when parties set down their agreement in a clear, complete document, their writing should . . . be enforced according to its terms.'" (alterations in original)); *In re* Wallace v. 600 Partners Co., 658 N.E.2d 715, 717 (N.Y. 1995) (stating that plain meaning rule imparts stability to commercial and property transactions "where commercial certainty is a paramount concern").

[253] *See supra* notes 147–50 and accompanying text.

choice-of-law preference evinces the degree to which sophisticated parties (who are able to access the forum of their choice) value the plain meaning interpretation of their express contract terms.

D. Excuse of Conditions and Rules Versus Standards

Corthell illustrates how courts can subvert formal doctrines by manipulating the process of contractual interpretation and applying those formal doctrines in a manner that covertly overrides formal contract terms. In this case, the court used interpretation to avoid application of the illusory promise doctrine to the parties' contract. We have argued that the *Corthell* court felt constrained to circumvent the contract's formal terms in order to maintain fidelity to the parties' intended ends at the time of formation. But the *Corthell* decision could also be explained as an application of the ex post forfeiture norm as embodied in the equitable doctrine governing excuse of conditions. On this view, the question of whether Corthell knowingly took the risk of forfeiture when he entered into the agreement would be irrelevant. Equitable excuse doctrine permits a court to invalidate express conditions "[t]o the extent that the nonoccurrence of a condition would cause disproportionate forfeiture . . . unless its occurrence was a material part of the agreed exchange."[254] According to this ex post version, the court simply refused to enforce a term that clearly created a condition of the promisor's obligation because the court believed that the condition was not a material part of the parties' agreement, and thus its enforcement would impose a substantial reliance loss on Corthell. Courts regard such losses as forfeitures and the resulting gains for the promisor as a form of unjust enrichment.[255]

The excuse account of the *Corthell* decision requires an explanation for why the court would determine that the express condition was not a material term of the parties' agreement. Although the court does not provide an explicit account of how the excuse doctrine applies to the case, the reasons that explain the ex ante account also explain why the court would have reached the same conclusion under the excuse doctrine. If a court fails to appreciate the information barriers to legal enforcement and the role that relational mechanisms can play in enforcing commitments that are not legally binding, it would understandably fail to see why a term granting unfettered discretion to determine the reasonable value of the promisor's performance was

---

[254] RESTATEMENT (SECOND) OF CONTRACTS § 229 (1981).

[255] *See, e.g.*, *id.* § 371; E. ALLAN FARNSWORTH, 3 FARNSWORTH ON CONTRACTS §§ 12.19, 12.20a (3d ed. 2004). For discussion of forfeitures, see cases cited in note 207 *supra*.

material to the design of the contract. As we have explained above, the *Corthell* court appeared to treat that term as anomalous, which suggests that the court did not regard it as material.

The above recharacterization of *Corthell* provides one illustration of how the doctrine governing excuse of conditions is premised on reasoning that fails to appreciate the role of intended contractual means in contract design. However, the degree to which the antiforfeiture norm undermines the structure of ex ante contractual intent can best be shown by the official illustrations included in § 229 of the Restatement (Second) of Contracts. Unlike *Corthell*, each of these illustrations expressly embraces the excuse doctrine. Consider the first illustration, loosely based on the celebrated case of *Jacob & Youngs, Inc. v. Kent*:[256]

> A contracts to build a house for B, using pipe of Reading manufacture. In return, B agrees to pay $75,000 in progress payments, each payment to be made "on condition that no pipe other than that of Reading manufacture has been used." Without A's knowledge, a subcontractor mistakenly uses pipe of Cohoes manufacture which is identical in quality and is distinguishable only by the name of the manufacturer which is stamped on it. The mistake is not discovered until the house is completed, when replacement of the pipe will require destruction of substantial parts of the house. B refuses to pay the unpaid balance of $10,000. A court may conclude that the use of Reading rather than Cohoes pipe is so relatively unimportant to B that the forfeiture that would result from denying A the entire balance would be disproportionate, and may allow recovery by A subject to any claim for damages for A's breach of his duty to use Reading pipe.[257]

The key to the application of the excuse doctrine in this illustration is the conclusion that the use of Reading pipe is "relatively unimportant" to B.[258] So the question presented in this stylized version of *Jacob & Youngs* is whether the court is correct in concluding that the condition was not material when the parties included it in their agreement. Implicit in the reasoning of the Restatement is the view that the condition is self-evidently nonmaterial on the facts as stated: No

---

256 129 N.E. 889 (N.Y. 1921). The Restatement omits facts critical to the outcome of the actual case. *See* discussion of those differences *infra* note 261.

257 RESTATEMENT (SECOND) OF CONTRACTS § 229 cmt. b, illus. 1 (1981).

258 Note that under § 229, the court can excuse the condition only if the parties regarded the condition as nonmaterial at the time of formation. The mere fact that the actual failure to satisfy the condition does not materially affect the promisor's interests ex post is irrelevant. *Id.* § 229 cmt. c. The question is not whether "the actual non-occurrence [of the condition that Reading pipe must be used] happened to involve a departure that was not a material part of the agreed exchange, [but rather whether] the occurrence of the condition was a material part of that exchange." *Id.*

reasonable person would believe that the condition of using Reading brand pipe was material to B because the Cohoes brand pipe that was in fact used "is identical in quality and distinguishable only by the name of the manufacturer which is stamped on it."[259] This reasoning, however, begs the question of why the parties chose to make B's payment obligations expressly conditional on A's installation of Reading brand pipes only, instead of simply including a promise by A to use Reading pipe, thereby subjecting it only to damages for any losses caused by its use of nonconforming pipe.

It is, of course, possible that the drafting was careless or made in ignorance of the legal implications of making a contractual obligation a condition rather than a promise. But such a conclusion would require some objective evidence to override the strong presumption that commercially sophisticated parties such as A, the contractor, exercise reasonable care in executing their agreements and know or should know the legal implications of the express contractual language to which they agree. Instead, the Restatement illustration invites the court to excuse the condition if it believes that the promisor did not regard the condition as relatively important at the time of agreement, even though the court concedes that the parties deliberately chose to create a condition instead of a promise. Such reasoning can be explained only by a deep-seated per se policy against the enforcement of conditions that create a forfeiture ex post.[260] The difficulty is that commercial parties have sound reasons for creating express condi-

---

[259] *Id.* § 229 cmt. b, illus. 1.

[260] This point is underscored by an analysis of the other illustrations in Restatement § 229. For example, consider the second illustration of § 229, based on *Del. Steel Co. v. Clamar S.S. Corp*., 378 F.2d 386 (3d Cir. 1967). RESTATEMENT (SECOND) OF CONTRACTS § 229 cmt. b, illus. 2 & reporter's note (1981). This illustration, in which a carrier receives only oral notice of damaged cargo when the contract requires written notice, implies that express conditions should be subject to a "no prejudice" standard: They should not be enforced if their purpose has been served by other means. *Id.* § 229 cmt. b, illus. 2. But § 229 does not, in fact, permit courts to set conditions aside based on a purely ex post "no prejudice" standard. As we have seen, the doctrine requires that the condition not have been regarded as material by the parties at the time of formation. In this case, the parties had ample reason ex ante to regard the condition as a material component of their agreement. The third, fourth, and fifth illustrations of § 229 each constitute variations on the theme of excusing the timing of the occurrence of a condition. *Id.* § 229 cmt. c, illus. 3–5. The Restatement says that a court may excuse the nonoccurrence of a condition "during the period of time in which it would otherwise have to occur, if it concludes that the time of its occurrence is not a material part of the agreed exchange. This conclusion is sometimes summed up by the phrase that 'time is not of the essence.'" *Id.* § 229 cmt. c (1981) (citations omitted). Even if parties expressly condition an obligation on a timing requirement and expressly state that the timing requirement is of the essence, equitable doctrine permits a court to set aside the condition in order to prevent a forfeiture. *See* Holiday Inns of Am., Inc. v. Knight, 450 P.2d 42, 43–45 (Cal. 1969) (excusing failure to comply with timing requirement where contract expressly made timing requirement essential but enforcement

tions, and thus courts have good reasons to enforce them. Indeed, it is easy to imagine reasons why B might have placed a high value on the use of Reading brand pipe, even though he knew other brands of equivalent quality were available. We argued in Part III that parties faced with the challenge of specifying performance standards choose between precise, rule-like terms and vague, standard-like terms. In doing so, they trade off front-end specification costs against back-end verification or enforcement costs. Even if B's sole objective was to ensure the installation of Reading-quality pipe in his house, he might have intentionally conditioned his payment obligations on the installation of Reading *brand* pipe, rather than Reading *quality* pipe, in order to lower the expected costs of enforcing that requirement. A term requiring Reading quality pipe sets out a standard that places on B the burden of proving that the pipe installed by the builder does not conform with the Reading quality standard. A term requiring Reading brand pipe instead sets out a precise rule that allows B to verify performance or nonperformance at relatively low cost.[261] Sophisticated parties wishing to focus more of their contracting costs ex ante would thus be frustrated with regard to this goal by courts' invocation of the excuse of conditions.

E. A Final Equitable Override: Waiver of Conditions

The final equitable doctrine in direct tension with the formal doctrine governing the enforcement of express conditions allows courts to avoid enforcement of express conditions that courts have acknowledged are undeniably contained in the parties' agreement and are not subject to excuse. The law governing waiver of conditions nonetheless allows a court to set aside conditions by holding that the promisor has

---

of condition would have caused forfeiture); RESTATEMENT (SECOND) OF CONTRACTS § 229 cmt. c, illus. 4 (1981) (based on *Holiday Inns of America*).

[261] It is also possible that B had other, atypical reasons for preferring the brand. Perhaps he was skeptical of claims that other brands were "just as good." Or perhaps he had an emotional, reputational, or business reason for preferring Reading brand pipe over pipe of equivalent quality. To be sure, in the actual case, Kent's payment obligations were conditioned not just on the use of Reading brand pipe, but on Jacob & Youngs' exact conformity with *all* the specifications in the building plans. *See* SCOTT & KRAUS, *supra* note 3, at 71 (excerpting original contract language). That fact makes both of the above accounts questionable. Indeed, because the condition applied to all specifications rather than just a few, Jacob & Youngs might plausibly be seen as having actually enforced the condition as written but as having interpreted the express condition to require the installation of Reading quality, rather than Reading brand, pipe. So interpreted, Jacob & Youngs did not breach and actually satisfied the condition. The facts in the illustration, however, state only that B's payment obligation is conditioned on the use of Reading brand pipe. On the facts so described, there is no justification either for interpreting the condition to require installation of Reading quality pipe or for setting aside the condition.

waived the condition. A promisor can waive a condition explicitly by promising that he will not enforce it, even if he receives no consideration in return. Or a promisor can waive a condition implicitly by failing to object if the promisee fails to satisfy it or by performing despite the nonoccurrence of the condition.[262]

The doctrine of waiver illustrates how equitable reasoning can undermine judicial respect for express terms that clearly indicate the parties' determination to avoid the application of equitable doctrines to their agreement. *Fritts v. Cloud Oak Flooring Co*.[263] provides a paradigmatic example. In *Fritts*, a commercial landlord and commercial tenant entered into a commercial lease agreement that granted the landlord the right to terminate the lease after ten days following the tenant's default and the landlord's written notice of default.[264] The lease also expressly provided an antiwaiver clause and a time-is-of-the-essence clause.[265] The tenant originally fell behind in rent in late 1968 but subsequently arranged for payment of arrears.[266] For the next year, the tenant paid in full each month but on ten occasions failed to pay by the first of the month, usually paying a few days there-

---

[262] *See* RESTATEMENT (SECOND) OF CONTRACTS § 84 cmts. a & b (1981) (discussing waiver rationale for rule that promise to perform despite nonoccurrence of condition is generally binding); 2 FARNSWORTH, *supra* note 18, § 8.5 (describing excuse of condition by waiver).

[263] 478 S.W.2d 8 (Mo. Ct. App. 1972). Waiver by election bars the promisor from defending its own failure to perform on the ground that the promisee had previously failed to satisfy a condition of the promisor's prior performance. SCOTT & KRAUS, *supra* note 3, at 656. Waiver by estoppel bars the promisor from refusing to perform on the ground that the promisor had, implicitly or explicitly, waived not only in the past but in the future as well. *Id.*

[264] 478 S.W.2d at 9–10. The court described the commercial parties to this dispute as follows:

> [P]laintiff L. C. Fritts d/b/a B & C Leasing Service (the landlord) sought restitution of certain leased premises, to wit, a tract in Springfield and the business building situate thereon, together with double the monthly rents and profits of said premises, because (so it was alleged in plaintiff's petition filed on January 9, 1970) "the lease . . . was terminated at January 1, 1970, for failure of defendant [Cloud Oak Flooring Company, a corporation, the lessee-tenant] to pay the rental reserved." Plaintiff-landlord appeals from the judgment for defendant-tenant.

Id. at 9 (citation omitted).

[265] According to the court,

> *Paragraph 25* [of the contract] stated that no failure of the landlord to exercise any power given him under the lease, or to insist upon strict compliance by the tenant with its obligations thereunder, and no custom or practice of the parties at variance with the terms of the lease should constitute a waiver of the landlord's right to demand exact compliance with the terms thereof. [*P*]*aragraph 26* [of the contract] declared that "[t]ime is of the essence of this agreement."

Id. at 10.

[266] *Id.* at 10.

after.[267] The landlord accepted these payments without objection until January 1970, when he returned the tenant's check dated January 5, 1970, and instituted an unlawful detainer action.[268]

The court found that the landlord had accepted and cashed "ten consecutive tardy monthly rental payments . . . [without ever warning] the tenant of his intention to seek a forfeiture of the leasehold estate for any future tardiness in payment."[269] The court then held that, notwithstanding the time-is-of-the-essence and antiwaiver clauses, this conduct constituted a waiver of the landlord's right to terminate upon late payment without first providing advance notice of his intention to strictly enforce the payment deadline. The court stated that "[b]y the mere act of including an essence provision . . . the landlord did not immunize or insulate himself from the legal effect and consequence [of cashing the tardy checks]."[270] It then quoted *Corbin* for the proposition that

> a provision that an express condition of a promise or promises in the contract cannot be eliminated by waiver, or by conduct constituting an estoppel, is wholly ineffective. The promisor still has the power to waive the condition, or by his conduct to estop himself from insisting upon it, to the same extent that he would have had this power if there had been no such provision.[271]

Fritts is a dramatic illustration of the pervasive effect of the law governing waiver of conditions. The law of waiver permits courts to ignore the parties' expressed intentions *even when* parties draft meta-terms providing clear instructions to courts about how they intend the express terms in their agreement to be interpreted. Application of the waiver doctrine in the face of such meta-terms undermines the ability of parties to control the interpretation, and thus the design, of their agreements.[272] To see why, consider the late rental payments in *Fritts.* Suppose that the parties expect that late payments will sometimes occur under conditions where the expected value of the loss to the landlord is low. At other times, a late payment can impose a substantial expected loss. Ideally, the parties would condition the landlord's

---

[267] *Id.* at 10–11.

[268] *Id.* at 11.

[269] *Id.* at 13.

[270] *Id.*

[271] *Id.* at 14 (quoting 3A ARTHUR LINTON CORBIN, CORBIN ON CONTRACTS: A COMPREHENSIVE TREATISE ON THE WORKING RULES OF CONTRACT LAW § 763, at 531 (1960)).

[272] *See* Pa. Ave. Dev. Corp. v. One Parcel of Land, 670 F.2d 289, 292–94 (D.C. Cir. 1981) (granting commercial tenant larger share of condemnation award so as to prevent forfeiture, notwithstanding explicit terms of condemnation clause). For discussion, see Victor P. Goldberg, Thomas H. Merrill & Daniel Unumb, *Bargaining in the Shadow of Eminent Domain: Valuing and Apportioning Condemnation Awards Between Landlord and Tenant*, 34 UCLA L. Rev. 1083, 1120–25 (1987).

right to terminate on a late payment that causes the landlord to suffer a large loss. However, the size of the loss caused by any given late payment is likely to be observable but not verifiable—the expected costs to the landlord of proving the loss in court are likely to be greater than the expected benefits of litigation. Under these circumstances, parties reasonably might agree to give the landlord an unconditional right to terminate, subject to the legally unenforceable (but relationally enforceable) understanding that the landlord will not exercise this right unless the expected loss caused by the late payment is large.[273] But given the existence of the equitable doctrines governing waiver of conditions, the only way for the parties to achieve this result is to indicate expressly their desire to prevent these doctrines from being applied to their agreement. This is the *only* possible purpose to be served by including the "time-is-of-the-essence" and "antiwaiver" clauses in an agreement. Thus, when sophisticated commercial parties include these express clauses in their agreements, there is simply no justification for the judicial practice of refusing to enforce them on equitable grounds.

In short, when sophisticated commercial parties incur costs to cast obligations expressly in a written and unconditional form, they do so to permit a party to stand on its rights under the written contract when doing so is necessary to protect the party from incurring a substantial loss. But sometimes the only way to make such a right practically effective is to make it unconditionally enforceable. Thus, when sophisticated commercial parties create express conditions and then include an express clause unmistakably prohibiting waiver of those conditions by subsequent conduct, short of modification, there are good reasons for courts to enforce those conditions strictly. The theory of contract design therefore teaches that ex ante efficient contracts may sometimes result in forfeiture ex post. To abhor a forfeiture is to abhor the formal contract doctrines that make it possible for commercial parties to design their contracts efficiently. By seeking to redistribute losses ex post, the antiforfeiture doctrines not only under-

---

[273] The argument here parallels our argument above that, in *Hunt Foods*, Doliner might rationally have agreed to give Hunt Foods an unconditional option subject to the legally unenforceable condition that Hunt Foods will not exercise its option unless Doliner shops its bid. *See supra* Part II.B. Thus, the landlord's commitment is relationally enforceable because typically he will be a repeat player with an interest in preserving his reputation and social esteem. In addition, as we explained above, "[e]xperimental evidence shows that a preference for reciprocity—the willingness to reward cooperation and to punish selfishness—can motivate cooperation even in arms-length interactions between complete strangers." *See supra* note 133. Finally, because commercial real estate attorneys who represent landlords are repeat players in their locality, they can bond their clients' nonlegal commitments with their reputations.

mine efficient contract design but may also unfairly deprive one party of gains that it paid to retain in order to compensate another party for losses that it was paid to bear.

Conclusion

The current two-stage regime of contract adjudication treats formal contract terms as prima facie instructions that can be set aside when necessary to vindicate the parties' contractual ends. In this Article, we have sought to demonstrate why sophisticated commercial parties would instead prefer a regime that strictly enforces formal contract terms absent an express invitation for judicial intervention. Our argument is premised on the fundamental distinction between the parties' intended contractual means and their intended contractual ends. Contractual means are the mechanisms by which commercial parties create contracts designed to best achieve their intended ends. Preservation of the efficacy of these means requires a regime in which courts adhere to the formal doctrines governing contractual interpretation and resist the temptation to vindicate contractual ends irrespective of the parties' chosen contractual means. However well-intentioned the instinct to soften the hard edge of formal legal rules, it is ultimately self-defeating: This instinct may redistribute risks the parties were paid to bear, and, in any event, it undermines efficient contract design.

Contracts that otherwise might seem inexplicable can be rationalized in light of the principal tradeoffs parties face in designing contracts efficiently. Because commercial parties often prefer to condition their performance obligations on events that are observable but not easily verifiable, in many cases they will rely on both legal and relational enforcement to regulate different obligations arising out of a single agreement. By using express terms and invoking interpretive rules such as the parol evidence rule and the doctrine of integration, parties can partition their agreement into legal and relational segments. When they do so, they knowingly expose themselves to the risk that they will have no legal recourse should their counterparty breach an obligation governed only by the relational norms of their agreement. With respect to the legally enforceable terms of the agreement, contract design theory explains that parties face the further choice between using their superior knowledge of their contractual ends to specify precisely the terms of the contract ex ante or to delegate to a court with the benefit of hindsight the task of selecting an appropriate proxy for the contract performance ex post. The former strategy motivates parties to use express terms with a plain meaning to

create contractual rules that confine the discretion of a court in subsequent litigation.

The challenges of contract design thus require parties to use various combinations of legal and relational norms, as well as contractual rules and standards, to optimize their contract in light of uncertainty about the future and the difficulties of proof. But they can do this efficiently only if courts reliably apply the formal interpretive doctrines of contract law. Since design choices are best made by the parties ex ante, our theory holds that commercial parties would prefer a regime in which the second stage of the two-stage adjudicatory system—an inheritance of the centuries-old conflict between law and equity—is invoked only if specifically requested.

Our argument has been largely based on theory, but there is growing evidence to support the claim that commercial parties in fact exhibit the preferences that theory predicts. For example, a common provision in many alliance agreements specifies the parties' express preference that courts resolve any disputes by relying exclusively on formal interpretive rules.[274] In addition to anecdotal evidence from individual contracts, Lisa Bernstein has shown that parties who are members of trade associations—and thus who rely on both relational enforcement and third party enforcement—carefully preserve formal contract doctrine and reject equitable principles in assessing performance, breach, and liability.[275] Bernstein argues that this single-stage regime can best be understood as a mechanism for preserving the space for both formal and relational norms to operate. Finally, as we have suggested above, recent work by Eisenberg and Miller that showed strong party preferences for selecting New York over California for both choice of law and choice of forum clauses provides

---

274 *See, e.g.*, E. I. du Pont de Nemours & Co. and EarthShell Corp., Alliance Agreement art. 12(h), at 7 (July 25, 2002), http://contracts.onecle.com/earthshell/dupont.collab.2002.07.25.shtml. The Agreement reads:

> The Parties' legal obligations under this Alliance Agreement are to be determined from the precise and literal language of this Alliance Agreement and not from the imposition of state laws attempting to impose additional duties of good faith, fair dealing or fiduciary obligations that were not the express basis of the bargain at the time this Agreement was made.
>
> The Parties are sophisticated business entities with legal counsel that have been retained to review the terms of this Alliance Agreement and the Parties represent that they have fully read this Alliance Agreement, and understand and accept its terms.

Id.

275 Bernstein, *Private Commercial Law*, *supra* note 130, at 1735–37 (describing formalistic approach adopted by cotton industry arbitration tribunals); *see generally* Bernstein, *Merchant Law*, *supra* note 130 (describing formalistic adjudicative methods of National Grain and Feed Association arbitrators and why merchants find this approach to adjudication preferable).

further support for the claim that commercial parties prefer interpretations based on the formal doctrines of the common law.[276] The significance of this striking differential in party preference is shown by Miller's analysis of the differences in contract law between New York and California: New York uses formal doctrine to enforce bargains strictly and displays little tolerance of equitable principles that seek to balance interests ex post;[277] California, by contrast, is far more willing to revise contracts ex post on the grounds of fairness, equity, or public policy. Miller concludes that "[t]he revealed preferences of sophisticated parties support arguments by Schwartz, Scott[,] and others that formalistic rules offer superior value for the interpretation and enforcement of commercial contracts."[278]

The cornerstone of contractual interpretation is respect for the parties' contractual intent. But judicial respect for the parties' contractual intent entails fidelity to the doctrines on which parties must rely to select the contractual means for pursuing their contractual ends. Under the current two-stage regime, courts presume that the parties intended them to realign their contract terms with their contractual ends whenever the parties' formal contract terms either never did, or no longer do, constitute a rational means of pursuing those ends. Indeed, as the law of conditions dramatically illustrates, courts sometimes treat this presumption as conclusive by refusing to give effect to express contractual language directing courts not to apply equitable doctrine to their agreement. We have argued, however, that, at least between sophisticated commercial parties, this presumption is rarely, if ever, justified. Although commercial parties might sometimes find it rational to delegate such discretion to courts, they would prefer an adjudicative regime that gives them the option of using contract design strategies that are effective only if courts will reliably enforce the contract's formal terms, even when such enforcement produces a result radically misaligned with the parties' contractual ends. Such an adjudicative regime requires courts to reject the

---

[276] *See* Eisenberg & Miller, *supra* note 148, at 19, 34 (discussing study showing over 40% of parties in large data base preferred New York for choice of law and choice of forum clauses as against only 8% and 7% selecting California for choice of law and choice of forum clauses, respectively).

[277] Miller, *supra* note 146, at 5. In addition to identifying the preference in New York for strict interpretation of plain language in commercial contracts, a preference for a "hard" parol evidence rule, and reluctance to grant relief on grounds of mistake or excuse, Miller reports that "the trend of recent New York cases has been to enforce forfeitures in the absence of some other grounds for invalidity, such as unconscionability." *Id.* at 25 n.139 (citing Fifty States Mgmt. Corp. v. Pioneer Auto Parks, Inc., 389 N.E.2d 113, 115 (N.Y. 1979) (enforcing provision in lease contract providing for forfeiture by tenant of all possessory rights upon failure to tender rent payments for two months)).

[278] *Id.* at 1.

hoary principle that equity trumps formal doctrine and instead equate the parties' contractual intent with their chosen contractual means. Nothing less will satisfy contract law's foundational commitment to honoring contractual intent.