

Conceptualizing Contractual Interpretation

Alan Schwartz¹ and Joel Watson²

Abstract

Contract interpretation is an important subject for lawyers but there are few rigorous economic analyses. We use a formal model to derive the interpretive rule that an “ideal” legal enforcer would adopt. Commercial parties make contracts to maximize gains from trade. Our enforcer is ideal, in the sense that he interprets contracts to maximize welfare over the set of contracting relationships. We then compare how actual interpretation matches the ideal. We derive the following results, among others: (a) An optimal interpretive rule induces parties to make efficient tradeoffs between the gain from accurate interpretation and the costs of (i) writing contracts; (ii) investing in the deal, which also conveys intent; and (iii) trials; (b) The optimal rule induces inefficient relationships not to form and permits the enforcer sometimes to exclude trial evidence and decide on reduced evidentiary bases; (c) Expertise in interpretation comes in several forms, including the ability to infer party intent from a performance and from the written words; (d) Parties choose between arbitrators and courts on the bases of their respective competencies and the cost of writing a contract meant for a court and the cost of writing a contract meant for an arbitrator; (e) Real courts maximize accuracy in interpretation rather than welfare, which creates a number of inefficiencies; (f) Party attempts to constrain the evidence courts use would cause courts to act more as the ideal enforcer acts, so courts should obey party interpretive instructions. We also test several of the model’s predictions on a data set of almost 43,000 actual contracts. The data appear to be consistent with the theory.

¹Yale University Law School and Management School.

²UCSD Economics. We are grateful for comments received at a Columbia Law School Contract Theory Workshop and the NBER Summer Law and Economics Workshop.

1. Introduction

Contract interpretation has been an important subject in the law school world for decades. There have been few law and economics analyses and very few formal economic treatments, however.³ This paper does three things. First, it creates a model to show how contract interpretation works when an ideal “enforcer” does it.⁴ The enforcer is ideal in two senses: (i) Motivation: The enforcer’s goal is to maximize welfare: the expected value of commercial transactions over the set of contractual relationships; (ii) Expertise: The enforcer can make the cost/benefit calculations that this maximization exercise requires. In the model, a population of sellers make private sunk cost investments that stochastically increase the buyers’ values for a product or a service. Hence, our ideal enforcer interprets contracts to induce sellers to invest efficiently. We analyze the model to see how ideal contractual interpretation works as premise for understanding parties’ and decision makers’ actual interpretive practices and for making policy suggestions as to how to improve those practices.

Regarding the interpretive task itself, to induce efficient investment the probability that the seller is paid must increase in the probability that she has rendered a compliant performance. The enforcer observes the actual performance; the interpretive issue is whether the parties ex ante intended the performance he sees. Parties contract to maximize expected surplus so they prefer accurate interpretations. The task for parties, then, is to make the correct tradeoff between

³The law and economics literature is summarized in Schwartz and Scott (2010) and (2003) and discussed briefly below. A recent survey of the economics literature, Zhao (2010), discussed only two papers extensively.

⁴The enforcer can be thought of as a court; we use the more neutral term because later we ask when parties would prefer a judicial enforcer or an arbitrator. The enforcer can develop his interpretive rule through a series of decisions, as a court would, or the legislature could enact the rule. See UCC §2-202. The important feature is that potential parties know the rule when choosing a contract.

the gain in investment incentives from a more accurate interpretation and three sets of costs: (i) the cost of writing the contract, which conveys information about intent; (ii) the cost of investing in performance, which (a) increases its value, and (b) increases the likelihood that the seller will be found to comply;⁵ and (iii) the cost of a possible trial, which also reveals intent. An optimal interpretive rule induces parties to make these tradeoffs in a socially welfare maximizing way. In addition, parties can choose the enforcement institution (e.g., court or arbitrator) that will later interpret their agreements. Understanding interpretation requires, at least, an understanding of how an ideal enforcer (or an ideal set of interpretive rules) would affect parties' cost/benefit tradeoffs and institutional choices.

Our second task is to use this understanding to analyze some significant interpretive practices. As examples, there is a strong, though not complete, congruence between the ideal enforcer's goal and the parties' goals: the enforcer and the parties both prefer optimal tradeoffs on the three relevant margins. This implies that parties seldom would want to tell the enforcer how to interpret the contract and utilize evidence. If actual parties often attempt to tell actual enforcers what to do, then real world interpretation may deviate in material ways from the ideal. As a second example, enforcement institutions differ in the nature of the expertise they bring to bear. We analyze the nature of interpretive expertise to better understand how parties choose interpretive institutions.

Our third goal is to provide a basis for making interpretive policy. To illustrate, let courts and parties actually differ regarding how much and what types of evidence courts should consider when making interpretations. An analysis of ideal interpretation may illuminate whose

⁵An actual performance is evidence of the intended performance.

preferences should prevail.

We generate the following results:

A. Even an ideal enforcer cannot induce first best investment incentives. First best requires absolutely correct interpretations, but the costs of reducing uncertainty regarding the parties' ex ante intentions to zero exceed the incentive gains. Further, the effect of residual uncertainty cannot be eliminated by punishing probabilistically noncompliant performances, if penalties are not permitted as aids to contract enforcement.

B. An optimal interpretive rule induces parties to choose efficient contracts. If, say, a particular set of contracting relationships can best indicate their intentions by writing extensive contracts, the rule will induce them to write these contracts rather than write simple contracts and rely on trials to reveal intent. Along this line, an optimal interpretive rule correctly induces some potential relationships not to form. The cost of conveying intent, for them, would exceed the investment efficiency gain.

C. Expertise in interpretation comes in three forms: (i) expertise in extracting meaning from words; (ii) expertise in inferring party intentions by applying industry knowledge to evaluate a seller's performance; and (iii) expertise in evaluating trial evidence.⁶

D. Institutional changes that lower contract-writing costs and changes that increase expertise in evaluating a tendered performance are complements; on the other hand, changes that lower contract-writing costs and trials are substitutes. Regarding complements, a more extensive contract description better permits an enforcer to match the actual performance to the intended

6. Enforcers see similar evidence before and at trials, but evidence is evaluated differently at these litigation stages. For example, evidence at a court trial is presented through a prism of evidentiary rules that it takes an expert adjudicator to apply. Thus a court may be better than an arbitrator at evaluating trial evidence but less good at just looking at a performance and inferring correctly whether the performance was intended.

performance. Regarding substitutes, an extensive trial can make up for a sketchy contract. Furthermore, under some conditions, the enforcer optimally restricts trial evidence.

E. In the empirical part of the paper, we use the model to formulate and test three predictions: (i) There is an inverse relation between contract length and the number of contracting relationships that choose a contract of that length; (ii) There is a positive relation between contract complexity, proxied by length, and the selection of an (expert) arbitrator for dispute resolution; and (iii) There are, over a wide range, more extensive context descriptions in contracts intended for arbitration. The empirical tests, which are preliminary, are primarily done to show that a contract theoretic treatment of interpretation can generate testable empirical predictions. The data offer tentative support for the predictions, which suggests that our model may capture some aspects of actual behavior.

F. When the enforcer maximizes accuracy in interpretation rather than welfare over the population of contractual relationships, (i) welfare is lower; (ii) parties' contracting choices are biased, often unexpectedly in the direction of writing more complicated contracts; and (iii) some contracting relationships suboptimally fail to form.

G. Many courts do attempt to maximize accuracy⁷ and actual parties commonly send interpretive instructions to them.⁸ These instructions almost invariably attempt to reduce the probability and cost of trials.⁹ Party instructions, if followed, thus would make courts function

⁷See Posner (2004).

⁸Instructions are included in the parties' contract.

⁹For example, an instruction may direct the court not to consider oral statements the parties made to each other during a negotiation. Claims regarding these statements often are contested, so making statement evidence admissible increases the likelihood of a trial.

more as the ideal enforcer functions; hence, courts should obey them. Many courts today do not.

We conclude this Introduction with two remarks and a brief description of the literature. First, our analysis is preliminary. Interpretation is understudied formally, so we make an early cut at a complex subject.

Second, the subject is important. Contract theory models exclude interpretive issues by assumption. In these models, the parties either describe the subject of trade perfectly or do not describe it at all. Perfect descriptions need not be interpreted and null descriptions cannot be interpreted. In reality, contracts describe the subject of sale more or less effectively. The typical issue that actual cases present, therefore, is not whether a particular payoff relevant variable is verifiable or not. The issue rather is what parties intended that variable to be, which is an interpretive question. Richard Posner thus estimated that 80% of the contract cases that come before the Seventh Circuit involve interpretive issues.¹⁰

Our paper utilizes the same foundation as does Shavell (2006), whereby (i) there is a population of contractual relationships with different “types,” (ii) contracting is costly, and (iii) contract selection serves to partition the space of relationship types and thus provides coarse information about type to the interpreter.¹¹ We are also similarly focused on optimal interpretation to maximize social welfare. However, there are major differences between Shavell (2006) and this paper. First, there is no investment in Shavell’s model, so the interpreter there is concerned only with ex post optimality. In contrast, we analyze the choice of interpretive practices to induce efficient investment, where there are trade-offs between costs realized at

¹⁰Judge Posner’s views regarding interpretation are set out in Posner (2004).

¹¹The second feature follows the prior work on costly contracting and limits on describability (Dye 1985, Anderlini and Felli 1999, Battigalli and Maggi 2002, Schwartz and Watson 2004).

different stages of productive relationships. Second, Shavell assumes that all terms have an exogenously defined “literal” meaning, and his results focus on demonstrating that it is sometimes optimal for the enforcer to depart from literal interpretation.¹² We argue, in contrast, that context commonly is indispensable to interpretation when the interpreter’s task is to induce efficient party behavior, and we determine the optimal interpretive rule absent exogenous constraints on the meaning of contract terms. This allows us to explore the implications of court bias toward accuracy. It also allows us to provide results on trade-offs in various dimensions of evidence and expertise, and on the optimal use of multiple enforcers.

Listokin (2010) considers a problem related to ours. He shows how a court that wants to maximize accuracy in interpretation would use Bayesian analysis to interpret contract terms. The enforcer in our model also does Bayesian updating, but our model differs in at least three significant respects from Listokin. First, our enforcer maximizes welfare, not accuracy. Second, our enforcer interprets contracts in the service of inducing investment; there is no investment in Listokin’s model. Third, a court cannot be a good Bayesian unless it has an informed prior. Listokin assumes that courts have such priors. In our model, the informativeness of the enforcer’s prior is a facet of his expertise. This permits us to model the parties’ choice of an enforcer.

Finally, Kvaloy and Olsen (2009) analyze a model with investment where parties can affect the verifiability of payoff relevant variables by the contracts they write and the investments themselves. This paper is illuminating regarding contract choice, but it assumes that values are deterministic and, importantly, that litigation is costless. The last assumption is limiting because

¹²Shavell’s interpreter generally follows, though he sometimes overrides, “specific” terms and fills gaps in contracts with efficient terms. Maggi and Staiger (2008), in the context of international trade agreements, also assume that aspects of the enforcement regime are exogenously fixed.

the common question for parties, and for our enforcer, is when is it efficient to reveal their initial intentions in a costly contract or in a costly trial, or in some combination of these.

Part 2 below sets out the model, which Part 3 solves. Part 4 presents our preliminary empirical test. Part 5 concludes.

2. The Model

2.1 A technical description

There is a population of contractual relationships, each one a match between a buyer and a seller. Relationships are differentiated by a *type* parameter t , that summarizes what the parties agree to trade and other payoff relevant aspects of their relationship. The type space is drawn from a uniform distribution over the interval $[0, 1]$.¹³ Disputes between the parties are adjudicated by an external “enforcer” who has three tasks: (a) to develop an “interpretive rule” that governs how he identifies a relationship’s type; (b) to make a finding as to type in contested cases; and (c) to compel monetary transfers between the parties (or not) as a function of his type decision. The enforcer’s goal, in performing these tasks, is to maximize aggregate welfare over the population of contractual relationships.

Interaction among the three agents in the model — buyer, seller and enforcer — is as follows:

Date 1. The enforcer creates and publishes his interpretive rule.

Date 2. A buyer and seller agree to trade a product or a service. They then choose a controlling contract that contains a price, normalized to 1, and a set of nonprice terms that, more

¹³This is a continuum, so there is an infinite number of possible types.

or less, describe the subject of sale and other governing conditions. The nonprice terms are indexed by the integer k , where $k \in \mathbf{P} = 0, 1, 2, 3, \dots$. The parties jointly pay the contracting cost $y_k > 0$. Contracts are ordered so that y_k is increasing in k .¹⁴ We assume that $y_0 = 0$ and for every κ there is a k such that $y_k > \kappa$. This assumption implies that contract costs may exceed possible trading gains for some potential contractual relationships.

Date 3. The seller privately chooses an investment level $q \in [0, 1]$ at cost $c(q)$. The probability that the seller's performance is compliant — i.e., it is the performance a type t relationship intends — is increasing in q .

Date 4. The seller tenders performance, which the buyer observes. With probability q , the seller's performance is compliant. With complementary probability $1 - q$, her performance is noncompliant. The buyer realizes a benefit of 1 if performance is compliant and zero if not.

Date 5. The buyer either accepts performance and pays the contract price, in which case the game ends, or the buyer rejects. A dispute follows rejection.¹⁵

Date 6. The enforcer observes information that the litigation yields, *except* evidence that only a trial would reveal. Pre-trial evidence includes the contract k and a signal, denoted x , that the seller's physical performance sends.¹⁶ A compliant performance commonly is probative

¹⁴Note that $k = 0$ represents “no contract”: a potential buyer and seller reject a relationship.

¹⁵The seller can sue the buyer for the price or the buyer can sue the seller for damages.

¹⁶The common distinction in the legal world is between “intrinsic” and “extrinsic” evidence. Intrinsic evidence corresponds to what we mean by pre-trial evidence: the parties' litigation narratives; the contract; and the performance, all of them evaluated with whatever expertise the enforcer possesses. Extrinsic evidence is what we mean by trial evidence: the parties' practice under prior contracts; their practice under the existing contract; evidence of precontractual negotiations; precontractual memoranda and preliminary contract drafts; and industry custom. Some extrinsic evidence may be revealed in discovery proceedings, but evidence in these categories is usually contested and so settled at trial. For example, the buyer may claim that the seller's performance was noncompliant, as measured by the custom in the parties' industry. The seller may deny the existence of a custom or claim that she satisfied it. We refer to extrinsic evidence as trial evidence because its weight and materiality usually is settled at

regarding the parties' type t — i.e., what they agreed to trade. Consequently, we suppose that if the performance is compliant then x equals the relationship's type t with probability $s > 0$, and with complementary probability x is drawn from a uniform distribution over the entire type interval $[0, 1]$. If the performance is noncompliant, x is drawn from the uniform distribution for sure.¹⁷ The three agents in the model observe the contract and the performance signal x .

Date 7. The enforcer forms a belief about the parties' type t and whether performance was compliant on the basis of the litigation documents, the contract, and the performance signal x . He then orders the buyer to pay, or not. If the enforcer orders payment, then he also decides whether to permit trial evidence if the buyer offers it.¹⁸

Date 8. The buyer decides whether to submit trial evidence, at cost $\gamma \geq 0$, if the enforcer permits it. We assume for convenience that trial evidence perfectly reveals the type t . The enforcer, at this stage, then makes a final determination of whether to order the buyer to pay, either on the basis of all evidence or with pre-trial evidence only.¹⁹

Welfare is maximized when the seller chooses q to solve $\max_q q - c(q)$. The cost function c is twice continuously differentiable, with the standard properties $c' > 0$, $c'' > 0$, $c''' > 0$, and $\lim_{q \rightarrow 1} c(q) = \infty$. Let q^* denote the solution to this maximization problem, so that $c'(q^*) = 1$.

trial. For a further discussion see Schwartz and Scott (2003) and authorities cited therein.

¹⁷The signal is unilluminating in this event because the type space is uniform.

¹⁸The enforcer's interpretive rule governs when trial evidence is permitted.

¹⁹Trial evidence provision is inefficient ex post because at stake then is only whether the buyer should make a transfer or not. As we show below, an interpretive rule that creates efficient investment incentives makes the correct decision as to whether to permit evidence.

2.2 Explanation and additional notation

The seller complies by tendering a performance that parties of type t would want, so it is helpful, in understanding how interpretation functions, to clarify what we mean by a type. We proceed by example. Let there be three party types, all of whom trade raw cotton to be made into cloth. There are seven cotton grades, ranging from grade one, which is the best — Egyptian superfine — to grade seven, which is the worst — Bulgarian. A type t_1 relationship trades only grade one cotton, because the buyer produces high quality fabric for clothes sold in boutique stores. In contrast, a delivery of grades two or three cotton would be compliant for a type t_2 relationship; this buyer produces fabric that is used for moderately priced clothing. Grade one cotton would make the buyer's cloth too costly while grades four through seven would yield cloth of insufficient quality. For a type t_3 relationship, however, only grade seven cotton is compliant. The buyer produces work clothes and so requires only coarse, inexpensive cotton.

Now let the enforcer observe that the seller tendered grade four cotton. The enforcer would not order the buyer to pay unless he formed a mistaken belief, for a tender of grade four cotton would be noncompliant for any of the three relationship types. If the enforcer instead observed a tender of grade three cotton, he should order the buyer to pay *only if* he believed that the parties before him constituted a type t_2 relationship. A buyer party to either a type t_1 or a t_3 relationship would have an incentive to contest with trial evidence an order to pay for grade three cotton.

The enforcer's adjudicatory task, when there is a dispute, is to locate the parties on the $[0,1]$ type interval. He has recourse to four evidentiary sources: (a) the parties' litigation documents, which commonly provide narratives of the deal and its breakdown; (b) the contract;

(c) the seller's physical performance, which generates the signal x ; and (d) trial evidence (testimony, documents) if the enforcer permits it. The enforcer may observe some of these categories at approximately the same time — the contract may be attached to the seller's complaint — but it is convenient to suppose that the enforcer observes these categories in the order set out here.

Returning to the cotton example, the enforcer learns from the litigation documents that he should search only over cotton trading types. Formally, the parties' type t is drawn from an interval that includes possible cotton traders. An expert enforcer may narrow this interval further. Thus, the narrative may permit him to exclude types that trade cotton blankets. The enforcer next observes the contract. These sometimes detail the parties' circumstances and usually describe, more or less precisely, what the seller is to do.

The litigation narrative and the contract thus provide information about (i.e., are “signals” of) the parties' type. Evidence in these evidentiary categories permits the enforcer to reduce the type space to a subinterval of $[0,1]$ in which, he believes, the litigation parties' type likely lies. We let S denote the subinterval that evidentiary categories (a) and (b) permit the enforcer to identify. To be clear, the enforcer continues to believe that relationship types are uniformly distributed, but now only over the subinterval S .

The absolute number of types in S , denoted σ , also matters. The enforcer may locate the parties in the type space on the basis of their narratives and contract, but different types may use contracts that appear similar to the enforcer and so S could contain multiple types.²⁰ Such residual uncertainty is a real possibility when, as in our formal model, there is a continuum of

²⁰This is an implication of positive contracting costs, which may preclude perfect individuation.

types (our cotton trading example assumed a discrete, finite set of types for simplicity).

The enforcer next observes x — the seller’s physical performance. Suppose initially that x is in the previously formed S subinterval. Recalling that a compliant performance is probative of what a compliant performance is supposed to be, the enforcer thus puts positive probability mass on the type being equal to x exactly. To be precise, he believes:

$$\begin{aligned} \text{Prob}[x = t \mid x \in S] &= \text{Prob}[x \in S \mid x = t] \text{Prob}[x = t] / \text{Prob}[x \in S] \\ &= qs / [qs + q(1 - s)\sigma + (1 - q)\sigma] \\ &= qs / \sigma(1 - qs). \end{aligned}$$

The enforcer believes with complementary probability that the parties’ type is some other point in S , but he cannot put strictly positive probability on any such point because the S interval is a continuum.²¹

Suppose next that x is not in S . Then the enforcer’s posterior probability about t will not move: a noncompliant performance cannot indicate what a compliant performance was supposed to be. More precisely, when x is outside of S , the enforcer continues to believe that the parties’ type is in S , but because the distribution in S is uniform, he does not put positive probability on any particular type in this interval.

The enforcer thus conditions his judgment regarding whether the buyer is to pay or not on the set S and the signal x (which yields a posterior belief about the relationship’s type). As a

²¹The variable q appears in the enforcer’s Bayesian calculation though q is private information. The enforcer infers q from his prior information, and his inference turns out to be correct in equilibrium. To anticipate later results, the Bayes’ rule calculation indicates that the enforcer’s ability to identify type is increasing in s . Intuitively, the more “precise” the performance, the less likely the enforcer is to draw the wrong inference. If the contract says “Rembrandt” and the seller delivers a Rembrandt, the enforcer is likely to find that the seller complied. On the other hand, the enforcer’s ability to identify type is decreasing in σ , the number of types in S who appear similar on the basis of their contracts and narratives. If the contract only says “Rembrandt”, then x may equal s with a lower probability because there are different Rembrandts. Would the delivery of a self portrait be compliant? Or did the parties intend to trade a portrait of the artist’s wife?

consequence, the enforcer can create a wedge between the seller's expected payoff for a compliant performance and her payoff for a noncompliant performance. The existence of this wedge, in turn, affects the seller's choice of an investment level. We show below that the enforcer optimally orders the buyer to pay without trial evidence only when x is in S . This interpretive practice functions imperfectly, however, because the three signals on which the enforcer decides may not identify the true t with certainty. The enforcer is mistaken if x is in S but t does not equal x : that is, the signal misleadingly indicates a compliant performance. In this event, the enforcer orders the buyer to pay, which creates an incentive for the buyer to introduce trial evidence.

There are three investment margins in this model. First, the parties invest in contract writing; second, the seller invests in performance quality; third, the buyer invests in trial evidence if the enforcer permits it. Regarding the relationship among these margins, parties have an incentive ex ante to write illuminating contracts: the more "type" information the contract contains, the smaller is the enforcer's initial partition S and so the more illuminating the performance signal becomes. When S is narrow, a performance signal x can be in it only if the performance is highly likely to be compliant. Hence, the parties' ability to create efficient investment incentives — to ensure that sellers are rewarded only for compliance — is increasing in the contract's informativeness. The seller's investment in quality increases the likelihood that her performance is compliant. The enforcer, in turn, is more likely to infer from a compliant performance that the parties' true type requires that performance rather than any other performance. Hence, a seller's investment has two related effects: (a) it increases the likelihood that her performance is compliant; and (b) it increases the likelihood that her performance is

found to be compliant. Finally, the buyer has an incentive to introduce evidence when otherwise he would be required to pay for a noncompliant performance.

The parties thus must make three tradeoffs: (a) between the cost of writing a more extensive contract and the more accurate adjudications a more extensive contract permits; (b) between the seller's investment cost and the expected benefits of investment, as just described; and (c) between the cost of introducing trial evidence and the accuracy gain. Parties make all of these tradeoffs with knowledge of the enforcer's interpretive rule. Hence, if the rule is optimal, parties make optimal tradeoffs on these three margins.

Before solving the model, it will be helpful to have a glossary of notation.

t : A contractual relationship's type.

S : The interval on the type space in which the enforcer believes the parties' type falls after observing their litigation story and their contract.

σ : The number of types in an interval S (that is, the length of this interval).

q : The investment level the seller chooses and the probability that the seller tenders a compliant performance.

$c(q)$: The seller's cost of choosing investment level q .

x : The performance signal the enforcer infers from the seller's physical performance.

s : The probability that the performance signal x equals t conditional on a compliant performance.

γ : The buyer's cost of providing trial evidence.

We conclude the model description with three comments. First, holding party type constant, the parties' costs are decreasing over the range of the enforcer's expertise. The more

expert the enforcer is, the better able he is to locate the parties' type from their story; their contract; and trial evidence. Thus, stories, contracts and trials can be shorter when the enforcer is an expert. On the other hand, the enforcer's expertise declines in the uniqueness of the parties' type. The more individual a type is, the less helpful general industry expertise is in evaluating it. Hence, unique types may incur large contracting costs to convey type information. We explore these possibilities in the later empirical section.

Second, the usual legal interpretive enterprise focuses on a contract's words. There is controversy among lawyers, for example, regarding whether words should be given their dictionary meanings or read in light of applicable trade usages. We do not elide such issues, but our view of interpretation is slightly different. We begin with the basic premise that the state should *enforce* contracts to induce parties to make sunk cost investments in a contract's subject matter. It follows that adjudicators should *interpret* contracts to realize the investment-inducing goal. An ideal enforcer therefore should attempt to recover the parties' type, and then only reward a performance that the type he believes the parties before him to be would intend to trade. On this view of interpretation, the contract is an important part, but not the only part, of the record the enforcer uses. Another way to put this view is that a contract must be interpreted in light of *all* relevant signals: the one it sends, and the signals that the parties' story, the performance, and trial evidence also send.²²

²²Some contracts require less interpretation than others because some descriptions require less context to evaluate than others. For example, a contract requires the seller to deliver "200 grams of iodine". Using the vocabulary here, the interval S for the relationship that uses this contract is very narrow. A delivery of 200 grams of iodine sends the signal "x in S" while a delivery of 200 grams of sulfur sends the signal "x not in S". On the other hand, the contract description "a bottle of iodine" requires substantial interpretation; bottles come in many sizes. In our language, the type space for bottle traders is large so S can be big. We return to this issue below when we analyze procedural instructions. Some procedural instructions require little interpretation — "no depositions permitted" — while others may require a lot — "do not use extrinsic evidence if the contract is complete".

Third, although the formal model supposes that disputes are resolved at a trial, our results would also hold if there can be settlement. We can interpret γ as the buyer's costs of preparing and producing evidence at the pre-trial stage. That is, the buyer will have to engage in some manner of evidence production before settlement negotiation takes place. The outcome of settlement negotiation is influenced by the anticipated outcome of trial, which itself is influenced by the contract, performance, and evidence. So, the basic logic of the model holds if settlement is permitted.

3. Analysis and Results

The enforcer maximizes welfare across the population of contractual relationships. Thus, he wants parties to respond optimally regarding contract choice to the interpretive rule he creates. As a consequence, interpretation presents a single-person decision problem — a “planner’s problem” — that involves the simultaneous choice of interpretive rule and contract for each relationship type. The optimal interpretive rule — the solution — induces parties to sort themselves, through the contracts they choose, into subsets of the type space. To see the relevance of this process for interpretation, recall the example above. The enforcer can use the litigation documents to infer that the type space includes cloth traders but not blanket traders. When the enforcer observes a cloth trader relationship that uses contract k , he may infer that the parties’ type is in a smaller subset: traders of cloth to be used for fine clothing (as opposed, say, to traders of cloth to be used for work outfits).

3.1 When the enforcer should require payment

The enforcer’s substantive goal is to induce sellers to invest efficiently. This objective is

best served by paying the seller if and only if her performance is compliant. In this model, first best is achieved exactly if the seller expects to receive 1 (the buyer's value) when she complies. Anticipating this price, the seller maximizes $q - c(q)$ and selects the optimal investment level q^* .

An insight of the model is that even an optimal interpretive rule cannot induce parties to choose first-best investment levels. There are two reasons. First, the seller expects to receive less than 1 in return for compliance because the enforcer cannot identify a compliant performance with certainty. Positive contracting costs may preclude perfect individuation (the contract cannot identify type exactly) and the performance signal x correlates with type imperfectly.²³ The seller's marginal benefit from investment falls as she becomes less certain of payment for compliance. Second, the investment margin cannot be scaled up because the price is capped at 1. We assume a cap because one is the buyer's value, and penalties (contractual recoveries above the gain a party would realize from performance) are not enforced in modern states. Also, as a practical matter, parties seldom can pay arbitrarily large transfers.²⁴

Turning to when the enforcer should order the buyer to pay in this imperfect world, recall that the enforcer uses the litigation documents and the contract (which his rule may induce) to locate the parties in an interval S of the type space. The enforcer then observes the seller's performance. The signal it sends can yield two outcomes: the signal x is in S or it is not in S . In relative terms, the event $x \in S$ (where the performance signal is consistent with what the enforcer has already inferred about type) is "good news" about compliance. On the other hand, the event x

²³Recall from the model description that when the seller tenders a compliant performance, $x = t$ with probability s , but $0 < s < 1$.

²⁴The assumption that transfers are bounded is not critical to the ideas we explore. One can construct other versions of the model without a transfer constraint and with features similar to those discussed here, but these are more complex to analyze.

$\notin S$ is “bad news” (indicating noncompliance). An optimal interpretive rule thus maximizes the seller’s marginal benefit from investment if it directs the enforcer to order the buyer to pay if $x \in S$ and to allow the buyer to exit if $x \notin S$.

The interpretive rule also may allow trial evidence. We assume that evidence sometimes is cost justified: $\gamma < 1$. In the model, evidence is assumed to be perfectly revealing. Therefore, the buyer produces evidence if and only if the performance signal satisfies $x \in S$ (so that payment would be ordered in the absence of evidence) but $x \neq t$. The enforcer thus reverses his payment decision when he sees evidence.

3.2 When the enforcer should admit evidence

The enforcer’s rule should admit trial evidence when admission maximizes the value of relationships. We show here that admission is optimal when evidence costs are low; the performance signal x is relatively inaccurate (a compliant performance fails to indicate clearly which performance was intended); and/or a relatively large number of types use similar contracts.

To begin, when evidence is admissible the seller’s expected payoff from investing q is $qs - c(q)$. Because evidence is revealing, the seller only gets paid if she complies and the enforcer ultimately believes her to comply (that is, if $x = t$). She complies with probability q when she invests q and the enforcer believes her to comply with probability s when she actually complies. The seller’s optimal investment level, q^E , thus satisfies the first-order condition $s = c'(q^E)$, with E denoting the admission case.

The expected joint value of a relationship the enforcer knows to be in the set S , gross of contracting costs, is

$$v^E \equiv q^E - \gamma(1 - sq^E)\sigma - c(q^E).$$

The first term on the right-hand side is the value of the seller's performance.²⁵ The second term is the expected cost of introducing trial evidence: evidence cost γ times the probability that the buyer will want to introduce evidence. This in turn is the probability that (i) the enforcer observes $x \in S$, which is σ , times (ii) the probability that the signal is inaccurate — that x does not equal the true type t .²⁶ This second probability is $1 - sq^E$.²⁷ The third term is the seller's effort cost. Referring again to the seller's first order condition, because $s < 1$ (i.e., the performance signal is not perfectly accurate) the seller chooses a suboptimal investment level.²⁸ As a consequence, the value of a relationship, v^E , is increasing in q^E .

Suppose next that the enforcer's rule disallows trial evidence. The seller's expected payoff from choosing q is then

$$q[s + (1 - s)\sigma] + (1 - q)\sigma - c(q),$$

The first term characterizes the two cases in which the seller tenders a compliant performance (with probability q) and the enforcer orders payment: (i) the performance signal is accurate ($x = t$); and (ii) the performance signal is unilluminating (with probability $1 - s$) but by chance the signal falls in the interval S . The second term is the probability that performance is noncompliant but the enforcer fails to recognize noncompliance because the performance signal also is in S .

²⁵The probability of compliance is q when the seller invests q . The buyer realizes a value of one from a compliant performance. Hence, the expected gross value of the contract is q times one, or q^E , when the parties expect evidence to be admitted.

²⁶The relevant type space is normalized to be between 0 and 1. The set S the enforcer creates is a subinterval of this space. We denote the length of the S interval (that is, the mass of types in S) as σ . Because the type distribution is uniform, σ is the probability that the performance signal is consistent with compliance given the enforcer's belief that the parties' type t is in S . The enforcer, however, can be mistaken.

²⁷The seller complies and the performance signal accurately represents compliance with probability q^E times s .

²⁸Recall that the optimal investment level is one.

Evidence is disallowed here so the enforcer again orders the buyer to pay. The last term is the seller's cost.

The first-order condition for effort when trial evidence is barred is $s(1 - \sigma) = c'(q)$. Let q^N denote the solution. The expected joint value of a relationship in the set S , gross of contracting costs, then equals

$$v^N \equiv q^N - c(q^N).$$

A comparison of the two first-order conditions is illuminating.

$$\textit{Evidence:} \quad s = c'(q^E)$$

$$\textit{No Evidence:} \quad s(1 - \sigma) = c'(q^N)$$

These first-order conditions have three intuitive implications. First, the seller exerts less effort when evidence is excluded.²⁹ Second, the seller exerts less effort as s falls — that is, as the performance signal becomes less accurate. The less accurate the signal is, the less likely the seller is to be rewarded for a compliant performance so the lower is her reward for investing. Third, the seller exerts less effort as σ increases — as more types seem similar to the enforcer without trial evidence. Again, the seller is less likely to be rewarded when the enforcer is less able to distinguish among possible contracting types.³⁰

Our last step is to compare the aggregate value of relationship types, with and without evidence, for a given set S that choose the same contract. Initially comparing v^E and v^N , v^E is reduced by the expected value of introducing evidence. On the other hand, the seller works

²⁹Because $\sigma > 0$, the LHS in the evidence case is larger than the LHS in the no evidence case, implying that on the margin the seller chooses a higher investment level when she anticipates that evidence will later be admitted.

³⁰When the number of types is large, the true t can be almost anywhere. Note that σ appears only in the no evidence first-order condition because evidence is perfectly revealing. Unsurprisingly, q^N converges to q^E as σ goes to zero.

harder when she anticipates that evidence will be admitted.³¹ The enforcer's accuracy is increasing in the seller's effort. The enforcer's rule thus should trade off saving trial costs against the fall in effort incentives that less accuracy yields.

To see how the enforcer proceeds, it is convenient to write the value expressions as functions of σ , the length of the S interval, because the effort levels q^E and q^N depend on σ , and these effort levels in turn affect accuracy. Welfare gross of contracting costs as a function of σ is

$$w(\sigma) \equiv \sigma \max\{v^E(\sigma), v^N(\sigma)\}.$$

There is a number $\underline{\sigma} > 0$ such that welfare is increasing for $\sigma < \underline{\sigma}$, and is decreasing for $\sigma > \underline{\sigma}$. This is because it is suboptimal for a large number of relationship types to choose the same contract; rather, some of these relationships should not contract at all.³²

Turning to the main issue, it is optimal to disallow evidence when σ is small and s is high but to allow it as σ grows or s falls. We express this conclusion in

Lemma 1: As a function of σ , v^E is affine and decreasing, and v^N is strictly concave and decreasing. There is a number $\sigma^E \in [0, 1]$ such that $v^E(\sigma) > v^N(\sigma)$ for $\sigma > \sigma^E$ and $v^E(\sigma) < v^N(\sigma)$ for $\sigma < \sigma^E$. A necessary and sufficient condition for $\sigma^E = 0$ is $\gamma(1 - sq^E)c''(q^E) < s(1 - s)$, which holds in particular if γ is close to 0. More generally, σ^E becomes smaller as γ falls or if s becomes moderately high.³³

The intuition has been partly explained above: (i) Trial evidence increases accuracy, which is

³¹See note 28 above. As further explanation, when evidence is admitted, the seller anticipates being paid only if she tenders a compliant performance. Trial evidence, in the model, eliminates the possibility that the seller is paid for a noncompliant performance. Thus, when evidence is admitted, the seller's incentive to choose a high effort level is maximized.

³²Note that $q^E(0) = q^N(0)$ and so $v^E(0) = v^N(0)$; the latter is the slope of w at $\sigma = 0$.

³³The proof of Lemma 1 and the proofs for the Propositions that follow are in Appendix A.

good, so an optimal interpretive rule should admit more evidence as evidence production costs fall; (ii) When the performance signal is very accurate — s is high — the enforcer may not need a trial to determine type; (iii) When a large number of types use the same contract, the other pretrial signals of type — litigation narrative and performance — may be insufficient to determine type accurately. Then a trial may be helpful. We illustrate these conclusions in Figure 1.

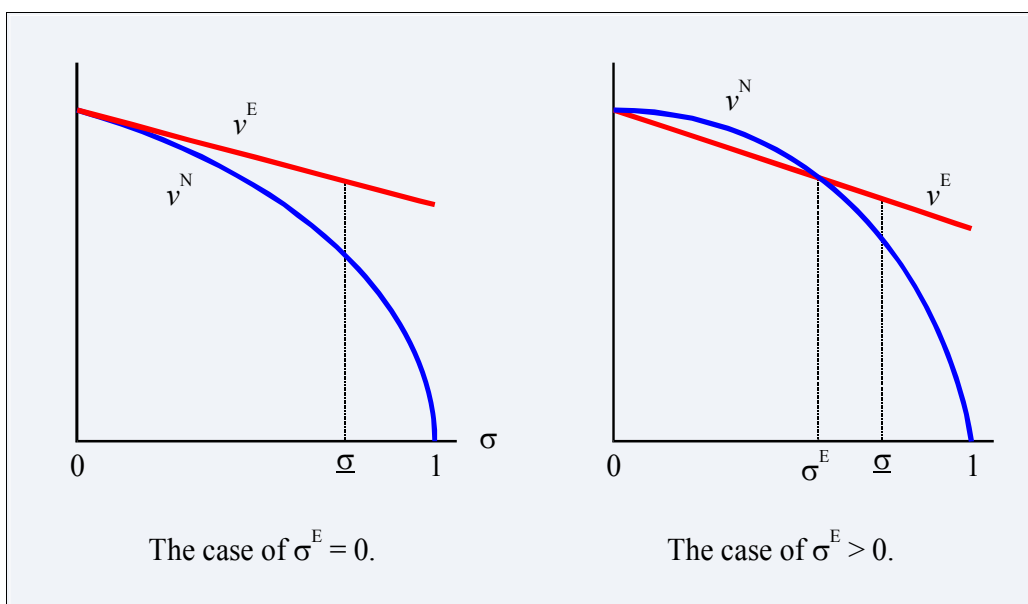


Figure 1: Values with and without evidence.

The realistic case has $\sigma^E > 0$ because γ — trial cost — is positive. The right side figure thus shows that the value of relationships without evidence exceeds the value with evidence when σ is small. As σ gets larger — more types use a particular contract — trial evidence becomes helpful for distinguishing among contracting relationships.

3.3 How the interpretive rule optimally partitions types

In the previous subsection, we derived the optimal interpretive rule for any set S of types that choose the same contract. We next describe how the rule partitions the type space into sets that choose distinct contracts. The optimal partition solves the following “planner’s problem”:

Proposition 1: The optimal interpretive rule and the relationships’ equilibrium behavior solve the problem of selecting an integer K and numbers $\sigma_0, \sigma_1, \sigma_2, \dots, \sigma_K$ to maximize $\sum_{k=1}^K [w(\sigma_k) - \sigma_k y_k]$

subject to $\sum_{k=0}^K \sigma_k = 1$. For each $k = 1, 2, \dots, K$, σ_k is the size of the set of relationships that select contract k . The solution has the following properties:

- (a) The derivative w' exists at σ_k for each $k = 1, 2, \dots, K$.
- (b) The difference $w'(\sigma_k) - y_k$ is constant across $k = 1, 2, \dots, K$, and equals zero if $\sigma_0 > 0$.
- (c) $w'(0) - y_{K+1} \leq w'(\sigma_K) - y_K$.

Property (b) states that if the enforcer’s interpretive rule partitions potential contracting parties into appropriate sets, there would be no welfare gain from further reshuffling those sets: on balance, any gain in welfare from a switch would be matched by a corresponding contracting cost increase. Property (c) adds that when the longest (or most complex) contract the enforcer’s rule induces is contract K , welfare would be reduced if a potential relationship wrote contract $K + 1$. The practical implication of Property (c) is that no relationship would write this contract.

Remark 1:

When parties create a contract, they know what the enforcer will later see — the context, the contract, the performance — and the interpretive rule he uses to identify their type. Hence, each relationship type prefers the contract meant for it. Some relationships also may not form. For example, potential parties may realize that, under the interpretive rule in place, the enforcer cannot identify a conforming performance without the aid of an extensive contractual description

or a trial.³⁴ The contracting or trial cost may exceed particular parties' gains from better effort incentives. An optimal interpretive rule induces parties to make these benefit cost calculations correctly.

Remark 2:

The optimal rule may or may not create asymmetries between sets of relationships that use different contracts, depending on the relation between v^E and v^N . For the parameterization that produced Figure 1, the welfare function w is not concave: $\sigma_E > 0$. For some classes of contracting costs (for instance, $y_k = 0$ up to some large K and y_k large for higher values of k), the optimal interpretive rule creates a partition with equally sized components and the enforcer allows evidence. In general, however, the optimal interpretive rule may induce $\sigma_k > \sigma_j$ — more relationships use contract k than contract j — and it may allow evidence for party types that use k but not for party types that use j .

Remark 3:

The enforcer's and the parties' goals are largely congruent. Both he and they must trade off contract writing, trial and investment cost against the gains from greater interpretive accuracy; and both want these tradeoffs made efficiently. The enforcer, however, maximizes over the set of contractual relationships while individual relationships maximize individual gains. A conflict could arise if, say, a particular relationship prefers not to have a trial while the enforcer wants to face similar types with the prospect of a trial. He could have this preference if relationship types respond to the prospect by writing more informative contracts, and it is cheaper to induce the

³⁴The signal that a conforming performance sends, x , equals t with probability s . In the example described, s is low. A conforming performance, that is, identifies imprecisely the performance that was intended.

efficient effort level, in the context at hand, through contracts than through trials. A concrete implication of this conflict would have an ideal enforcer sometimes refusing to follow parties' interpretive instructions: for example, an instruction to bar all trial evidence.

3.4 A less than ideal enforcer

Actual enforcers may depart from the model's ideal enforcer in systematic ways. Courts are the principal example: Many courts maximize accuracy in interpretation — the correct identification of type — rather than efficiency in investment. A reasonable model of such courts is that they choose interpretive rules to maximize $\sum_{k=1}^K [\sigma v^E(\sigma_k) - \sigma_k y_k]$ rather than $\sum_{k=1}^K [w(\sigma_k) - \sigma_k y_k]$. The judicial enforcer, that is, maximizes the net value of relationships conditional on evidence always being admitted; he does not maximize social welfare.

The characterization results of Proposition 1 continue to hold if we substitute $\sigma v^E(\sigma)$ for $w(\sigma)$ in the statement of the Proposition. This permits us to compare the “optimal” interpretive rule that an accuracy-driven enforcer creates with the optimal rule of the welfare-maximizing enforcer. The comparison yields a surprising result. Intuition apparently suggests that because the accuracy-driven enforcer permits more trial evidence, the incentive of potential relationships to differentiate themselves through contracts falls. As a result, relationships create a smaller set of (simpler) contracts, and fewer relationships contract. This intuition is incomplete, however. The higher evidence costs that the accuracy-driven enforcer induces are proportional to σ ; that is, the larger is the set of relationships that appear similar to the enforcer on the basis of pre-trial evidence, the more frequently the enforcer will admit evidence, and the more evidence he will admit. Put another way, the marginal cost of an increase in σ — the set of similar relationships — is higher for the accuracy-driven enforcer. This creates an incentive for the enforcer, who pursues

efficiency subject always to admitting evidence, to choose an interpretive rule that forces relationships to write different contracts, creating smaller partitions of the contract type space. Such contracts can reduce the trial cost that a rule always admitting evidence would otherwise cause parties to incur.

Depending on whether the factor driving simple contracts and the factor driving complex differentiated contracts dominates, the optimal number of contracts that the accuracy-driven enforcer induces may be larger or smaller than the optimal number that the welfare-maximizing enforcer induces. The next Proposition shows that the number is larger when the evidence cost parameter γ is sufficiently great.

Proposition 2: Assume that $\sigma^E > 0$. The following conclusions hold for γ sufficiently large. Relative to a welfare-maximizing enforcer, an enforcer that always admits trial evidence but is otherwise welfare-maximizing induces (a) fewer relationships to contract and (b) those that contract separate themselves with a larger set of more complicated contracts: σ_0 and K are both weakly higher under the accuracy-driven enforcer. Furthermore, (c) regardless of the parameters, welfare is lower under an enforcer who always admits trial evidence.

Remark 4:

Trials, especially in the United States, are quite costly; that is, γ is large. This explains parts (a) and (b) of the Proposition. These, together with (c), imply that courts should obey parties' interpretive instructions. Under accuracy maximization, there are too many trials relative to the efficient amount. Interpretive instructions almost invariably truncate the evidentiary base that a court would otherwise use to interpret a contract, and so reduce the need for trials. Courts

that follow interpretive instructions thus behave more like the ideal enforcer in the model.³⁵

Therefore, when parties anticipate that a court will follow interpretive instructions, the efficiency with which parties make the requisite tradeoffs increases. Proposition 2 also provides a prediction regarding which type of court parties prefer.³⁶ If it is optimal in reality to disallow trial evidence for some relationships ($\sigma^E > 0$), then we expect that parties prefer courts that constrain the admission of evidence. There is data consistent with this prediction.³⁷

3.5 Enforcer expertise and contracting costs

Our model permits us to disaggregate the concept of enforcer expertise. The parameter s may take on higher values as the enforcer becomes more expert. Recall that, when a performance is compliant, the signal it sends, x , equals the parties' type t with probability s . An expert enforcer may be better at inferring type from performance than a generalist court; hence, s is higher for the expert. Similarly, γ , the cost of using trial evidence, may fall as the enforcer becomes more expert. An expert enforcer may need less evidence to identify type, or be better able to apply evidentiary rules, than would a novice enforcer. Finally, the contracting cost vector y may decrease in enforcer expertise. A particular relationship may need fewer contractual details to

³⁵Interpretive instructions commonly are relatively context free. For example: "A court shall not use written or oral pre-contractual communications between the parties when interpreting this Agreement." A court could infer from such an interpretive instruction only that the parties are in the large set of types that give interpretive instructions. For further explanation, see note 21 above.

³⁶Parties can choose the interpreting court by including a choice of law clause in their contract, which is commonly enforceable regardless of where a transaction is conducted. For example, a contract between Illinois and Kentucky parties will be interpreted under New York law if the parties chose that state's law to govern contracting disputes.

³⁷In the US, the California courts' interpretive practices permit parties almost always to introduce trial evidence. In contrast, the New York courts tend to admit trial evidence only when, to use our terminology, other signals of party type are inconclusive. Our analysis shows that full trials are helpful only in a subset of cases. Hence, we speculate that commercial parties prefer their contracts to be interpreted under New York law. Geoffrey Miller and his collaborators show that actual parties prefer New York by a factor of five (see, e.g., Eisenberg and Miller 2009).

individuate itself before an expert enforcer than it would have to include for a lay enforcer. We summarize this reasoning with

Proposition 3: Social welfare is weakly (a) decreasing in the vector y ; (b) increasing in s ; and (c) decreasing in γ .

We next consider how these parameters combine to affect social welfare. Contracting cost and enforcer expertise could be *complements* or *substitutes*, depending on the nature of enforcer expertise. As complements, when contracting costs fall, parties can write more extensive descriptions. An expert enforcer can combine his ability to evaluate commercial performances with a full description to read the performance signal accurately. As substitutes, the marginal benefit of less costly contracts may decrease in the level of expertise. Parties can send a coarser, and thereby cheaper, contractual signal when the enforcer can identify type using little trial evidence. There is a straightforward relation among these parameters for large parameter shifts.

Proposition 4: For sufficiently large parameter shifts, (a) expertise measured by γ and lower contracting costs (measured by y) are substitutes; (b) expertise measured by s and lower contracting costs are complements.

Proposition 4 shows that enforcer expertise operates at two stages of a contracting dispute and invokes three possibly distinct types of enforcer expertise: (i) Pre-trial: The enforcer can combine his ability to read contracts and his ability to evaluate performance signals with an extensive contract description to recover type with high accuracy, thereby obviating the need for trial evidence; (ii) Trial: The enforcer can combine his ability to evaluate sketchy contracts with his ability to evaluate trial evidence again to determine type with high accuracy, thereby obviating the need for extensive contracts. It is an open question whether an enforcer who is expert at the

pretrial litigation stage also is expert at the trial stage.³⁸

3.6 Optimal Use of Multiple Enforcers

We extend our model to consider two enforcers: a court and an arbitrator. The arbitrator is assumed to have greater expertise in evaluating performances because it is more common for arbitrators to have pertinent industry knowledge. Formally, we let s^C and s^A denote the levels of expertise of the arbitrator and the court, with $s^A > s^C$. For reasons given immediately below, we assume that, holding contract type constant, to use an arbitrator imposes an additional contracting cost of $\alpha \geq 0$ for each relationship served.³⁹ All other parameters remain as before.

We motivate the addition of the contracting cost α as follows: The state provides detailed procedures that govern the pretrial and trial stages of litigation in courts. These procedures are long standing and are glossed by many precedents. Parties can opt into rule systems that govern arbitration — the rules of the American Arbitration Association for example — but those rules are much less detailed and less predictable in application than the state’s procedural rules.⁴⁰ Parties that choose arbitration thus have to create some procedures of their own.

In this setting, an interpretive rule would determine the enforcer to which parties using particular contracts are directed. The interpretive rule also would specify when the buyer must pay and when trial evidence is allowed, as before. Proposition 1 therefore continues to characterize the optimal interpretive rule, with one modification: The welfare function w is now defined by

³⁸See note 6 above.

³⁹As an example, if parties agree to trade ten widgets for \$ x each, their contract would cost p to write if the parties intend judicial enforcement and $p + \alpha$ if the parties intend arbitrator enforcement.

⁴⁰For example, an arbitration panel’s interpretation of an AAA rule often is not publically reported. Hence, precedents seldom inform the application of these rules. In contrast, there is a set of reporters – the Federal Rules Decisions – that report interpretations of the Federal Rules of Civil Procedure.

$$w(\sigma) \equiv \sigma \max \{v^{CE}(\sigma), v^{CN}(\sigma), v^{AE}(\sigma) - \alpha, v^{AN}(\sigma) - \alpha\},$$

where v^{CE} and v^{CN} denote the values of a relationship under court enforcement with and without trial evidence allowed, and v^{AE} and v^{AN} denote the respective values under arbitration. This expression represents that, under the optimal rule, a set of relationships of size σ — those that use the same contract — can be directed to either the court or the arbitrator, and context evidence may be either allowed or disallowed.⁴¹

The modified welfare expression suggests why two enforcement systems can coexist. For sets of relationships that use different contracts, some may benefit more from arbitrator expertise and would optimally be directed to the arbitrator, while others would be optimally directed to the court. To make this precise, we investigate how the marginal benefit of expertise depends on the size σ of a set of relationships that choose the same contract.⁴² We proceed by providing two lemmas that help sort out (i) the relation between σ and whether relationships are optimally sent to arbitration or to court and (ii) the relation between σ and the slope of the welfare function w . In the next section we use Proposition 1 and these lemmas to obtain two predictions for our pilot empirical study.

To produce the first lemma, for convenience we limit attention to a special case in which the cost function for the seller's investment in quality is $c(q) = q^2/2$.⁴³ Take as given an expertise

⁴¹Further variations of this extension are also possible; they amount to variations in the definition of w . For example, to represent the case in which the court maximizes accuracy in interpretation (and thus always admits context evidence), we remove v^{CN} — the value of relationships in court when the court excludes context evidence — from the definition of w .

⁴²Recall that this set is endogenous.

⁴³The analysis can be done more generally but requires some technical assumptions.

level s . We can show that an increase in s causes v^N to rise for values of σ near zero — when few relationships use a similar contract — and to fall for σ close to 1. The increase in s also causes v^N to become more concave. These properties imply the following result.

Lemma 2: Assume $c(q) = q^2/2$ and take as given an arbitration cost α and expertise levels s^C and s^A satisfying $s^A > s^C$. There are numbers z^0 and z^1 satisfying $0 \leq z^0 \leq z^1 \leq 1$, such that

- (i) $v^{AN}(\sigma) - \alpha > v^{CN}(\sigma)$ for all $\sigma \in (z^0, z^1)$ and
- (ii) $v^{AN}(\sigma) - \alpha < v^{CN}(\sigma)$ for all $\sigma \in (0, z^0)$ and for all $\sigma \in (z^1, 1)$.

Furthermore, if α is sufficiently small, then $z^0 = 0$. Also, $z^1 - z^0$ is decreasing in α .

Part (i) states that when the enforcer excludes context evidence, then, holding contract type constant, the value of relationships is greater in arbitration when moderately few parties use this contract, unless the cost of specifying arbitration (α) is large. Part (ii) holds that this result reverses when the number of parties using a particular contract grows.⁴⁴ Regarding the intuition, the relation between party type and contract performance becomes attenuated as more party types use the same contract. The arbitrator's expertise consists in inferring type from performance. Hence, as this ability becomes less valuable, parties do better going to court and saving the cost α of writing arbitration specifications. This result also holds when comparing $v^{AN}(\sigma)$ and $v^{CE}(\sigma)$: that is, when holding σ constant, an accuracy driven court would admit evidence that an arbitrator would disallow. When contracting cost is relatively low and trials are relatively costly, it is optimal to send type sets with a small σ to the arbitrator and sets with a large σ to the court.

The second lemma concerns the relation between σ and the slope of w . Since w is not strictly concave, there may be multiple points with the same slope, so we look at the set of points at which the slope is some value m , $\{\sigma \mid w'(\sigma) = m\}$.

⁴⁴ This comparison between $v^{AN}(\sigma)$ and $v^{CN}(\sigma)$ is relevant when γ is relatively large so that $\sigma^E > 0$.

Lemma 3: The set $\{\sigma \mid w'(\sigma) = m\}$ is decreasing in m in the sense that as m increases, the endpoints of the set $\{\sigma \mid w'(\sigma) = m\}$ decrease.

This Lemma holds that, generally speaking, when σ is small — that is, relatively few types use a similar contract — then welfare is increasing relatively sharply. Lemma 3 is independent of the specific functional form that $c(q)$ may take.

Remark 5:

In this section we have argued that multiple enforcers can be explained by differences in their expertise and direct costs. The benefit of multiple enforcers can also be motivated by differences in expertise and evidence or contracting costs. An enforcer with a lower value of γ — i.e., cheaper trials — can partition types conveniently when contracts are relatively expensive to create. Parties with high contracting costs thus may prefer the low γ system. Therefore, Proposition 4 implies that in equilibrium different enforcement systems that offer different types of expertise will exist when there is heterogeneity of contracting cost across parties.

Our conclusion about the benefit of multiple enforcers is consistent with the existence, in almost every industry studied, of a positive fraction of contracts that contain arbitration clauses (e.g. Eisenberg and Miller 2007, Drahozal and Ware 2010). The use of arbitrators and variations in interpretive rules across enforcement systems and jurisdictions, allows parties more precisely to convey context information, so each contractual relationship has an enhanced probability of receiving an interpretive style tailored to its type.

4. Empirical Implications and Pilot Study

4.1 Predictions

Arbitration has been studied extensively, but seldom in a contract theoretic framework. We attempt here to show that this framework can generate interesting empirical predictions. We test certain implications of Proposition 1, the extension in Section 3.6 and a part of Proposition 4. Recall that the contract space is ordered by cost, so that for any two contracts $k > j$, the cost y_k of contract k is weakly larger than is the cost y_j of contract j . This implies, practically, that contract k is longer and more complicated than contract j . We shall measure contract length by the number of words.

To produce our first prediction, we combine Proposition 1(b) with Lemma 3. Property (b) of Proposition 1 holds that longer contracts are associated with larger marginal values of the welfare function w . That is, for contracts k and j with $y_k > y_j$ (so contract k has more words), the optimal interpretive rule creates partitions σ_k and σ_j satisfying $w'(\sigma_k) > w'(\sigma_j)$. Lemma 3 then implies that, up to an equivalence set, σ_k is smaller than σ_j . We thus obtain

Prediction 1: Across all contracts written in the population of relationships, there is an inverse relation between contract length (measured by number of words) and the number of relationships choosing a specific contract of this length.

The prediction refers to the frequency with which relationships use a specific contract. For a given number of words, λ , there are a number of potential contracts with exactly λ words. The number of such contracts is increasing in λ . Thus, if we found that the number of relationships choosing contracts with λ words is itself increasing in λ , the finding would not contradict Prediction 1. The relation implied in Prediction 1 would be confirmed, however, if the data shows

an inverse relation between λ and the number of relationships choosing contracts with λ words.

Lemma 2 helps to generate our second prediction. Under the assumption that α is small and γ is relatively large, the optimal interpretive rule directs type sets with small values of σ to the arbitrator and sets with large values of σ to the court. Because σ is inversely related to contract length, we expect longer contracts to specify arbitration more frequently.⁴⁵ A contributing factor to this Prediction is that parties who choose arbitration sometimes must specify litigation procedures; this makes contracts longer. A possible countervailing factor is that, as shown in Lemma 2, if α and/or γ is moderate, litigation in court may be optimal for some sets of relationships when σ is small (i.e., a few relationships write very long contracts). Thus, we obtain

Prediction 2: There is a positive relation between contract length and the choice of arbitration, except perhaps for the longest contracts.

Proposition 4 holds that arbitrator expertise in inferring type from performance and contract descriptions of context are complements. The Proposition thus implies that contracts directed to arbitrators should contain more “whereas” clauses, which directly convey type information.⁴⁶ This reasoning yields

Prediction 3: Contracts that are enforced by arbitrators should contain more “whereas” clauses than contracts intended for courts.

⁴⁵We are presuming here that, to be directed to arbitration, the parties have to specify arbitration in their contract. Our model is silent on the actual words used in the contract. Note that a positive correlation between a measure of complexity and the probability that parties choose arbitration is found by Drahozal and Hylton (2003) and Drahozal and Ware (2010).

⁴⁶As an example, a whereas clause may recite: “Whereas, the parties herein want to develop a widget that meets USDA specification 4.35, and”

4.2 Results

The data for our preliminary analysis consists of contracts filed with the Securities and Exchange Commission. Using the SEC's EDGAR database portal, we formed a sample of 42,822 documents spanning years 2001-2005.⁴⁷ Analysis is restricted to documents with word counts between 300 and 30,000, documents that do not contain "Amendment" in the title, and to the industries noted in Appendix B. For each document in the sample, an algorithm was used to count the occurrence of the following keywords: arbitration (and variants), whereas, court, appeal, mediation, litigation, warranty, guaranty, specification and deposition.

To verify that word counts were a suitable measure for further analysis, research assistants assessed the performance of the algorithm for 250 of the documents in the sample. While other details of these documents were explored, the results of this assessment indicated that word units did provide an accurate proxy. For example, the contextual "whereas" content classified by research assistants corresponded to increases in the automated count of "whereas."

We performed a Probit regression on the probability of observing at least one arbitration word (which we take as suggesting that the contract includes an arbitration clause); these are reported in Appendix B. The regressions were done for the entire sample and industry by industry (construction, manufacturing, minerals, retail, services, wholesale). Two alternative specifications were used. The first uses only word count and word count squared as explanatory variables (Table

⁴⁷Overdahl (1991) provides a summary of the SEC filing requirements. Material contracts were recovered through the SEC's ftp server, where all filings are listed in quarterly indexes which include form type. The contents of each filing are also indexed and indicate exhibits included with the filing. This filing level index was used to identify all "Exhibit 10" documents included with the filing for any filing with a form type on the list which requires the inclusion of material contracts. Finally, each exhibit 10 was recovered from the location indicated in the filing index. Additional information recovered for each contract includes the filing company's name, SEC identifier, the filing date for each form filed, and the filing company's SIC code and IRS number.

3-1 in Appendix B). The second uses word count and word count squared as well as the other word counts (Table 3-2). Under either specification, the coefficients on word count and word count squared have the same magnitudes and are statistically significant at the 1% level. In all cases, likelihood ratio tests indicated that the coefficients were jointly significant. Many of the other word counts were also found to be significant.

The predictions described above are generally borne out in the data. As Figures 2 and 3 show, except for very short contracts, there is an inverse relation between contract length (word count) and frequency, consistent with Prediction 1. We also find an increasing, concave relation between contract length and arbitration, consistent with Prediction 2. Further, we find a significant positive relation between arbitration and (i) the number of “whereas” terms in a contract, consistent with Prediction 3, and (ii) the terms “deposition,” “court,” and “mediation,” which is consistent with our assumption that parties who use arbitration must specify litigation procedures.

Figure 2 illustrates the relationship between the probability of observing at least one arbitration word and the length of the document under specification two for the entire sample (the figure is similar for specification one). All other explanatory variables were fixed at their respective means. The number of documents containing at one least arbitration word is also presented. As is apparent from the Figure, the probability that parties use arbitration is lowest for short contracts, it increases as contracts get longer, and it is concave. The figure also shows the distribution of contracts in the data set by total number of words.⁴⁸

That the curve in Figure 2 bends downward at the right is a bit misleading, because the

⁴⁸Most of the contracts contain fewer than 10,000 words.

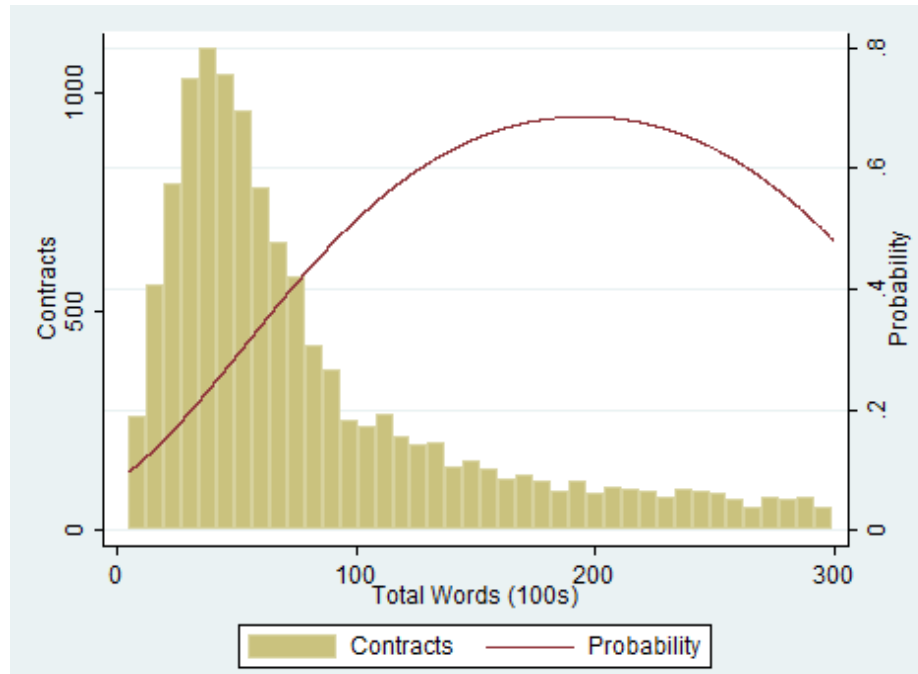


Figure 2: Regression fit of the probability of arbitration, and number of contracts in the data set, by document length

curve's shape is driven mainly by the large number of observations in the lower range of word counts. Figure 3 shows the actual fraction of contracts containing at least one arbitration word, for twelve groups of contracts that are defined by total word counts (0 – 2500 words, 2501 – 5000 words, etc.). The figure also shows the distribution of contracts that contain at least one arbitration word. The relationship between word count and the fraction of contracts containing at least one arbitration word is increasing and concave, although it is not negative at high word counts.

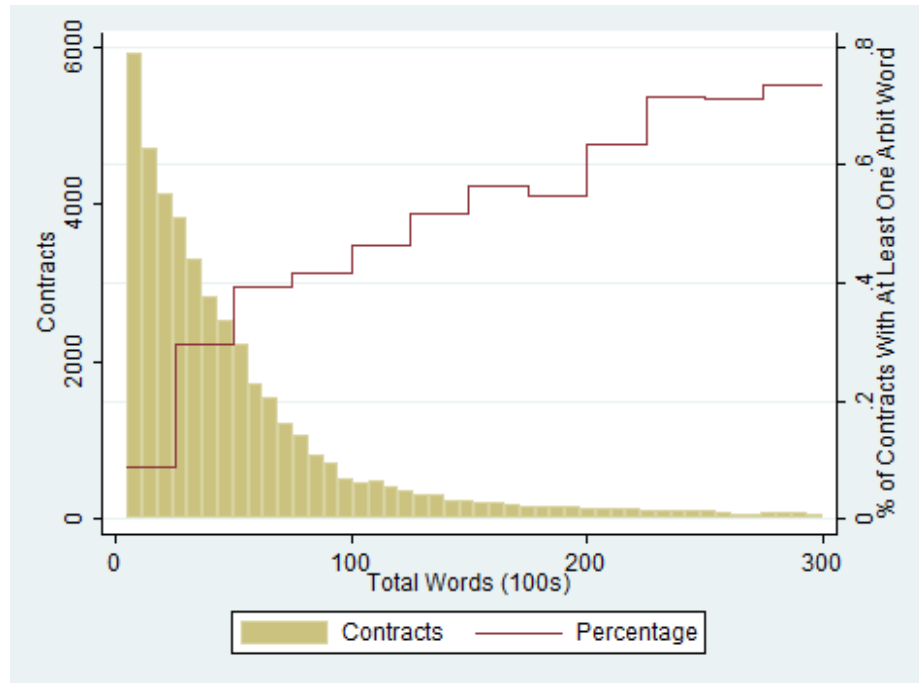


Figure 3: Fraction of contracts with an arbitration word, and number of contracts with at least one arbitration word, by document length

5. Conclusion

This paper analyzes contract interpretation formally. We initially show how a welfare maximizing enforcer induces contracting parties to make socially efficient tradeoffs between the gain in better effort incentives that accurate interpretation yields against the costs of contract writing, investment in the deal and trial cost. We also disaggregate enforcer expertise into its three aspects: (i) the ability to extract meaning from words; (ii) the ability to infer ex ante intention from tendered performances; and (iii) the ability to infer intention from trial evidence. This exercise permits us to show that contract writing costs and evaluating performances are complements while

those costs and trials are substitutes. Finally, an optimal interpretive rule induces parties to choose enforcement institutions — court or arbitrator — so as to balance the marginal benefits of the sets of relationships that are partitioned by the contracts.

Courts are not optimal enforcers but the cost factors that drive enforcer and party choice are at play in the real world. Hence, we expect parties to behave, at least roughly, as our model would predict. It turns out to be relatively convenient to test this view in two dimensions: the complexity of the contracts that actual parties use and their choice of court or arbitrator. There is support in the contracting data for our predictions in these dimensions. Finally, we show that enforcers who maximize accuracy in interpretation reduce welfare in contrast to enforcers who maximize welfare directly. The consequences of accuracy maximization include distortions in the extent of contractual specifications, in the fractions of relationships choosing to contract, and in the number of trials, which are more frequent than welfare maximization would imply. Many courts attempt to maximize accuracy. It follows that courts should follow interpretive instructions that parties send when those instructions, as is commonly the case, would have courts reduce the number and scope of trials.

Appedix A: Analysis Not Presented in the Text

Proof of Lemma 1: This result follows from the seller's first-order conditions and the first and second derivatives of the value functions v^E and v^N , using the implicit function theorem and our assumptions on c . ■

Proof of Proposition 1: We must show that, if the enforcer partitions relationships by contracts in order to maximize aggregate welfare, then each contracting relationship prefers the contract meant for it. First, notice that it could not be optimal for a relationship to switch from its prescribed contract (or no contract) to another contract, holding fixed the interpretive rule. The enforcer would believe this relationship to have a type in some set S when in fact its type is not in that set. The seller would then have no incentive to invest because the enforcer would believe that her performance was noncompliant. Hence, the contracting relationship would have no value but it still would incur a contracting cost. Parties to this relationship would be better off not forming a contract. Second, observe that it could not be optimal for a relationship to switch from its prescribed contract to not contracting at all; if it could gain by doing so, then the enforcer would induce the parties not to contract (which contradicts the assumption that the enforcer maximizes aggregate welfare). This step also uses the fact that decreasing σ , as would occur when relationships move from contracting to not contracting, has a strictly positive effect on those relationships that remain in the set S . These relationships are more to have their contracts interpreted accurately.

Regarding the properties of the solution, property (a) follows from the fact that w is the upper envelope of concave functions. The only point where the derivative of w would not be defined is the positive point of intersection between $\sigma v^E(\sigma)$ and $\sigma v^N(\sigma)$, which is σ^E . At σ^E the slope of w from the right is strict greater than is the slope from the left. If $\sigma_k = \sigma^E$ for some k then the enforcer could slightly raise or lower σ_k and make a corresponding adjustment to some other σ_i , resulting in a strict increase in total welfare (a contradiction). The same logic implies Property (b). Note that if $\sigma_0 > 0$, there must be no way to increase total welfare with a marginal increase in the sizes of any set of relationships choosing the same contract, which means $w'(\sigma_k) - y_k$ must be zero for all $k = 1, 2, \dots, K$. Property (c) holds that, at the margin, total welfare cannot be increased by shifting some relationships from an existing contract to the least-cost unused contract. ■

Using Proposition 1, we can provide an algorithm for calculating the solution in the two special cases of $\sigma^E = 0$ (when it is optimal to always admit context evidence) and $\sigma^E v^{E'}(\sigma^E) + v^E(\sigma^E) < y_1$ (when it is optimal to always disallow context evidence). We first define, for any $\beta \geq 0$, a sequence $\{\mu_j^\beta\}_1^{j^\beta}$ as follows.⁴⁹ For each positive integer j , we look for a number μ_j^β that solves $w'(\mu_j^\beta) = y_j + \beta$. Because w is strictly concave and differentiable over the relevant region (in both cases considered here), the number μ_j^β either does not exist (which is the case if and only if $w'(0) <$

⁴⁹In the expressions to follow, where β is shown as a superscript, it refers to an index rather than an exponent.

$y_j + \beta$) or it uniquely exists. By strict concavity of w , that $w'(0) = v^N(0) = v^E(0)$, and that y_i is unbounded as j becomes large, there exists an integer J^β such that μ_j^β is defined if and only if $j \leq J^\beta$. Furthermore, μ_j^β is weakly decreasing in j .

If it is the case that $\sum_{j=1}^{J^0} \mu_j^0 \leq 1$ then the solution is defined by $K = J^0$ and $\sigma_k = \mu_j^0$ for all $k = 1, 2, \dots, K$, and in this case we have $\sigma_0 = 1 - \sum_{j=1}^{J^0} \mu_j^0$. Otherwise, we choose β to maximize $\sum_{j=1}^{J^\beta} [w(\mu_j^\beta) - \mu_j^\beta y_j]$ subject to $\sum_{j=1}^{J^\beta} \mu_j^\beta = 1$, and this gives the solution.

Proof of Proposition 2: Recall that the accuracy-driven court uses $\sigma v^E(\sigma)$ in place of $w(\sigma)$. Clearly, $\sigma v^E(\sigma) \leq w(\sigma)$, which proves the welfare claim. Regarding parts (a) and (b), note that $w(\sigma) = \sigma v^N(\sigma)$ for $\sigma \leq \sigma^E$, so for such values of σ we have to examine the relation between using $\sigma v^N(\sigma)$ and $\sigma v^E(\sigma)$ in the solution algorithm. Note that $\sigma v^N(\sigma)$ and $\sigma v^E(\sigma)$ both equal 0 at $\sigma = 0$. Recall that q^E is constant in σ , whereas q^N solves $s(1 - \sigma) = c'(q^N)$ so we write $q^N(\sigma)$. The derivatives of $\sigma v^N(\sigma)$ and $\sigma v^E(\sigma)$ are, respectively,

$$v^N(\sigma) + \sigma v^{N'}(\sigma) = q^N(\sigma) - c(q^N(\sigma)) - s\sigma[1 - c'(q^N(\sigma))]c''(q^N(\sigma))^{-1}$$

and

$$v^E(\sigma) + \sigma v^{E'}(\sigma) = q^E - c(q^E) - 2\gamma(1 - sq^E)\sigma.$$

The calculation for the former derivative utilizes the implicit function theorem to establish that $\partial q^N / \partial \sigma = -sc''(q^N(\sigma))^{-1}$.

These derivatives are positive and equal at $\sigma = 0$, because $v^N(0) = v^E(0)$. Also, both derivatives are strictly decreasing, continuous, and take negative values for large enough σ . Thus, the inverse functions of the derivatives are well defined and continuous on $[0, v^E(0)]$. Write $g^N(m) = \sigma$ for the inverse of $v^N(\sigma) + \sigma v^{N'}(\sigma)$, and denote by $g^E(m) = \sigma$ the inverse of $v^E(\sigma) + \sigma v^{E'}(\sigma)$. For γ sufficiently large, we can guarantee that for every $m \in [0, v^E(0))$, it is the case that $g^E(m) < g^N(m)$. This further implies that $\sigma v^N(\sigma) > \sigma v^E(\sigma)$ for all $\sigma \in (0, g^N(0))$, so the welfare-maximizing enforcer always disallows evidence.

From the notes following the proof of Proposition 1, we know that the optimal interpretive rule of the welfare-maximizing enforcer satisfies $g^N(y_k + \beta) = \sigma_k$ for some $\beta \geq 0$ and all $k = 1, 2, \dots, K$. Applying the same K and β values to the accuracy-driven enforcer yields strictly lower values of σ_k , because $g^E(y_k + \beta) < g^N(y_k + \beta)$. If $\beta > 0$ for the welfare-maximizing enforcer then the accuracy-driven enforcer's optimal interpretation is characterized by a lower value of β and hence a larger value of K , and σ_0 must also be weakly higher because this is zero for the welfare-maximizing enforcer. If $\beta = 0$ for the welfare-maximizing enforcer, then the accuracy-driven enforcer's optimal interpretation is characterized by the same value of β and hence the same value of K , and σ_0 must be strictly higher. This proves claims (a) and (b). ■

It is clear what interpretive rule would be ‘‘optimally’’ adopted under the constraint that evidence always be allowed. Because v^E is an affine and decreasing function, $\sigma v^E(\sigma)$ is quadratic and concave. The interpretive rule thus would partition contracting relationships into equally sized elements, so that $\sigma_1 = \sigma_2 = \dots = \sigma_K$.

Proof of Proposition 3: For each favorable parameter shift described (γ lowered, s or γ increased), welfare would weakly rise under the optimal interpretive rule prescribed for the *original* parameter values (that is, before the shift). Thus, by adjusting the interpretive rule to its new optimum, welfare is constant or increases. In particular, (a) parties are more able to communicate their type in the contract when contracting costs decrease; (b) enforcer accuracy increases as an enforcer is better able to infer type from the seller's physical performance (s is high); and (c) parties supply more context evidence as evidence production cost falls. ■

Proof of Proposition 4: For a fixed value of s , given the constraints in the model, an upper bound on the effort level that can be induced is the value q^H that solves $s = c'(q^H)$. This follows from the seller's first-order conditions in both the evidence and no-evidence cases; q^E actually achieves this bound and q^N is generally below it. Let $v^H \equiv q^H - c(q^H)$ denote the joint value for a relationship in this case, gross of contracting costs and without evidence costs. With s fixed, v^H is a type's highest possible joint value.

Respecting part (a) of the Proposition, start from any given values of the other parameters. As γ approaches zero, the seller's effort comes arbitrarily close to q^H , so the relationship's joint value comes arbitrarily close to v^H regardless of the parameter γ . This result also obtains if contracting cost γ converges to zero, so that relationships could distinguish themselves finely by their contract selection; that is, as $s\gamma$ approaches the zero vector, the seller's effort comes arbitrarily close to q^H and the relationship's joint value comes arbitrarily close to v^H regardless of the parameter γ . Thus, low values of γ and y are substitutes in the large.

Respecting part (b), an almost efficient level of effort q^* can be induced only if s is high and at least one of the other shifts just discussed occurs. Recall that x is the signal that the seller's performance generates. If the seller complied, then $x = t$ with probability s ; and x is uninformative with probability $1 - s$ (that is, x is uniformly distributed over $[0, 1]$). If the seller breached, x is similarly uninformative. Hence, a compliant performance under a low s is equivalent, in the enforcer's view, to a breach. It follows that a seller cannot be given good incentives unless s is high. But if s is close to 1 then sufficiently lowering the vector y will motivate the seller to select an effort level that is arbitrarily close to q^* . Hence, expertise measured by s and improved language are complements for large enough shifts. ■

Proof of Lemma 2: Note that with $c(q) = q^2/2$, we get $q^N = s(1 - \sigma)$. Substituting this into the expression for v^N yields

$$v^N(\sigma) = s(1 - \sigma)[2 - s(1 - \sigma)]/2.$$

We have $\partial^2 v^N / \partial^2 \sigma \partial s = -2s < 0$. Thus, as s increases, v^N becomes more concave and the value of $v^N(0)$ rises. An implication is that if $v^{AN}(\sigma) - \alpha = v^{CN}(\sigma)$ for some value of σ where $v^{AN}'(\sigma) < v^{CN}'(\sigma)$, then $v^{AN}(\sigma') - \alpha < v^{CN}(\sigma')$ for all $\sigma' > \sigma$. We conclude that $v^{AN}(\sigma) - \alpha$ and $v^{CN}(\sigma)$ intersect at most twice, with $v^{AN}(\sigma) - \alpha > v^{CN}(\sigma)$ in the interior region. This proves claims (i) and (ii) of the lemma. Note also that $\partial v^N / \partial s = (1 - \sigma)[1 - s(1 - \sigma)] > 0$, so $v^{AN}(0) > v^{CN}(0)$. This implies that when α is small $z^0 = 0$. The last statement of the lemma follows from the concavity of $v^{AN}(\sigma)$. ■

Proof of Lemma 3: This lemma follows from the fact that w is the upper envelope of concave functions $\sigma v^{CN}(\sigma)$, $\sigma v^{CE}(\sigma)$, $\sigma v^{AN}(\sigma) - \sigma\alpha$, and $\sigma v^{AE}(\sigma) - \sigma\alpha$. Note also that the derivative of w is maximized at $\sigma = 0$. ■

Appendix B: Empirical Details

This appendix provides details of the pilot study discussed in Section 6.2. Table 1 compares the length of each document and the number of occurrences of arbitration words. Bins for word count are set to each contain 20% of the sample. As the table indicates, longer contracts often contain more arbitration words with frequency of occurrence increasing with contract length.

"Arbitration" words	Word Count					Total
	500-1492	1493-2811	2812-4579	4580-7725	>7726	
None	8,139	7,253	5,991	5,216	4,165	30,764
% of Total	95.07%	84.68%	69.95%	60.89%	48.63%	71.84%
1-2	215	465	667	801	1,504	3,652
% of Total	2.51%	5.43%	7.79%	9.35%	17.56%	8.53%
3-4	75	215	273	315	547	1,425
% of Total	0.88%	2.51%	3.19%	3.68%	6.39%	3.33%
5-9	89	328	723	782	686	2,608
% of Total	1.04%	3.83%	8.44%	9.13%	8.01%	6.09%
10 or more	43	304	911	1,452	1,663	4,373
% of Total	0.50%	3.55%	10.64%	16.95%	19.42%	10.21%
Total	8,561	8,565	8,565	8,566	8,565	42,822
% of Total	100	100	100	100	100	100

Table 1 : Arbitration count by length of document

Below is a list of the industries that were included in the sample:

2 Digit SIC	Count	%	Description	Group
10	155	0.36	Metal mining	Mineral Industries
12	127	0.3	Coal mining	Mineral Industries
13	1199	2.8	Oil and gas extraction	Mineral Industries
14	66	0.15	Nonmetallic minerals, except fuels	Mineral Industries
15	223	0.52	General building contractors	Construction Industries
16	261	0.61	Heavy construction contractors	Construction Industries
17	172	0.4	Special trade contractors	Construction Industries
22	274	0.64	Textile mill products	Manufacturing
23	531	1.24	Apparel and other textile products	Manufacturing
25	229	0.53	Furniture and fixtures	Manufacturing
26	296	0.69	Paper and allied products	Manufacturing
27	695	1.62	Printing and publishing	Manufacturing
28	5330	12.45	Chemicals and allied products	Manufacturing
29	205	0.48	Petroleum and coal products	Manufacturing
30	650	1.52	Rubber and miscellaneous plastics products	Manufacturing
31	173	0.4	Leather and leather products	Manufacturing
32	157	0.37	Stone, clay, glass, and concrete products	Manufacturing
33	681	1.59	Primary metal industries	Manufacturing
34	722	1.69	Fabricated metal products	Manufacturing
35	3,001	7.01	Industrial machinery and equipment	Manufacturing
36	4,168	9.73	Electrical and electronic equipment	Manufacturing
37	937	2.19	Transportation equipment	Manufacturing
38	2,906	6.79	Instruments and related products	Manufacturing
39	528	1.23	Miscellaneous manufacturing industries	Manufacturing
50	1329	3.1	Wholesale trade--durable goods	Wholesale Trade
51	887	2.07	Wholesale trade--nondurable goods	Wholesale Trade
52	91	0.21	Building materials, hardware, garden supply, & mobile	Retail Trade

53	397	0.93	General merchandise stores	Retail Trade
55	249	0.58	Automotive dealers and gasoline service stations	Retail Trade
56	474	1.11	Apparel and accessory stores	Retail Trade
57	283	0.66	Furniture, home furnishings and equipment stores	Retail Trade
70	490	1.14	Hotels, rooming houses, camps, and other lodging places	Service Industries
72	378	0.88	Personal services	Service Industries
73	9,821	22.93	Business services	Service Industries
78	232	0.54	Motion pictures	Service Industries
79	970	2.27	Amusement and recreational services	Service Industries
80	1,698	3.97	Health services	Service Industries
82	276	0.64	Educational services	Service Industries
87	1561	3.65	Engineering and management services	Service Industries

Table 2: Documents by two-digit SIC code

Tables 3-1 and 3-2 present the marginal effects from the Probit regression on the probability of observing at least one arbitration word. The regressions were done for the entire sample and industry by industry. Two alternative specifications were used. The first uses only word count and word count squared as explanatory variables (Table 3-1). The second uses word count and word count squared as well as the other word counts (Table 3-2). Under either specification the coefficients on word count and word count squared have the same magnitudes and are statistically significant at the 1% level. In all cases, likelihood ratio tests indicated that the coefficients were jointly significant.

	All	Minerals	Construction	Manufacturing	Wholesale	Retail	Service
Word Count (100s)	0.00621*** (0.000115)	0.00580*** (0.000595)	0.00889*** (0.000934)	0.00586*** (0.000159)	0.00676*** (0.000491)	0.00811*** (0.000590)	0.00635*** (0.000200)
Word Count Squared	-1.59e-05*** (4.77e-07)	-1.63e-05*** (2.36e-06)	-2.53e-05*** (3.89e-06)	-1.42e-05*** (6.61e-07)	-1.69e-05*** (1.94e-06)	-2.17e-05*** (2.41e-06)	-1.72e-05*** (8.32e-07)
Observations	42,822	1,547	656	21,483	2,216	1,494	15,426
Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1							
Table 3-1: Probit regression of appearance of Arbitration word (Specification 1)							

	All	Minerals	Construction	Manufacturing	Wholesale	Retail	Service
Word Count (100s)	0.00456*** (0.000125)	0.00223*** (0.000684)	0.00589*** (0.00104)	0.00434*** (0.000172)	0.00573*** (0.000565)	0.00768*** (0.000700)	0.00454*** (0.000221)
Word Count Squared	-1.28e-05*** (4.98e-07)	-1.04e-05*** (2.58e-06)	0.05*** (4.38e-06)	-1.14e-05*** (6.89e-07)	-1.65e-05*** (2.13e-06)	-2.07e-05*** (2.92e-06)	-1.31e-05*** (8.80e-07)
Whereas Count	0.0128*** (0.00105)	0.0138** (0.00577)	0.00525 (0.00921)	0.0123*** (0.00144)	0.0128** (0.00504)	0.0262*** (0.00571)	0.0138*** (0.00182)
Court Count	0.0198*** (0.000599)	0.0438*** (0.00406)	0.0719*** (0.00846)	0.0177*** (0.000876)	0.0116*** (0.00293)	0.00861*** (0.00259)	0.0220*** (0.00103)
Appeal Count	5.63e-05 (0.00146)	0.0262** (0.0118)	-0.0123 (0.00757)	0.00264 (0.00192)	-0.000545 (0.00431)	-0.0296** (0.0120)	-0.00365 (0.00309)
Mediation Count	0.0538*** (0.00226)	0.0818*** (0.0230)	0.0368 (0.0329)	0.0353*** (0.00236)	0.177*** (0.0349)	0.442*** (0.0824)	0.0835*** (0.00631)
Litigation Count	-0.00218* (0.00111)	0.0331*** (0.00927)	0.0218 (0.0168)	-0.00955*** (0.00147)	0.0178*** (0.00645)	0.0154 (0.0107)	0.00442** (0.00225)
Waranty Count	0.00290*** (0.000467)	0.00329 (0.00279)	-0.0109** (0.00548)	0.00485*** (0.000642)	0.00387* (0.00201)	0.00119 (0.00233)	-9.65e-05 (0.000849)
Guaranty Count	-0.000915*** (9.74e-05)	-0.00229*** (0.000547)	0.000244 (0.000424)	-0.000843*** (0.000146)	-0.00105** (0.000428)	-0.000697* (0.000416)	-0.00108*** (0.000168)
Specification Count	0.00105 (0.000656)	0.0126* (0.00645)	0.00300 (0.0101)	0.00124* (0.000728)	-0.00302 (0.00537)	-0.00995 (0.00863)	0.000106 (0.00165)
Deposition Count	0.0798*** (0.00819)	-0.135* (0.0762)		0.0659*** (0.00961)	0.139** (0.0549)	0.239*** (0.0763)	0.128*** (0.0184)
Observations	42,822	1,547	656	21,483	2,216	1,494	15,426
Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1							
Table 3-2: Probit regression of appearance of Arbitration word (Specification 2)							

As Table 3-2 indicates, there is a significant positive relation between arbitration and the number of “whereas” terms in a contract. There is also a significant positive relation between arbitration and the terms “deposition,” “court,” and “mediation,” which we associate with the specification of litigation procedures. Other terms are significant (positively or negatively) in some industries; we have not conjectured about these relations.

Figure 3 in Section 6.2 illustrates the relationship between the probability of observing at least one arbitration word and the length of the document under specification two for the entire sample. The figure is similar for specification one. The number of documents containing at one least arbitration word is also presented. Under the second specification all other explanatory variables were fixed at their respective means. Figure 5 shows the relationship between the number of arbitration words and document length for the entire sample. This figure shows that, conditional on containing an arbitration word, the number of such words in a contract document is not significantly increasing with total word count. Thus, the important thing is the presence of at least one arbitration word, the probability of which is what our regressions are measuring.

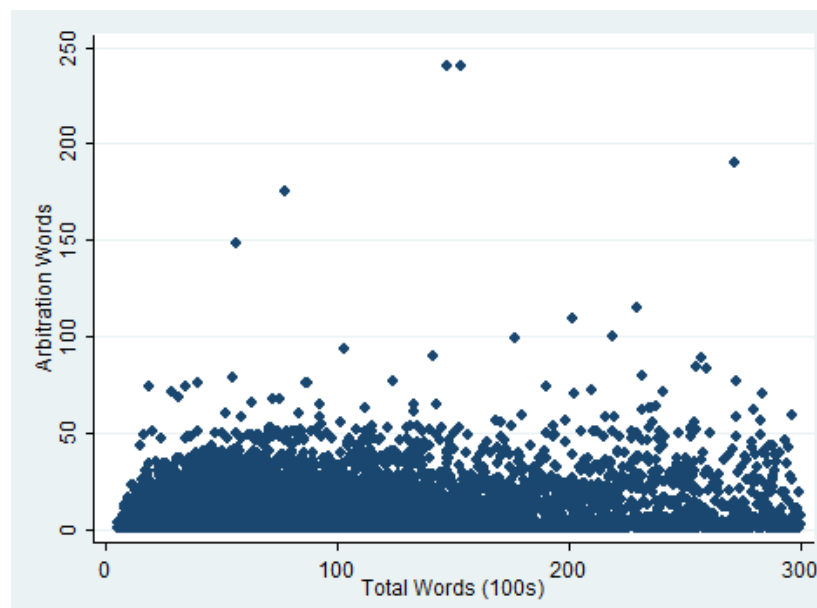


Figure 5: Number of arbitration words by document length

Figures 6 and 7 present the information on probability of arbitration for Manufacturing and Minerals industries, respectively.

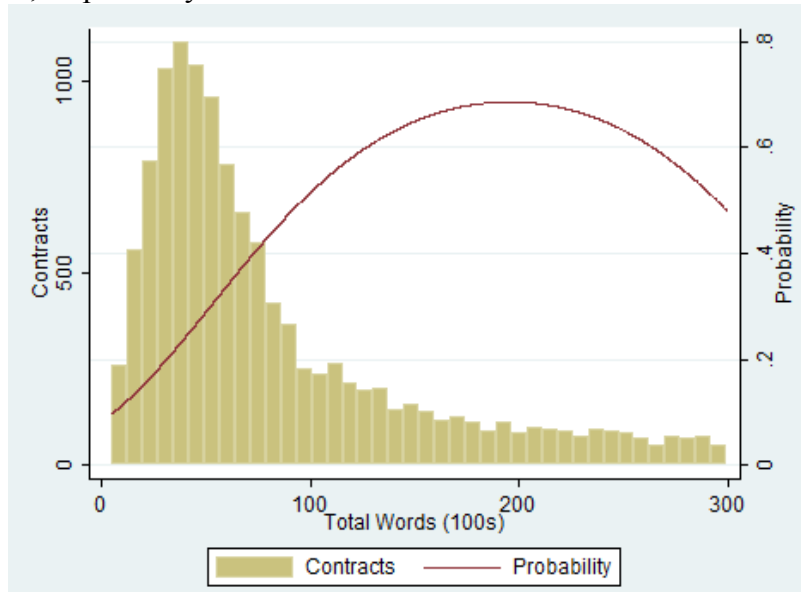


Figure 6: Probability of arbitration – Manufacturing

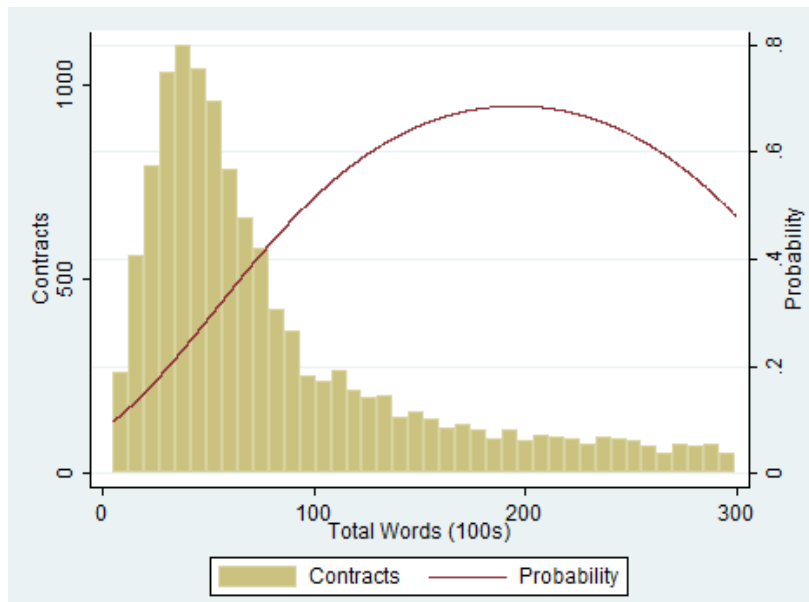


Figure 7: Probability of arbitration – Minerals

References

- Anderlini, L. and L. Felli (1999), “Incomplete Contracts and Complexity Costs,” *Theory and Decision* 46: 23–50.
- Battigalli, P. and G. Maggi (2002), “Rigidity, Discretion, and the Costs of Writing Contracts,” *American Economic Review* 92: 798–817.
- Drahozal, C.R. and S.J. Ware (2010), “Why Do Businesses Use (or Not Use) Arbitration Clauses?” *Ohio State Journal on Dispute Resolution* 25, forthcoming.
- Drahozal, C.R. and K. Hylton (2003), “The Economics of Litigation and Arbitration: An Application to Franchise Contracts,” *Journal of Legal Studies* 32: 549-584.
- Dye, R.A. (1985), “Costly Contract Contingencies,” *International Economic Review* 26: 233–50.
- Eisenberg, T. and G.P. Miller (2007), “Flight from Arbitration: An Empirical Study of Ex Ante Arbitration Clauses in the Contracts of Publicly Held Companies,” *DePaul Law Review* 56: 335-374.
- Eisenberg, T. and G. P. Miller (2008), “The Flight to New York: An Empirical Study of Choice of Law and Choice of Forum Clauses in Publicly-Held Companies’ Contracts,” NYU Center for Law and Economics working paper.
- Kvaloy, O. And T.E. Olsen (2009), “Endogenous Verifiability and Relational Contracting”, *American Economic Review* 99: 2193-2208.
- Listokin, Yair (2010), “Bayesian Contractual Interpretation”, 39 J. Legal Studies 359-374.
- Maggi, G. and R. Staiger (2008), “The Role of Dispute Settlement Procedures in International Trade Agreements,” *American Economic Review*, forthcoming.
- Overdahl, J. A. (1991), “A Researcher’s Guide to the Contracts of Firms Filing with the SEC,” *Journal of Law and Economics* 34: 695-701.
- Posner, E. (1998), “The Parol Evidence Rule, the Plain Meaning Rule, and the Principles of Contractual Interpretation,” *University of Pennsylvania Law Review* 146: 533-577.
- Posner, R. (2004), “The Law and Economics of Contract Interpretation,” *Texas Law Review* 83: 1581-1614.
- Schwartz, A. and R.E. Scott (2003), “Contract Theory and the Limits of Contract Law”, *Yale Law Journal* 113: 541- 619.
- Schwartz, A. and R.E. Scott (2010), “Contract Interpretation Redux”, *Yale Law Journal* 119: 926-964.
- Schwartz, A. and J. Watson (2004), “The Law and Economics of Costly Contracting,” *Journal of Law, Economics and Organization* 20: 2–31.
- Shavell, S. (2006), “On the Writing and the Interpretation of Contracts,” *Journal of Law, Economics and Organization*, 22: 289-314.

Zhao, Xiaojian (2010), "Economics of Contract Interpretation: A Survey", Manuscript

December, 2011