Contract Interpretation and Contextual Asymmetry

Alan Schwartz and Joel Watson

This version: February 2010.

Abstract

A problem in economic contract theory is to know when parties can create efficient investment incentives with contracts. A problem in legal contract theory is to develop rules to guide courts in interpreting contracts. These problems are related because the law’s interpretive rules affect how parties describe what they want to trade in their contracts, and when parties will offer evidence at trial of their ex ante intentions. The more illuminating contacts and evidence production are, the better contracting works. Hence, some interpretive rules are more efficient than others. We study these rules and show, among other things, that (a) an optimal interpretive rule trades off accuracy in recovering the parties’ ex ante intentions against the costs of contract writing and evidence production; (b) an optimal rule sometimes prevents parties from introducing relevant evidence and deters some parties from writing contracts (in these cases, communicating the parties’ intentions to an adjudicator would not be worth its costs); (c) different enforcement institutions — courts and arbitrators — exist in equilibrium; (d) contract writing and evidence production are substitutes, while enforcer expertise (appropriately defined) and contract writing are complements; (e) under an optimal enforcement scheme, parties need not give interpretive instructions to enforcers because the enforcers and the parties share the same goal (to create efficient incentives to invest); and (f) current courts are more interested in accuracy than in incentives, so parties may want to send interpretive instructions to them.

1. Introduction

We begin with a description from a case:

Plaintiff Lewis has been doing business as a sawmill operator ....
In order to meet competition, Lewis decided to convert his power equipment to hydraulic equipment. He purchased a used system .... Following the installation [of the system], Lewis requested from Frank Rowe, a local

1Schwartz: Yale School of Law; Yale School of Management. Watson: Department of Economics, UC San Diego. The authors thank Jonathan Barnett, Gillian Hadfield, Dan Klerman, Shmuel Leshem, Nina Walton, and workshop participants at the American Law and Economics Association 2009 Meeting, Columbia, Duke, USC, and UC San Diego for comments. Financial support provided by the Southern California Innovation Project is gratefully acknowledged.
Mobile oil dealer, the proper hydraulic fluid to operate his machinery. The only information given to Rowe by Lewis was that the machinery was operated by a gear type pump. Rowe sold plaintiff a product known as Ambrex 810. This is a straight mineral oil with no chemical additives.

Within a few days after operation of the new equipment commenced, plaintiff began experiencing difficulty with its operation. The oil changed color, foamed over, and got hot. Approximately six months after operations with the equipment had begun, the system broke down. Ultimately, on the pump manufacturer’s and Mobile’s representative’s recommendation, a new oil was used which contained chemical additives, principally a “defoament”. Following these changes, plaintiff’s system worked satisfactorily.

1.1 The interpretive task

We present a new view of contract interpretation. On the traditional view, the court is supposed to recover the parties’ original meaning. Contract law’s “interpretive rules” determine how courts do this. The rules thus select among the oral and written terms the parties may proffer just which terms constitute the legally enforceable contract; and the rules govern which evidence is admissible to find the meaning of those terms. As an example of the former function, a party may claim that the written contract was supplemented by a prior oral understanding. The court must decide whether the enforceable contract includes the oral understanding. As an example of the latter function, a party may argue that an ambiguity in the

---

2 Lewis v. Mobile Oil Company, 438 F2d 500 (8th Cir. 1971). The Ambrex 810 oil was well made, but unsuitable for plaintiff’s equipment. The court held that Mobile breached the contract because it supplied the wrong oil. In a similar case, plaintiff purchased a herd of cattle that turned out to have “brucellosis”, a disease that did not affect the cattle’s fitness for beef but made the cattle unfit for breeding. Plaintiff purchased the cattle for breeding. The sellers were Berry, a rancher, and Pershbacker, a trader, and the issue was whether either warranted — i.e., guaranteed — that the cattle would be fit for the plaintiff’s intended use. In excusing Berry, the court said: “... this sale was the dealing in a different classification of stock than this cow and calf for resale. This was a sufficiently different type of business and type of goods than [Berry] theretofore sold. ... It is not sufficient to say that Berry had always dealt in ‘cattle’, as such a category includes too many entirely different ‘goods’”. Berry’s merchant “type”, that is, did not include the offering for sale of breeding cattle. The court remanded for findings concerning Pershbacker. See Fear Ranches, Inc. v. H.C. Berry, d/b/a Berry Ranch Co., 470 F.2d 905 (10th Cir. 1972), affirmed on rehearing 503 F.2d 953 (10th Cir. 1974).
written contract could be clarified by considering early drafts. The court must decide whether to admit those drafts into evidence. The court’s animating, and only, goal in resolving such questions is accuracy: the task is to reconstruct intent, not to encourage or discourage particular behavior.

The interpretive rules are mandatory. Parties may have preferences over how the court should find their meaning but courts choose and apply the rules independently of those preferences. Legal scholars commonly accept the two features of the current approach, that the rules should facilitate reconstructing party intent and should be mandatory. The scholars, though, sometimes divide over how much and what kinds of evidence yield more accurate interpretations. The current approach is flawed because accuracy in interpretation is a good, and goods do not come free. An economic view would ask how best to make accuracy/cost tradeoffs.

We posit an “enforcer” who resolves interpretive disputes arising from commercial contracts. Under our paradigm contract, parties agree to trade a good or service whose quality can be enhanced by the seller’s investment. A conforming performance by the seller requires that the seller’s investment match the buyer’s needs. As in the introductory case, a lubricating oil must, by definition, lubricate a subset of machines; that is, every oil must be intrinsically good for something. The key issue is whether the seller invested so as to provide the oil that fit the buyer’s machine, or whether the cattle in the case described in note 2 above are useful to the buyer, conceding that they are useful to some buyers. On this view, the enforcer’s interpretive

---

3 The state of legal scholarship is reviewed and challenged in Schwartz and Scott (2010) and (2003).

4 We refer to an enforcer rather than a court because we attempt to explain why different enforcement institutions — court, arbitrator, arbitration panel — exist.
task is to recover not the meaning of words alone, but to recover the “type” of productive relationship the parties have; he then can infer from that type what the parties intended to trade.

Parties prefer accurate interpretations because they enhance incentives. A seller thus invests more efficiently when she anticipates being paid if, and only if, she produces a performance that matches the contracting relationship’s type (as examples, to trade a particular category of oil or cattle with particular properties). Our enforcer goes a little beyond the preferences of particular parties, however. His goal is to recover party types in order to induce optimal investment by parties that constitute the total population of contractual relationships.

The enforcer’s task is illuminated by considering the information structure. Contract theory models commonly focus on information asymmetries between parties and an external enforcer regarding outcome variables — events that materialize or actions that parties take after the contract is written. For example, parties may observe a seller’s investment choice but be unable to verify that choice to a court. In contrast, we focus on information asymmetry regarding context: those aspects of the parties’ shared knowledge when they contract that the enforcer later may not know. The existence of such “contextual asymmetry” helps to explain the need to contract: if the enforcer could recover the parties’ context perfectly, he could create the contract they would have written, so parties could then delegate the contract writing job to him.\(^5\) That the enforcer is partly ignorant of context thus helps to motivate the existence of contracts, which can convey context information. The more context the enforcer knows, the better able he is to identify the parties’ type; and thus the better able he is to compare what the parties intended to trade with what the seller actually supplied.

\(^5\)Shavell (2006) also makes this point.
Turning to costs, parties can convey context information to the enforcer in three ways: by their physical performance; the contract they write; and the evidence they introduce at trial. Contract writing and adjudication are costly, however, so parties prefer interpretive rules that make optimal tradeoffs among accuracy in recovering type, contracting cost, and adjudication expense. These rules would induce parties to spend the efficient amount on contract writing and to introduce context evidence into trials only when the accuracy gain exceeds the evidence production cost. Our enforcer shares this preference: he too wants to create optimal, rather than merely accurate, interpretive rules. The model below explains how he goes about doing this.

1.2 Interpretation and contract theory

Parties cannot induce efficient sunk cost investment when payoff relevant variables are unverifiable. The verifiability problem can be a function of contextual information asymmetry. To see why, we distinguish between outcome variables that are “free standing” and those that are not. An outcome variable is “free standing” if the enforcer can evaluate it only from its designation and its properties. For example, a contract may require a seller to deliver “100 grams of iodine”. An enforcer can verify the seller’s compliance by observing the quantity and the physical properties of what the seller supplied. The enforcer need not know any more than the contract description and the physical performance.

Many outcome variables are not free standing; the enforcer cannot verify compliance by observing only a variable’s contract description and its physical properties. As an illustration, let a contract require the seller to deliver “a bottle of iodine”. The enforcer must know what the parties meant by “bottle”; that is, he must observe context. Here, Bolton and Dewatrapont’s explanation of why sellers may underinvest in quality is illuminating (2005 at 569):
“Think, for example, that were the seller to deliver the good, the buyer could always claim it is not of “appropriate quality”. If quality is unverifiable, the court can only observe whether trade took place ... but cannot distinguish who is responsible for the lack of trade.”

In this example, the court observes the failure to trade and the seller’s physical performance. The latter could be the properties of the oil or the nature of the cattle in the introductory case examples. But as in those examples, whether the seller complied with the contract cannot be inferred from the physical properties of her performance alone. Put another way, contract “quality is unverifiable” largely when the court does not know what the parties intended to trade.

Hart and Moore capture some of this idea when they say (1999 at 125):

Even though the buyer and seller know at state 1 which widget is the special widget [they intend to trade], they have no words to describe it, other than the N [large number of] names, any one of which may turn out to be appropriate at stage 3.

Using our terminology, the quality variable in Hart and Moore’s example is not free standing because an enforcer could not infer from the physical properties of the widget the seller delivered, and the name the seller attached to that widget, that the seller delivered the right widget. The quality variable becomes unverifiable in their example because the parties could not convey enough context information in their contract to permit the enforcer to evaluate the variable: that is, to know whether the widget the seller supplied was the widget the parties intended the buyer to get.

To sum this discussion up, an outcome variable materializes after the contract is made. The variable is free standing when the enforcer can evaluate it from a contract description and the variable’s physical properties. Parties contract on free standing outcome variables when they expect the enforcer later to observe them. Other outcome variables, such as product quality,
seldom are free standing. An enforcer who observes the product or service the seller supplied cannot evaluate the seller’s performance unless he also can access the relevant context variables: information about the parties when the contract was made and their world that permit the enforcer to recover the parties’ type. The variable is unverifiable when he cannot access this information. This discussion supports two observations: (i) In many cases, the verifiability problem and the description problem reduce to the same problem; (ii) The existence of contextual information asymmetry helps to explain both why parties contract and contracting’s limits. In turn, the law’s interpretive rules deserve study because they influence how much context information enforcers get, and how costly it is for parties to provide that information.6

1.3 Model and results

We model contractual relationships in which a buyer and a seller contract to perform a productive action such as providing a service or producing a product. The relationship’s type represents its contextual information. The seller chooses an effort level, which is private information. The seller’s performance is observed by everyone. A dispute arises when the buyer asserts that the seller delivered the wrong or an insufficient quality. The enforcer’s role is to determine whether the seller’s performance matches the parties’ type. If the match is sufficiently

6Evaluating outcome variables that are not free standing poses a common problem in Sales Law. Section 2-314 of the Uniform Commercial Code provides that a seller’s performance is compliant if it “passes without objection in the trade” or is “fit for the ordinary purposes for which it is used”. The former criterion requires the court to identify the goods that types in the relevant industry commonly trade, and to see whether the parties before it are an industry type. For example, a certain type of grain is acceptable to a maker of animal feed but not to a maker of bread. A seller who delivers that grain is in compliance if the parties are a “feed type”. The latter fitness criterion requires the court to locate the set of intentions that parties in the relevant industry attempt to realize, and then to find whether the seller’s performance realized one of those intentions. For example, grain intended for baking must rise to a certain level when mixed with yeast. The seller delivers grain with particular chemical properties. If the parties are a “baking type” — they have a baking intent — the court can find whether those properties match baking requirements. Returning to the first criterion, if the product in question satisfies common trade intentions but the buyer rejects it, the buyer may have had a special intention. If so, the seller’s performance need not suit that intention unless she should have recognized it. See UCC §2-315.
close, the enforcer rewards her for a compliant performance. The probability that the seller actually complied is increasing in the seller’s investment, and the seller’s belief that she will be paid for a compliant performance is increasing in the enforcer’s accuracy. A contract therefore cannot give the seller good incentives unless it and other information the enforcer can access will resolve much contextual asymmetry.

The enforcer’s interpretive resources consist of the pleadings in a law suit, the enforcer’s general (or industry specific) knowledge; the contract; the performance the seller tenders; and additional context evidence when he permits litigating parties to introduce it. The pleadings identify the parties and provide some context regarding their deal. The enforcer uses this information, together with his general knowledge, to create an initial partition of the type space in which the parties may be located. The contract functions as an additional signal of the parties’ type, which the enforcer uses to refine his partition. A contract signal’s quality depends on the informativeness of its language, in light of other information that the enforcer has. The seller’s performance is a third signal of type.

To see what is meant, let the enforcer’s initial partition include party types that trade Class A to Class F widgets. The contract may cause him to refine the partition to include only types that trade Class A or Class B widgets; then delivery of either widget type would be a compliant performance. Suppose, then, that a Class B widget is delivered. This leaves the enforcer’s partition unchanged because performance is consistent with “Class A or Class B.” The, buyer, however, may be able to provide context evidence to show that the parties’ type is such that only a Class A widget is compliant. An optimal interpretive rule admits such evidence when the gain in further refinement is worth its costs. Finally, the enforcer compels transfers
between the parties on the basis of his updated belief about the parties’ type. In this illustration, the enforcer orders the buyer to pay if he does not admit evidence, but he will let the buyer off if the enforcer admits the buyer’s evidence and it is persuasive.

We note two methodological points. First, there is an interaction among these interpretive resources. For example, the better able the enforcer is to use his knowledge to infer type from the pleadings and the performance, the less parties have to invest in contract writing and evidence production. Second, when parties contract, they know their context and the enforcer’s interpretive rule. Hence, parties can predict where in the contracting “type space” the enforcer will put them given the agreements they write, the performances they render and the evidence they are permitted to introduce. It is the ability to predict what the enforcer will do that permits parties to make efficient decisions regarding which contract terms to create, what evidence to produce, and which investment level to choose.

Regarding results, we explore trade-offs among the three information channels the enforcer uses to recover a relationship’s type: the contract, contextual evidence, and the enforcer’s expertise. These channels are costly to exploit (or in the case of expertise, exogenously limited), so the optimal interpretive rule will emphasize or de-emphasize a particular channel as a function of the costs of all of the channels. This analytic method yields some simple comparative statics.

1) An optimal interpretive rule induces parties to make optimal investments in contract writing, investment levels, and litigation. Conversely, a suboptimal rule can induce inefficiencies in all three areas. This is why interpretation matters.

2) It is sometimes optimal for the enforcer to disallow context evidence. The enforcer
optimally shuts down the evidence channel (that is, prevents parties from introducing context evidence) when the marginal increase in enlightenment from further narrowing the parties’ position in the type space is not worth the trial cost.

3) Multiple enforcers exist in equilibrium because the choice of enforcer is a signal of type. The phrase “we will use arbitration” thus can substitute for the context informing effect of contractual “whereas” clauses. The existence of multiple enforcers with different interpretive styles thus expands the language in which contracts can be written.

4) Enforcer expertise and contract language are related, both as substitutes and as complements. As substitutes, raising the external enforcer’s technical sophistication — his baseline ability to differentiate types, and his ability to evaluate evidence — reduces the parties’ need to signal their type through the contract; then the parties optimally use a coarser contractual language. A parameter shift that lowers the cost of evidence production has the same effect. As complements, the better able the enforcer is to evaluate the seller’s performance (which we call probatory depth), the more productive contract descriptions and context evidence become.

5) Parties in our model have no need to send interpretive instructions to an enforcer regarding how he should find the contract’s terms and which evidence categories he should consider at trial. Interpretive instructions are unnecessary here because parties and enforcers share the same goal: to maximize welfare by creating efficient investment incentives. Thus, our enforcer admits context evidence when parties would want it admitted but otherwise not. Actual parties, however, may prefer to give interpretive instructions to courts because contract law’s

---

7 An interpretive instruction may direct the court not to consider such evidence as prior contract drafts or pre-contractual conversations, but to interpret technical words according to the customs of the parties’ industry. Because the interpretive rules are mandatory, courts do not feel themselves bound by such instructions. Hence, judicial interpretation today is largely court controlled, not party controlled.
current interpretive rules do not attempt to help parties maximize expected surplus, nor are
courts as expert as are our enforcers at making cost/benefit tradeoffs.

There are few analyses of contract law’s interpretive rules in the economics and law-and-
economics literatures. Our modeling exercise takes as an input the prior work on costly
contracting and limits on describability (Dye 1985, Anderlini and Felli 1999, Battigalli and
Maggi 2002, Schwartz and Watson 2004). This paper is most closely related to Shavell (2006).
Certain core elements of our model are in his (both having contextual asymmetry and a limited
language). There are also significant differences between our studies. Our model includes
evidence production and expertise, and we focus on the trade offs between language, evidence,
and expertise; on the comparison and interaction of multiple enforcers; and on the various
elements of language costs and welfare consequences. Shavell does not analyze these issues.
We also differ regarding the nature of meaning in contract. Shavell assumes that words have
exogenously defined “literal meanings” and some of his results follow from this premise. In our
model, most payoff relevant variables are not free standing: that is, their meaning turns on the
category variables that the enforcer is able to recover.8

2. The Model with a Single External Enforcer

There is a population of contractual relationships consisting of buyers and sellers. Each

8Shavell’s premise leads him to conclusions such as that “opt-out” clauses (that would prohibit the external
enforcer from using certain evidence) should be enforced verbatim. We show that whether interpretive instructions
should be obeyed — indeed, whether they will be given — is a function of what the enforcer is maximizing and how
he behaves. Another related paper is Maggi and Staiger (2008), which develops a model along the lines of Shavell’s
in the context of international trade agreements. Maggi and Staiger also assume that aspects of the enforcement
regime are exogenously fixed. See also Hadfield (1994) and Posner (2005) for a general discussion of issues in
interpretation of contracts, especially on the notions of gap-filling and judicial error.
relationship is distinguished by its type \( t \in [0, 1] \), where the interval \([0, 1]\) is the space of types and the type distribution is uniform. The type variable represents characteristics of an individual relationship (the seller is an oil company; the buyer needs a particular kind of lubricating oil for his pump). The seller agrees to perform a service or produce a product for the buyer. She chooses an action \( q \in [0, 1] \) where \( q \) is the seller’s effort level or investment. Performing at level \( q \) costs \( c(q) \). The seller’s performance creates a high value for the buyer with probability \( q \), and a low value with probability \( 1 - q \). A high value is worth 1; a low value is worth zero. The seller privately knows \( q \) and the buyer privately observes the value that performance has for him.

With transferrable utility, a relationship’s 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 \( \lim_{q \to 1} c(q) = \infty \). Let \( q^* \) denote the solution to this maximization problem: \( c'(q^*) = 1 \).

The seller’s performance is not free standing. The enforcer, that is, cannot know whether her performance complied with the contract without knowing the parties’ type. To capture this problem formally, the performance is a public signal \( x \) (public because both parties and the enforcer observe what the seller did). We call \( x \) the productive outcome of the relationship. To simplify the relationship between the signal and the type — i.e., between \( x \) and \( t \) — we let the space of productive outcomes be coextensive with the space of types: \( x \in [0, 1] \). When the seller’s performance has a high value, then \( x = t \) with probability \( s \in [0, 1] \); with complementary probability \( 1 - s \), \( x \) is drawn from a uniform distribution over \([0, 1]\) (i.e., \( x \) is unilluminating).
When the value is low, $x$ also is drawn from a uniform distribution over $[0, 1]$.

The $x$ and $t$ parameters jointly signal the value of the seller’s performance to the buyer. That is, $x = t$ is a positive signal of high value to the buyer, whereas $x \neq t$ is a negative signal. Thus, the enforcer must have some knowledge of the relationship’s type $t$ to help him infer the buyer’s value from the productive outcome. The precision of the signal is increasing in $s$.

A dispute may arise after the seller performs that causes one of them to invoke the enforcer. His task is to compel a monetary transfer between the parties as a function of the information that he acquires about their relationship. This information derives from three signals of the parties’ type and possibly evidence: (i) the signal that represents what the enforcer can infer from readily available context data; (ii) the signal the contract sends; (iii) the productive-outcome signal $x$ that the seller’s physical performance sends; and (iv) evidence about context that the enforcer permits a party to submit. Items (i), (ii) and (iv) help the enforcer to resolve contextual asymmetry; item (iii) does as well, but also permits the enforcer to decide whether the seller rendered a compliant performance.

Contracts are costly to write. The space of feasible contracts is the set of positive integers, $P = 1, 2, 3, \ldots$. For each $k \in P$, contract $k$ costs parties $y_k$ to form. Contracts are ordered so that $y_k$ is increasing in $k$, and we let $y = (y_1, y_2, \ldots)$. Here, $k = 0$ represent “no

---

9The distribution of $x$ treats pairs of types symmetrically: that is, if type $t$ gets an outcome $x \neq t$ then the type is equally likely to be any other type. In other words, each type is equally close to every other type in the distribution of $x$. Although this extreme form of symmetry is not particularly realistic, it helps to keep the model tractable.

10The source of the dispute does not matter in our model. The buyer can claim that quality was too low or the seller could claim that the buyer rejected the appropriate quality.

11This signal includes the pleadings in a law suit, and knowledge the enforcer may have acquired about deals in general or the parties’ industry in particular.
contract”: a potential buyer and seller reject a relationship.\textsuperscript{12}

Contracting cost can become large in $k$. Thus, contracts may not signal type perfectly.

To be precise, for every $\kappa$ there is a $k$ such that $y_k > \kappa$. When this condition is satisfied, then in equilibrium some contracting types will not differentiate themselves by their contract choice: that is, multiple relationship types choose the same contract $k$. The real-world analogue occurs when different relationship types use contracts that are sufficiently similar so that the enforcer cannot distinguish the relationship before him from other relationships by observing the contract alone.\textsuperscript{13}

For convenience, and with some realism, we assume that the enforcer initially observes information category (i) above: readily available context data, which the enforcer uses his knowledge to evaluate. His inferential process is modeled as a partition $\Lambda$ of the type space $[0, 1]$. The enforcer, that is, initially does not observe the type $t$ directly, but rather he observes a particular subset $\Lambda(t) \subseteq [0, 1]$ that contains $t$. Because $\Lambda$ is a partition, we have that $t' \in \Lambda(t)$ implies that $\Lambda(t') = \Lambda(t)$. For example, if the partition is given by $\Lambda = \{[0, 1/2], [1/2, 1]\}$ then, for any relationship type $t$, the enforcer could use category (i) information initially to decide whether $t < 1/2$. Thus, $\Lambda(t) = [0, 1/2]$ for every $t < 1/2$; and $\Lambda(t) = [1/2, 1]$ for every $t \geq 1/2$. The enforcer next observes the contract and then the seller’s performance (categories ii and iii above). He uses this data to update — that is to refine — the partition $\Lambda$.

\textsuperscript{12}For simplicity, we sometimes use the word “language” to refer to contracting costs, where “greater language” means lowering some components of $y$.

\textsuperscript{13}In a special case, there is a set $1, 2, \ldots, K$ of “common” contracts that the external enforcer is familiar with and that are relatively easy to specify. We might then assume that $y_k = 0$ for $k \leq K$. The question would then be whether a contractual relationship would have the incentive to expand the set of contracts in use by selecting contract $K + 1$ at some positive cost.
Finally, we model evidence by assuming that, at cost $\gamma \in [0, 1]$, the buyer can submit evidence that perfectly reveals his relationship type $t$.\(^{14}\) The buyer’s incentive to provide evidence depends on the seller’s performance — the signal $x$ — and on the other information that the enforcer has.

Three elements of the model represent the enforcer’s expertise. As just suggested, the partition $\Lambda$ represents what the enforcer can readily infer about context. A more refined partition — one that includes fewer party types — implies greater expertise; that is, if $\Lambda'$ is a refinement of $\Lambda$ (mathematically, each element of $\Lambda'$ can be written as the union of a set of elements in $\Lambda$), then $\Lambda'$ represents a greater level of expertise than does $\Lambda$. The probability $s$, which represents the enforcer’s ability to infer whether the seller’s performance had a high value to the buyer, also is an aspect of enforcer expertise. To see why, suppose the seller supplied high value. Then $x = t$ with probability $s$. The greater is $s$, the more likely the enforcer is correctly to find that the type matches the productive outcome; i.e., that the seller rendered a compliant performance. Hence, enforcer expertise is increasing in $s$. Finally, the parameter $\gamma$ represents expertise because less evidence is needed to reveal type to an expert enforcer. Hence, the lower is $\gamma$, the greater is the enforcer’s expertise. We refer to expertise in reference to $\Lambda$ and $\gamma$ with the term technical sophistication, and we refer to expertise in reference to $s$ with the term probatory depth.

The enforcer’s interpretive rule maps the enforcer’s information ($k$, $\Lambda(t)$, $x$, and the type $t$ if the buyer produced evidence) to the transfer $m$ the enforcer orders. We use two functions, $f$ and $g$, to describe the rule. Function $f$ gives the enforcer’s choice of transfer if evidence reveals

\footnote{We later show why only buyers submit evidence in the model. In our model, context evidence introduced at trial reveals a relationship’s type with certainty. If evidence only revealed type with positive probability, then since evidence is costly an optimal interpretive rule would preclude admission in more cases than the rule we derive below. The model’s qualitative results would not change, however.}
so \( m = f(x, k, t) \). Function \( g \) gives the enforcer’s choice of transfer without trial evidence: \( m = g(x, k, t) \), with the additional assumption that the function \( g \) is measurable with respect to \( \Lambda \) as a function of \( t \). The measurability requirement for \( g \) represents the enforcer’s limited direct information about the type. We assume that the enforcer can commit to his interpretive rule, and chooses it to maximize welfare over the population of contractual relationships.

The timing of interaction is as follows:
1. The enforcer chooses an interpretive rule \((f, g)\), which he makes public.
2. A buyer and seller meet and learn their type \( t \). They select a contract \( k \) and jointly pay the cost \( y_k \). They divide the expected surplus through the price.
3. The seller chooses \( q \) and incurs cost \( c(q) \).
4. The seller performs. The buyer accepts performance and pays the contract price or there is a dispute.\(^{15}\)
5. If the parties disagree, the buyer decides whether to provide evidence of \( t \) at cost \( \gamma \).
6. The external enforcer sees the dispute documents and creates the partition \( \Lambda(t) \).
   He then observes \( k \) and \( x \). If the enforcer permits evidence and the buyer provides it, then the enforcer also observes \( t \).
7. The enforcer compels a transfer \( m \) from the buyer to the seller.

We note two things about this scenario. First, though the seller knows what she did, she may not know whether the buyer regards what she did as suitable. If there is a dispute, the seller

\(^{15}\) The buyer may reject the performance as nonconforming. He then may sue for the loss the allegedly noncompliant performance imposed on him, or the seller may sue for the loss rejection caused her. The buyer also may accept but offer less than the price to compensate for the loss of a supposed noncompliant performance. The seller may then sue for the remainder of the price. In all of these cases, the issue is whether the seller’s performance complied with the contract. Since the seller’s performance is not free standing, the enforcer cannot resolve this issue without context information.
will learn more about the buyer’s view. Second, the formal model supposes that disputes are resolved at a trial, but the results hold if there can be settlement. To see why, suppose the seller initiates litigation. The buyer then sees the pleadings and he has previously observed the contract and the seller’s performance. Since the buyer knows the interpretive rule that the enforcer will apply, he can predict what the enforcer will do when he sees or does not see additional context evidence. The buyer cannot produce evidence at trial unless he reveals that evidence in discovery proceedings. Hence, the buyer will incur material evidence production costs, when he wants to produce evidence, whether the dispute settles or not. The anticipation of these costs affects the buyer’s decision of how much to invest in contract creation. Thus, the basic logic of the model holds whether we permit settlement or not.

The buyer has an incentive to provide evidence when the cost is less than the difference between the transfer the enforcer orders when he is relatively uninformed about type and the transfer he orders when he is perfectly informed: that is, when

\[ f(x, k, t) + \gamma \geq g(x, k, t). \]

The left-hand side is the transfer that an enforcer who knows the relationship’s type compels plus the buyer’s cost of informing the enforcer. The right-hand side is the transfer the enforcer compels when the buyer does not want to introduce context evidence. When the enforcer permits evidence production is analyzed below.

For simplicity, we restrict attention to the case in which, without trial evidence, the interpretive rule separates the outcome space into two sets, with some arbitrary transfer \( m^0 \) for one set and the transfer zero for the other. Recall that a high value performance yields one for the buyer. Penalties — 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. Hence, $m^0$ is bounded from above by 1.\textsuperscript{16}

In equilibrium, (i) each relationship forms the contract that maximizes its value, given the enforcer’s interpretive rule; (ii) the seller behaves optimally in choosing $q$ and the buyer behaves optimally in deciding whether to provide contextual evidence at trial; and (iii) the external enforcer selects an interpretive style to maximize the total value of the population of contractual relationships.

To make the analysis more concrete, and to preview where we are heading, consider a contract that requires “prompt delivery” of a piece of medical equipment. The seller delivers in ten days and the buyer rejects the equipment as not timely. The outcome variable “delivery time” is not free standing because it requires context to evaluate. A ten-day delivery would be a breach if the equipment were needed for urgent care but may be compliant if the equipment was to be stocked for future use. The enforcer first creates the partition $\Lambda(t)$. If the enforcer knows the industry, and parties in it usually require almost immediate delivery of the equipment at issue, $\Lambda(t)$ will comprise only such types. The enforcer then refines his partition in light of whatever context information the contract contains. Then he observes the seller’s performance: when the equipment was delivered. If the enforcer continues to believe that almost immediate delivery was required, he will find that $x \neq t$. The parties before him are a very fast delivery type; a ten day delivery time does not match that type. Therefore, the enforcer orders the buyer to make a transfer of zero. If, however, the enforcer lacks industry knowledge, then his $\Lambda(t)$ may include party types that do not require prompt delivery. If the parties at issue anticipate being

\textsuperscript{16}The assumption that transfers are bounded is not critical. More general versions of the model without a transfer constraint have features similar to those discussed here but are more complex to analyze.
before such an enforcer, they may write a more extensive contract, that more fully discloses their type. Were the enforcer’s partition nevertheless erroneously to include party types for whom a two week delivery would be compliant, he would order the buyer to pay. The buyer then has an incentive to introduce evidence that the parties at issue are not type $t$ (i.e., that a ten day delivery is not compliant).

3. Characteristics of the Optimal Interpretative Rule

We assume that the enforcer seeks to maximize aggregate welfare over the population of relationship types. In this section, we characterize how the enforcer’s interpretive rule influences the incentives of the seller to provide effort. We analyze the conditions under which it is optimal to allow context evidence. The section concludes with a description of the enforcer’s optimal interpretive rule as the solution to a “planner’s problem.”

3.1 The enforcer’s goal and the parties’ behavior

The enforcer’s goal is give the seller an incentive to choose the optimal effort level. The enforcer must order a higher transfer when value is high than when value is low, but he cannot directly observe value: whether the seller complied. Hence, the enforcer must create a margin on the basis of the signal $x$ (the seller’s observable physical performance) and what he infers about the relationship’s type. The seller must be rewarded for values of $x$ that are associated with high value and not rewarded otherwise.

To see how the enforcer proceeds, recall that the enforcer initially distinguishes among different elements of the partition $\Lambda$ by applying his experience to publically available information. Parties will not use the contract or supply evidence in order to differentiate
themselves across partition elements that the enforcer can create on his own. Hence, we analyze individual elements of the partition \( \Lambda \) independently of the analysis of other elements. For example, if the enforcer can determine whether \( t \in [0, 1/2] \) or \( t \in [1/2, 1] \), then a type in \([0, 1/2]\) is concerned only to differentiate itself from other types in this set. Relationship types, however, cannot differentiate themselves arbitrarily finely through contract selection because contract costs are increasing in contract informativeness.

Given separability across elements of \( \Lambda \), we restrict attention to a single element of \( \Lambda \), denoted \( T \). We write \( T = \Lambda(t) \) where \( t \in T \). For instance, in the above example we could have \( T = [0, 1/2] \). Letting \( |.| \) denote cardinality (size), and noting that \( |[0, 1]| = 1 \), we see that \(|T|\) is the fraction of types in this element of the enforcer’s information partition. In the example above, given the assumption of a uniform distribution we have \(|[0, 1/2]| = 1/2\).

\( T_k \) denotes the subset of types in \( T \) that select contract \( k \), as pictured below. We calculate, for a fixed \( T, k, \) and \( T_k \), the optimal specification of \( f(\cdot, k, t) \) and \( g(\cdot, k, t) \) over all \( t \in T \). That is, we calculate the optimal interpretive rule restricted to \( T \) and \( k \) under the assumption that \( T_k \) is fixed. Although the enforcer’s interpretive rule influences which types select contract \( k \), as is explicated below, we show next that the optimal interpretive rule maximizes the total value of the relationships in \( T_k \) holding fixed this set.\(^{17}\)

\(^{17}\)If in equilibrium the enforcer is not maximizing as described, then there is another interpretive rule that improves the total value for types in \( T_k \) under the assumption that this is exactly the set of types that select contract \( k \). Some types in this set may now prefer to use a different contract and types outside of this set may want to switch to use contract \( k \). These adjustments will increase aggregate value.
Given the limitations we have imposed on the interpretive rule, two logical steps simplify analysis. Because $T$, $k$, and $T_k$ are assumed fixed for now, we write $g$ as a function of $x$ only.

Also, because $f$ is the enforcer’s interpretive rule when he sees trial evidence, we write $f$ as a function of $x$ and $t$. Let $S$ denote the set of outcomes for which the enforcer compels transfer $m^0$ in the absence of evidence. That is, we have $g(x) = m^0$ for all $x \in S$, and we have $g(x) = 0$ for all $x \notin S$. We need to determine $m^0$, the set $S$, and the optimal values of $f(x, t)$ for all $x$ and $t$.

Conditional on the parties’ type being in set $S$, a performance that generates the signal $x \in S$ is “good news”. To see this, note that if the seller’s physical performance $x$ also is in $S$, it raises the enforcer’s posterior belief that $x = t$; that is, the performance reinforces the enforcer’s belief that the seller complied with the contract. On this logic, any $x \notin S$ is “bad news” because it reduces the posterior belief that $x = t$.

Because outcomes in $S$ lead to the transfer $m^0$ in the absence of evidence, sellers whose relationship types are in the set $S$ are given an incentive to exert effort. Increasing the effort level raises the probability that $x = t$, thus raising the probability that $x \in S$ (in which case the seller gets the reward $m^0$). In contrast, sellers whose types are not in $S$ lack this incentive.
because increasing their effort cannot increase the odds that \( x \in S \). Since only types in \( T_k \) selected contract \( k \), the effort incentive is tailored to these types. Thus, we know that \( S \subset T_k \) is optimal for the enforcer. Furthermore, since it is costly to contract, relationships with types outside of \( S \) will prefer no contract to forming contract \( k \), and thus we expect that \( S = T_k \) will prevail.

Next, observe that the specification of \( f(x, t) \) affects the incentives only of a seller in a type \( t \) relationship. Thus, we can determine the optimal interpretive rule \( f(x, t) \) by calculating the value of this relationship. Recall that the buyer provides evidence when the outcome is \( x \) and

\[
f(x, t) + \gamma \leq g(x) = m^0
\]

Since \( x = t \) signals high value, the seller’s effort incentive is enhanced by rewarding her for generating this signal. Thus, when \( x = t \), the transfer \( f(t, t) \) should be big. The buyer will not provide evidence to increase the amount he must pay to the seller, however, so any transfer \( f(t, t) \) that is above \( m^0 - \gamma \) cannot generate additional context evidence. Therefore, the enforcer should set \( f(t, t) = m^0 \). The buyer has an incentive to provide additional context evidence only when \( x \in S \setminus \{ t \} \). In this case, the parties’ relationship is not of type \( t \) but the enforcer, without this evidence, would believe \( x \in S \) is good news and so would order the buyer to make the transfer \( m^0 \). By providing evidence, the buyer shows that \( x \) is actually bad news (because it does not equal \( t \)).

The question then is whether the interpretive rule gives the buyer a sufficient incentive to

\(^{18}\)To see more clearly what is claimed, recall the medical equipment example above. The enforcer may believe that parties who wrote the contract at issue commonly require delivery within two days. Let \( S \) denote these relationship types. The enforcer then observes that the seller delivered within two days; that is, the signal \( x \) is in the set \( S \). Hence, the enforcer will compel a transfer that rewards the seller for prompt delivery without requiring the type of evidence that would be introduced at a trial. On this reasoning, any \( x \notin S \) (i.e., delivery after two days) is bad news. Further, delivery in four rather than eight days will not help the seller.
provide evidence when appropriate. The transfer \( f(x, t) \) should be low enough to satisfy

\[
f(x, t) + \gamma \leq g(x) = m^0.
\]

When this holds, the buyer provides evidence so it is optimal to set \( f(x, t) = 0 \). When the buyer proves at trial that \( x \) incorrectly signals his type — the seller has not performed — the buyer should pay nothing. When the inequality \( f(x, t) + \gamma \leq m^0 \) does not hold, the buyer will not provide evidence so the exact value of \( f(x, t) \) does not matter. Therefore, we need to consider two possibilities: \( f(x, t) = m^0 \), in which case additional evidence is disallowed; or \( f(x, t) = 0 \leq m^0 - \gamma \), in which case evidence is admitted.

### 3.2 When context evidence should be allowed

To begin, suppose that \( x \neq t \) and evidence is admissible. The buyer produces evidence when \( x \in S \setminus \{t\} \) to avoid paying \( m^0 \). The seller’s expected payoff from choosing \( q \) thus is

\[
qsm^0 - c(q),
\]

The seller receives the transfer \( m^0 \) with probability \( q \) when she exerts high effort and when the outcome correctly matches her relationship’s type (with conditional probability \( s \)). The seller receives no transfer otherwise because \( x \) is not in \( S \) or because the buyer provides evidence indicating that \( x \neq t \). The seller maximizes her expected payoff by choice of \( q \), which yields the first-order condition:

\[
sm^0 = c'(q).
\]

Let \( q^E \) denote the solution. Types in \( T_k \setminus S \) get a value of zero under this interpretive rule (because their sellers have no incentive to exert positive effort). Each relationship whose type is in \( S \) gets an expected joint value gross of contracting cost of

\[
v^E = q^E - \gamma(1 - sq^E)|S| - c(q^E).
\]
The superscript $E$ denotes the case where evidence is admissible. The first term on the right-hand side is the value of the seller’s performance. Regarding the second term, recall that $|S|$ is the size of the set for which the enforcer orders the buyer to make the high payment without trial evidence. The variable $\gamma$ is trial evidence cost and the term in parenthesis is the probability that the enforcer is mistaken. Hence, the second term is the expected cost of evidence production (the buyer produces evidence to correct mistakes). The third term is the seller’s effort cost. Because $q^E$ is below the efficient effort level, $v^E$ is increasing in $q^E$. Combining this with $q^E$ increasing in $m^0$, it is optimal to require the highest possible transfer: $m^0 = 1$.

Next suppose that the external enforcer does not allow evidence: $f(x, t) = m^0$ for $x \in S \setminus \{t\}$. The seller’s expected payoff from choosing $q$ is then

$$qsm^0 + q(1 - s)|S|m^0 + (1 - q)|S|m^0 - c(q).$$

The first term is the seller’s payoff when value is high and the outcome is $x = t$. The second term describes the case in which value is high but $x \in S \setminus \{t\}$, which occurs with probability $q(1 - s)|S|$. The third term is the seller’s payoff when quality is low and $x \in S \setminus \{t\}$, which occurs with probability $(1 - q)|S|$. The seller receives the transfer $m^0$ in all three cases. The first-order condition for the seller’s effort-choice problem is:

$$s[1 - |S|]m^0 = c'(q).$$

Let $q^N$ denote the solution. The seller exerts less effort when evidence is excluded because the enforcer is less accurate then. Further, $q^N$ is decreasing in $|S|$ because as $S$ gets larger — i.e., more dissimilar types are lumped together — the strength of the good-news signal diminishes so the seller is less likely to be rewarded for a compliant performance$^{19}$. As with the

$^{19}$ Note that $q^N$ converges to $q^E$ as $|S|$ goes to zero
case of allowed evidence, types in $T_k \setminus S$ get a value of zero under this interpretive rule (they are using a contract not meant for their types). In contrast, each relationship whose type is in $S$ gets an expected joint value of

$$v^N = q^N - c(q^N).$$

As before, increasing $m^0$ has the effect of increasing $q^N$ and thus improving welfare. So we conclude here as well that it is optimal to set $m^0 = 1$. To summarize, it is optimal to set $m^0 = 1$ whether evidence is allowed or not.

It remains to calculate the optimal $S$ and to determine when it is best to allow additional context evidence. Note that $S$ enters the expressions above only through $|S|$; it is the size of this set that matters. We thus write $v^E$ and $v^N$ as functions of $|S|$, and for convenience write $|S| = \sigma$.

As an example, $S = [1/2, 5/8]$ implies that $\sigma = 1/8$.

The enforcer’s goal is to maximize the aggregate value of relationships, which is $\sigma$ times the larger of $v^E$ and $v^N$ (because the types in $T_k \setminus S$ get zero). Hence, he optimally allows evidence when $v^E(\sigma) \geq v^N(\sigma)$. For the set of types $T_k$ that select contract $k$, welfare under the optimal interpretive rule gross of contracting cost thus is

$$w(|T_k|) \equiv \max_{\sigma \in [\sigma(0), \sigma(1)]} \left[ \sigma \max \{v^E(\sigma), v^N(\sigma)\} \right].$$

It is easy to see that $v^E$ and $v^N$ are continuous. These variables decrease in $\sigma$ and converge to zero as $\sigma$ approaches 1. Sigma becomes larger as the set $S$ expands; $S$ includes every type when $\sigma = 1$. When that many types are included in a single partition element, the types will not be relevantly similar. Hence, observing the partition is unilluminating. The analysis implies:
Lemma 1: \( w \) is continuous, with \( w(0) = 0 \). There exists a value \( \sigma \in (0, 1) \) such that \( w \) is strictly increasing for \( |T_k| < \sigma \) and constant for \( |T_k| > \sigma \).

Regarding the intuition, \( w(0) = 0 \) because an infinitesimal group of types contributes an infinitesimal amount to aggregate welfare, though the contracts of those types are interpreted perfectly. Starting from the infinitesimal group, welfare increases as more types use a particular contract because the contract is revealing so the enforcer can give these types good incentives. This “quantity effect” is nonmonotonic, however, because, when the number of types that use the same contract becomes too large, the enforcer becomes inaccurate.\(^{20}\)

3.3 Interpretation and describability

This analysis allows us to explore two aspects of the optimal interpretive rule. The first concerns which types write contracts. For a given element \( T \) of the partition \( \Lambda \), suppose that the set \( T_k \) of types in \( T \) that select contract \( k \) has a mass that exceeds \( \sigma \). Put another way, \( |T_k| > \sigma \).

As Lemma 1 implies, the enforcer focuses only on the subset of size \( \sigma \), compelling the transfer 1 if and only if the signal \( x \) is in this subset of types. He orders no transfer for any other type in \( T_k \).

If it is costly to write contract \( k \) (\( y_k > 0 \)), disfavored types thus select no contract.\(^{21}\) We have:

Lemma 2: In equilibrium, \( |T_k| \leq \sigma \) for every \( T \) and \( k > 0 \).

The second aspect of the optimal interpretive rule concerns when the enforcer allows

\(^{20}\)Note that \( q^E(0) = q^N(0) \) and so \( v^E(0) = v^N(0) \); the latter is the slope of \( w \) at \( |T_k| = 0 \).

\(^{21}\)If \( y_k = 0 \), the disfavored types are indifferent between selecting contract \( k \) and not contracting, in which case we assume that they do not contract.
context evidence. This requires a comparison of $v^E$ and $v^N$ for a given $T$ and $k$. In general, there may be values of $\sigma$ for which $v^E(\sigma) > v^N(\sigma)$, so that it is optimal to allow context evidence, as well as values of $\sigma$ for which $v^E(\sigma) < v^N(\sigma)$, so that it is optimal to disallow context evidence. Under some mild additional assumptions, we can be more precise.

**Lemma 3:** As a function of $\sigma$, $v^E$ is affine and decreasing, and $v^N$ is strictly concave and decreasing. There is a number $\sigma^E \in [0, 1]$ so 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 $\gamma$ close to 0.

**Proof:** 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.

Regarding the intuition for Lemma 3, since more evidence is better than less, more evidence should be admitted as evidence production costs fall. On the other hand, when $s$ is high the enforcer is good at inferring high value from the seller’s performance (Whether performance matches the parties’ type) so trial evidence is less useful. Finally, context evidence is more valuable when $\sigma$ is large (i.e., a relatively large number of types would be treated the same without evidence). The evidence permits the enforcer to “break down” the type space and order transfers such that sellers expect to be paid if and only if they comply.

To illustrate the results so far, consider the cost function $c(q) = q^2/(1 – q)$. For this function and a fixed value of $s$, if $\gamma$ is small enough, then $\sigma^E = 0$ so the external enforcer should always allow context evidence; otherwise $\sigma^E > 0$ and the enforcer optimally disallows context evidence if the mass of types using a given contract is small enough. In this latter case, the
enforcer can infer type reasonably well without additional context evidence. Here is a depiction of the two cases:

In the right hand Figure, the value of relationships without evidence exceeds the value with evidence when $\sigma$ is small. As $\sigma$ gets larger — more types use a particular contract — evidence becomes helpful to distinguish among relationships. We expect that $\sigma^E > 0$ is the realistic case, since $\gamma$ will typically be large.

We now can characterize the optimal interpretive rule as the solution to a simple “planner’s problem”. Given the symmetry between types in the model, we need not keep track of which specific types select the same contract; instead, we can focus on the size of the subset of types that select each contract. For each element $T$ of the enforcer’s information partition $\Lambda$ and for each contract $k$, let $\mu_k = |T_k|$. As developed earlier, the enforcer’s optimal interpretive rule leads to welfare of $w(\mu_k)$, gross of the contracting cost, for the types that select contract $k$. Implied by the function $w$ is that $\mu_k \leq \overline{\sigma}$ (Lemmas 1 and 2) and that contextual evidence is
allowed if only if $\sigma > \sigma^E$ (Lemma 3). Thus, we have

**Proposition 1:** The optimal interpretive rule and the relationships’ equilibrium behavior solve, for each $T \subset \Lambda$, the problem of selecting $\mu_0^T, \mu_1^T, \mu_2^T, \ldots$ to maximize

$$\sum_{k=1}^{\infty} [w(\mu_k^T) - \mu_k^T y_k]$$

subject to

$$\sum_{k=0}^{\infty} \mu_k^T = T$$

and $w(\mu_k^T) \geq \mu_k^T y_k$ for every $k$. In the solution to this problem, for each $T \subset \Lambda$ there is an integer $K^T$ such that contracts $k = 1, 2, \ldots, K^T$ are precisely the contracts chosen by various types in the set $T$.

**Remark 1:** The enforcer effectively chooses the types that select the various contracts because his choice determines how each subset $T_k$ forms. The exact set of types that compose $T_k$ is indeterminate; it is the mass of types included in the set that matters. The last condition of Proposition 1 ensures that each type of relationship will pay the cost of the contract meant for it.

**Remark 2:** When a potential relationship type considers whether to contract, the type knows its own context as well as the interpretive rule that the enforcer later will apply. The type thus anticipates, for every contract in the possible set of contracts, when the enforcer will order the buyer to make a transfer without evidence and when the enforcer will admit evidence. Every type realizes that the enforcer will not order a transfer under contract $k$ if $x \notin T_k$ (i.e., if he believes that the performance the seller tendered is inconsistent with the contract the parties before him used). This interpretive practice implies that sellers in relationships that do not use the contract meant for their type have no incentive to provide effort; their contract could not increase value. Therefore, an optimal interpretive rule, which maximizes over the set of types, induces each type to choose an optimal contract.

**Remark 3:** The welfare term in Proposition 1 is net of evidence costs and the second
term reflects contracting costs. Recalling Lemma 1, the optimal interpretive rule induces some potential relationships not to make contracts. The cost to these types of describing context in a contract and supplying more context at trial exceed the gains from better effort incentives. Therefore, in this model the optimal interpretive rule is second best efficient regarding describability. Parties contract only when description is cost justified, but there would be more efficient contracts if costs were lower.

Remark 4: There may be multiple optimal interpretive rules. We let $W(c, y, s, \Lambda, \gamma)$ give the aggregate welfare of the population of relationships under an optimal interpretive rule, as a function of the model’s parameters.\(^{22}\)

Remark 5: If the applicable interpretive rule is suboptimal, parties have an incentive to overinvest in contract writing (to convey more context information) and in evidence production (to correct enforcer mistakes). Hence, society’s choice of an interpretive rule has important efficiency consequences.

4. Expertise and Language Costs

We next consider how the aggregate value of relationships under the optimal interpretive rule depends on contracting costs and the external enforcer’s expertise.

Proposition 2: Social welfare $W(c, y, s, \Lambda, \gamma)$ is weakly (a) decreasing in the vector $y$; (b) increasing in $s$; (c) decreasing in $r$; and (d) increasing as $\Lambda$ becomes more refined.

---

\(^{22}\)The Appendix contains an example with bounded language size, which serves to illustrate the analysis thus far.
Proof: For each parameter shift described, it is easy to see that 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 less able to communicate their type in the contract when contracting costs increase ($y$); (b) enforcer accuracy increases as an enforcer is better able to infer type from the seller’s physical performance ($s$ is high); (c) parties supply more context evidence as evidence production cost falls ($\gamma$); (d) accuracy increases and costs fall as the enforcer is better able to infer type from public information ($\Lambda$ is a finer partition of the type space).

There is an issue, relevant to policy, regarding how contractual language and the enforcer’s expertise interact. These could be *complements*: the marginal benefit of a less costly language is increasing in the level of the enforcer’s expertise. Expertise and language also could be *substitutes*: the marginal benefit of a larger language is decreasing in the level of expertise. It is difficult to obtain results on the relation between expertise and language in a local sense (for small parameter shifts), but there is a straightforward and intuitive relation for large parameter shifts. Furthermore, probatory depth and technical sophistication behave differently in relation to language.

**Proposition 3:** For sufficiently large parameter shifts, (a) technical sophistication (measured by $\Lambda$ and $\gamma$) and improved language (i.e., lowering contracting costs) are substitutes; (b) probatory depth and improved language are complements.

Proof: 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; in fact, $q^E$ achieves this bound and $q^N$ is generally below it. Let $v^H = q^H - c(q^H)$ 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 $\Lambda$ approaches the maximally fine partition, the seller’s effort comes arbitrarily close to $q^H$, so that the relationship’s joint value comes arbitrarily close to $v^H$. The same result also obtains if contracting cost $y$ converges to zero, so that relationships could distinguish themselves finely by their contract selection. Thus, $\Lambda$ and $y$ are substitutes in the large.

Respecting part (b), that probatory depth — expertise with respect to $s$ — is a complement for large parameter shifts follows from the observation that an almost efficient level of effort $q^*$ can be induced only if $s$ is close to 1 and at least one of the other shifts just discussed occurs. Recall that $x$ is the signal the enforcer gets from the performance the seller tendered. If the seller performed appropriately (value is high), the enforcer observes $t$ with probability $s$; otherwise, $x$ is uniformly distributed over $[0, 1]$ and thus is uninformative. If the seller did not perform properly, $x$ is again uninformative. Hence, a low $s$ when the seller produces high value is equivalent to the case in which the seller produces low value. As a result, a seller cannot be given good incentives unless $s$ is high. Intuitively, $s$ is high when what the seller did is revealing as to what the parties wanted the seller to do. As a corollary, it is difficult to give a party good incentives to render a particular performance when a conforming performance tells little about what a conforming performance was supposed to be. When $s$ is high, though, welfare can be
enhanced by improving any of the other parameters. Hence, probatory depth and an improved contracting language (or a finer partition) are complements. ■

**Remark 6:** Proposition 3 shows that an enforcement system can be improved by making large improvements in the available language (lowering contracting costs) and by increasing the enforcer’s expertise. The Truth in Lending Law, for example, greatly lowered contracting cost by requiring firms to quote interest rates in the form of an annual percentage rate. Such state created formulas are public goods; the analysis here suggests that they may be under-provided. Improving enforcer expertise may be more challenging, but an enforcement system can function effectively with either an expert enforcer or a rich language; it is not necessary to have both. On the other hand, probatory depth is a necessary condition for achieving high aggregate welfare.

### 5. Multiple External Enforcers

There is a welfare benefit to having multiple enforcement systems, such as courts and arbitration panels, in our model. To see why, realize that language relating to adjudication procedure is relatively acontextual. Consider the phrase “disputes under this contract will be resolved in binding arbitration”. This phrase is free standing in our sense: an enforcer does not need context to know that, if society supplies courts and an arbitration system, the parties prefer arbitration. On the other hand, the choice of arbitration may say something about the parties’ type.

To pursue this possibility, suppose that there are two enforcers: E1 and E2. Unless they are identical in all respects, the selection of enforcer can serve as a signal of a relationship’s type.
type. When types are heterogenous, a fraction of types contract for E1 and a fraction contract for E2. Consider the case in which the two enforcers have identical levels of expertise. Let \( y \) denote contracting costs when there is a single enforcer. We can represent the presence of two enforcers with an adjusted cost structure in the basic model: \( y' \), where for each even integer \( k \) we have \( y'_k = y_{k-1}' = y_{k/2} \). In words, each of the first two contracts costs \( y_1 \) to write, the first specifying E1 and the second E2; the next two contracts cost \( y_2 \) to write, one specifying E1 and the other E2; and so on. The presence of multiple enforcers thus expands the available language by lowering contracting costs (because \( y_k \) increases with \( k \)). Welfare therefore increases.

Multiple enforcement systems always raise welfare in equilibrium, as long as \( \sigma \) is not significantly lower, and \( \gamma \) is not higher, for the entrant enforcer. How relationships divide between enforcers is indeterminate; that is, there are multiple equilibria in which the enforcers apply different interpretive rules. We summarize this analysis with

**Proposition 4:** Consider an enforcement system E1 in which the enforcer lacks perfect direct information about types (\( \Lambda \) is not the most refined partition), and in which it is costly to produce evidence. (a) For any additional enforcement system E2 that has the same values of \( \sigma \) and \( \gamma \), both external enforcers get positive shares of the relationship types in equilibrium if the enforcers use optimal interpretive rules; (b) Aggregate welfare is strictly higher than when only E1 exists; (c) E1 and E2 optimally interpret contracts differently. (proof omitted)

6. Conclusion

6.1 Summary

Many outcome variables on which parties want to contract, such as product or service quality, are not free standing. An external enforcer cannot tell whether a seller supplied the
required quality unless he can recover enough context to know what the parties intended to trade. Any external enforcer will see three context signals: a signal derived from publically available information such as the pleadings in a law suit; a signal derived from the contract; and a signal derived from what a seller actually did. The enforcer also can receive additional context evidence at a trial, if he permits parties to introduce it. Parties and the state value accuracy in adjudication because the incentive of a seller to comply with her contract is increasing in the probability that she will be paid if and only she tenders a compliant performance; and that probability is increasing in the accuracy of interpretations. On the other hand, it is costly for parties to send the three signals of type, and costly to produce evidence at trial. We show that an efficient interpretive rule optimally trades off accuracy against cost such that parties send only cost justified public signals. For example, contracts reveal optimal context information, not full context information, under the optimal interpretive rule. Further, parties want to, and the enforcer permits them to, introduce context evidence only when it is cost justified to do so.

6.2 Interpretation in courts

The enforcers in our model pursue a substantive goal: to maximize welfare over sets of contracting relationships. Real courts engage in what one of us has called “goal neutral interpretation”. See Schwartz and Scott (2009). The judge, that is, sees her function as recovering the parties’ intentions, whatever those intentions may have been. Goal neutral interpretation is consistent with efficiency because parties, under any enforcement system, write contracts to maximize expected gains. A court thus advances the parties’ goal when it recovers the contract the parties intended to write.

Nevertheless, the courts’ interpretive rules are likely to be suboptimal for two reasons.
First, the court is unlikely to give parties as good incentives as the enforcers in our model give because the court is not trying to give incentives at all. Rather, the court is only trying to find what the parties said. Thus, the set of cases in which it is optimal for the parties to have the court hear evidence and the set of cases in which the court wants to hear evidence probably are disjoint. Second, the enforcers in our model are “ideal” in the sense that they know the costs to parties of introducing evidence and of writing contracts. Real judges are unlikely to be as well informed.

This analysis suggests that the courts’ interpretive rules should be defaults. A default is a rule that parties can vary by including instructions in their contract. There is no need for parties to send interpretive instructions to an enforcer in our model. The enforcer and the parties have the same goal — to maximize welfare — and the enforcer is competent to solve the relevant maximization problem. The enforcer therefore admits evidence when admission is optimal and excludes evidence when that is best. Relationship types have no need to instruct the enforcer regarding the evidentiary base because the enforcer is their perfect agent. In contrast, courts are not the parties’ perfect agents, and thus should permit parties to vary whatever interpretive rules a goal neutral interpretive theory would otherwise imply.

Another way to say this is that, in our model, enforcers exercise full discretion in designing how contracts are to be interpreted. Language and contract terms have no exogenous meaning. Meaning arises endogenously so as to maximize aggregate welfare. By contrast, we expect that in the real world some aspects of language are exogenously fixed, especially in the court system. Thus, we expect that judicial enforcers fail to maximize social welfare. It follows that for some groups of types, a different interpretive rule would be better. If an arbitration
clause is free standing and determined by the courts to have a fixed meaning, then these party sets will either seek to use other fixed meaning terms to affect the courts’ interpretive style or they will opt out by selecting arbitration. In either case, the court’s interpretive rule is effectively a default rule.

6.3 Empirical Implications

We conclude by briefly sketching some of the model’s empirical implications and links with the literature.

A. Proposition 4 offers an explanation for the observation that interpretive rules are heterogeneous across enforcement systems (courts and arbitrators) and across jurisdictions (such as different states). It is also consistent with the observation that in many industries a positive fraction of contracts contain arbitration clauses (Eisenberg and Miller 2007, Drahozal and Ware 2010). The variations in interpretive rules across enforcement systems and jurisdictions gives contractual relationships essentially an expanded set of contracts from which to choose, and thus allows them to more precisely convey context information, so each receives an interpretive style that is tailored to its individual type.

B. Our theory explains why courts and arbitrators actively manage the admissibility of context evidence at trial. Consistent with the model, enforcers have broad rules that restrict the introduction of context evidence but allow exceptions. Chief among these are the parol evidence rule and related doctrines such as plain meaning and unconscionability.23

C. The model predicts systematic differences between the interpretive styles of courts

23Posner (1998) attributes variations in courts’ application of the parol evidence rule to differences in error-making by adjudicators and contract-writing costs, which is broadly consistent with our model. Korobkin (2003) offers a story about standard-form contracts based on sellers making strategic contract offers to boundedly rational buyers, which gives a different rationale for the unconscionability doctrine.
and arbitrators, as well as parties’ selection between them, on the basis of these enforcers’ differential levels of expertise. From the stylized fact that arbitrators tend to have greater industry knowledge, we expect parties to prefer arbitration when the seller’s physical performance is difficult (requiring significant expertise) to evaluate. To see this, recall that the enforcer in our model observes $x$ — the performance — and correctly infers that $t = x$ with probability $s$. The parameter $s$ measures probatory depth, which relates positively to industry knowledge. An interpretive rule cannot give close to efficient investment incentives unless $s$ is high, which means that arbitrators have an advantage in these settings. Hence, the probability that parties use arbitration should vary directly with the complexity of the seller’s performance, a relation found by Drahozal and Hylton (2003) and Drahozal and Ware (2010).

Furthermore, evidence also is believed to be cheaper to introduce in arbitration ($\gamma$ is low) and since probatory depth and $\gamma$ are complements, we would expect that arbitrators tend to extract more context information than do courts in these industries. On the other hand, contracts intended for arbitration should contain fewer “whereas” and definition clauses than do contracts intended for courts. These clauses communicate context information to enforcers, but the more knowledgeable the enforcer already is — i.e., the finer is the partition $\Lambda$ — the less need there is to include such information in the contract.\footnote{Hill and King (2004) claim that enforcement in the German legal system facilitates shorter, less detailed contracts than occur in the U.S., owing to a more adversarial contract-drafting process in the U.S. Our model suggests investigating different stories, such as one based on differences in technical sophistication between courts in the two countries.}

D. Also consistent with the model is the existence of standard-form contracts, on which multiple types of relationships pool. Such contracts are easily recognized and classified by
courts. Form contracts may be disfavored in relationships with complicated technologies, so from point C above we expect that the choice of arbitration and the use of standard-form contracts may be negatively related. Eisenberg and Miller (2007) provide some evidence along this line.

E. In reality there may be variations across enforcers in the degree to which enforcement rules are optimally designed in the sense of our model. If real enforcers fail to provide the optimal interpretive rules that our model prescribes, we should see relationships differentially favor jurisdictions and enforcers whose rules are closer to optimal. More generally, our model leads us to expect that contracts with arbitration clauses should contain fewer instructions and specifications. This is both because arbitrators may have more expertise than do courts and because arbitrators are more likely than courts to share the parties’ contracting goals.

On all of these points, our model provides some logic to help organize empirical work. Further theoretical analysis would also be useful.

**Appendix: A Simple Example with Bounded Language Size**

Consider a case with a bound on language so that \( y_k = 0 \) for every \( k = 1, 2, \ldots, K \), for some positive integer \( K \), and where \( y_k \) is very large for \( k > K \). We can call \( K \) the *size of the language*. Then it is optimal to have the population of relationships utilize all \( K \) of the low-cost contracts to

\[ \text{39} \]
distinguish themselves. If we further assume that $\gamma$ is close to zero — introducing evidence is cheap — then it is optimal to allow evidence ($\sigma^E = 0$). Concavity of the function $w$, which follows from concavity of $v^E$, then implies that it is optimal to have $\mu_1^T = \mu_2^T = \ldots = \mu_K^T$. In words, the sets of types that select the various contracts are, in equilibrium, the same size. If $K\sigma < |T|$ then a mass $|T| - K\sigma$ of types in $T$ will not contract (that is, they will choose $k = 0$) and we have $\mu_k^T = \sigma$ for every $k = 1, 2, \ldots, K$. Since evidence would be allowed, it is the language that prevents these types from describing what they would like to trade. Otherwise, all types in $T$ form contracts and $\mu_k^T = |T|/K$ for every $k = 1, 2, \ldots, K$. This in the top figure on the following page.
The case of $\sigma = 0$.

The case of $\sigma > 0$. 

If $\sigma^E > 0$ ($\gamma$ is not sufficiently close to zero), the function $w$ is not concave because it is the upper envelope of $v^e$ and $v^N$ as shown in the right frame of the Figure above. Here, depending on $|T|$ and the location of $\sigma$, the optimal interpretive rule may specify that contextual evidence is allowed for some contracts but disallowed for others. This would arise, in particular, if $|T|/K$ is close to $\sigma^E$. The bottom figure above illustrates this case.
References


February 25, 2010