Vague Terms: Contracting when Precision in Terms is Infeasible

by
Benjamin E. Hermalin*

May 17, 2007

This article considers the consequence of incomplete contracts that arise due to difficulties in precisely describing potentially relevant contingencies. Unlike much of the literature, this article concludes that the resulting incompleteness could often be immaterial with respect to economic efficiency. Indeed, attempts to "improve" matters by increasing the accuracy with which the courts determine which events occurred can, in fact, be welfare reducing, casting doubt on the view that greater judicial accuracy is always beneficial (JEL: K12, K41 D86)

1 Introduction

Precision in language is often difficult to achieve, particularly with respect to matters that are difficult to quantify. While one can, for instance, be fairly precise about how much paint is to be used painting a house, it is far more difficult to specify the quality of the job. An instruction, as in a contract, to provide a "quality paint job" lacks precision insofar as reasonable people could dispute whether the way a house was painted did or did not constitute a quality job. And while greater precision can sometimes be achieved through providing more detailed instructions, greater detail entails greater costs in

*Research support provided by funds from the Thomas & Alison Schneider Distinguished Professorship in Finance and from the University of California, Berkeley's Committee on Research.

the form of time spent spelling out and absorbing these instructions (what are known in the contract literature as writing and reading costs). As a consequence, numerous exchanges occur under terms that are, to a greater or less degree, vague.

Vagueness of language, the difficult of precision, and consequent writing and reading costs are frequently invoked in the economics literature as justifications for assuming incomplete contracts (see HERMALIN, KATZ, AND CRASWELL, in press, for a partial survey and discussion).[1] A few authors have sought to model explicitly how these issues lead to incomplete contracts (see, *e.g.,* DYE, 1985, KATZ, 1990, ANDERLINI AND FELLI, 1994, and RASMUSEN, 2001). What has been less examined is whether the resulting incompleteness is material—do the parties to the contract do less well than they would were a complete contract feasible?

As this article seeks to demonstrate, the answer to this question could often be "no"; that is, the incompleteness due to the difficulties of precision in language need not preclude achieving the same result as the parties would absent such difficulties. This answer, in turn, raises the question of just how imprecise a contract can be and still achieve the same outcome as a complete contract?

The point of departure for this article is a straightforward contracting situation in which one party wishes to employ a second to do a task on behalf of the first. For instance, I wish to employ a handyman to paint my house. Although the vagueness of language could, in theory, make it difficult for me to reach a meeting of the minds with the handyman over what constituted a satisfactory paint job, I will assume here that the parties to the contract can achieve a meeting of the minds. There are two justifications for my doing so. First, in the literature on incomplete contracts, the parties to the contract are assumed—at least implicitly—to be in agreement about what the contract says. Hence, assuming a meeting of the minds here is in keeping with that literature. Second, in the model that follows, it wouldn't matter if the handyman didn't interpret the contract the same way I did provided I could predict how he would respond to it.[2]

[1] Other explanations also exist for contractual incompleteness including bounded rationality, complexity (SEGAL, 1999), asymmetric information (SPIER, 1992), and verification costs (TOWNSEND, 1979). Section 4 of HERMALIN ET AL. [in press] provides a general survey on incomplete contracts.

[2] Formally, because one party makes a take-it-or-leave-it offer to the other, the contract proposer (the principal) merely needs to know how the other party (the agent) will respond

The contract between the parties may lack precision (be incomplete) insofar as it does not fully differentiate among the possible actions the hired party (the agent) can undertake. For instance, rather than specify all possible qualities of how my house is painted, the contract could make the handyman's payment contingent on whether the job meets professional standards or not (*i.e.,* whether quality is above or below some threshold). Moreover, because of the vagueness of language, idiosyncrasies of interpretation, and so forth, neither the handyman nor I could be certain how, for any given quality of paint job, whether a judge (or other dispute adjudicator) would decide the job met or failed to meet professional standards. Nevertheless, if the probabilities with which the judge makes her decision vary with the true quality of the paint job in such a way that the judge's ruling would be informative—in a statistical sense spelled out below—about the job the handyman actually did, then the handyman and I can achieve the same outcome as we would have were our contract complete (*i.e.,* directly contingent on every level of quality).

This insight is related to and, indeed, draws upon earlier work by Hermalin and Katz [1991, 1993]. In the latter, my co-author and I suggested that parties to a contract could achieve the first-best solution even when judges made errors in determining what actions had happened. There was no formal analysis in that work, however. In particular, there was no attempt to model the precision of the underlying contract or how the likelihood of judicial error could be affected through either contract design or efforts of the judiciary. In Hermalin and Katz [1991], there is no judge (although the verifiable signal of the agent's action could be interpreted as being a judge's ruling) and no attempt to study the effect of the information structure nor the legal system on the ability of the contracting parties to achieve efficient outcomes using incomplete contracts.

That efficient solutions can be achieved with incomplete contracts suggests that efforts to improve contracts by making them more precise or improving the informativeness of events about the parties' actions can be without value: If, for instance, I achieve the complete-contract solution with a contract that sets the handyman's pay at one level if the job meets professional standards and another if it doesn't, then what can I gain by adding more tiers to the contract's assessments of quality (*e.g.,* have quality judged as dreadful, mediocre, or excellent)? In turn, this point raises questions

to the incentives in the contract, not how he interprets it.

about the wisdom of the courts refusing to enforce vague contracts (*e.g.,* on grounds of mistake; see Hermalin et al., in press, especially §2.5.2, for a survey).

In fact, improvements in judicial accuracy above some minimum level can be *detrimental* to economic efficiency. This is a view at odds with much of the literature, which argues that greater judicial accuracy is beneficial *ceteris paribus*.[3] However, as this article shows (see Section 3.4), the cost of a more accurate judiciary could be to require parties to write more detailed contracts.[4] In any case, one point that is worth emphasizing is there is a distinction between judicial accuracy and the informativeness, in a statistical sense, of a judge's rulings.

The main body of the article analyzes the problem abstracting away from some real features of the law. In particular, the costs of going to court, actual trial procedures, and limitations on the penalties that the contractual parties can impose on each other are ignored. Section 4 briefly discusses how the article's results extend if these features are taken into account. Although the analysis there only scratches the surface with respect to these issues, it does suggest that these features may temper some of the conclusions in the main part of the article but not reverse any of them.

2 Model

Consider a situation in which one party, the principal, wishes to employ a second party, the agent, to undertake a task that will benefit the principal. The timing and some other assumptions of the model are as follows:

1. The principal and agent agree to a contract. The precise bargaining game that occurs at this stage is not important to the results, but to keep the analysis straightforward assume that the principal makes a take-it-or-leave-it offer to the agent. There is some level of utility that the agent would enjoy if he chose *not* to work for the principal—his *reservation utility*—and the agent, thus, accepts the principal's offer if and only if his expected utility on the equilibrium path under the

[3] For example, consider Posner [1973, p. 401]: "Judicial error is therefore a source of social costs and the reduction of error is a goal of the procedural system."

[4] There could be other costs of greater accuracy arising from an increased burden on the judicial system itself (*e.g.,* longer trials, more cumbersome procedures, etc.). These are not modeled here.

offered contract is at least as great as this reservation utility. Without loss of generality, normalize the value of the reservation utility to zero.

2. Assuming a contract was agreed to in the first stage, the agent chooses an action $a \in \mathcal{A}$. The set $\mathcal{A}$ has K elements. In contrast to a hidden-action model (*e.g.,* Holmstrom, 1979, Grossman and Hart, 1983), assume that the principal observes the agent's choice of action.

3. The principal offers a payment, w, to the agent. If the agent accepts, the game is over. If the agent rejects, then the agent sues the principal for proper payment under the terms of the contract.

4. If the agent sues, a third-party adjudicator—hereafter known as the *judge* although it could be a jury, arbitrator, or other party empowered to render a verdict—rules on the case and specifies the payment the principal is to make to the agent.

The principal's payoff is $B(a) - t$, where $B : \mathcal{A} \to \mathbb{R}$ maps the agent's action into a monetary benefit (possibly an expected value) and t is the total transfer from principal to agent. The agent's payoff is $V(t) - C(a)$, where $V : (\underline{t}, \infty) \to \mathbb{R}$ and $C : \mathcal{A} \to \mathbb{R}$. Assume that $V(\cdot)$ is continuous, strictly increasing, and either affine or concave (*i.e.,* the agent is either risk neutral or risk averse). Assume further that $\lim_{t \to \underline{t}} V(t) = -\infty$ and $\lim_{t\to\infty} V(t) = \infty$. Observe that a consequence of these assumptions is that the inverse function $V^{-1}(\cdot)$ exists and is defined for all $v \in \mathbb{R}$.

Holding the agent to his reservation utility level (*i.e.,* 0), the social monetary cost of action a is $\Psi(a) \stackrel{\text{def}}{=} V^{-1}\bigl(C(a)\bigr)$. The first-best action is, therefore, the solution to

$$\max_{a \in \mathcal{A}} B(a) - \Psi(a)\,. \tag{1}$$

To insure there is a misalignment of preferences between principal and agent, assume

Assumption 1 *Let a^{FB} be a solution to (1). Then there exists an $a \in \mathcal{A}$ such that $C(a) < C(a^{\text{FB}})$; that is, a least-cost action for the agent is not a first-best action.*

Assume that the principal's benefit, $B(a)$, is unverifiable. Assume further that it is either (i) inalienable (*e.g.,* it is the benefit the principal enjoys from her painted house); or (ii) it is the expected value of a stochastic payoff and the agent is risk averse (*i.e.,* $V(\cdot)$ is strictly concave); or both. These assumptions rule out basing contracts on the realization of the principal's benefit and, moreover, ensure that the first-best solution cannot be achieved by "selling the store" to the agent (*i.e.,* transferring the principal's benefit to the agent in exchange for an upfront payment from the agent to the principal).

Let $P(\mathcal{A})$ denote a partition of $\mathcal{A}$.[5] Denote an element of $P(\mathcal{A})$ by A. Index the elements of $P(\mathcal{A})$ by n. Define $I : \mathcal{A} \to \mathbb{N}$ such that $I(a) = n$ if and only if $a \in A_n$. Observe $a \in A_{I(a)}$ by construction.

A *contract* is a mapping $w : P(\mathcal{A}) \to \mathbb{R}$ of payments from the principal to the agent. For example, $\mathcal{A}$ could be the set of possible quality levels that the agent could choose for a job, $P(\mathcal{A})$ could be the dichotomous division of those quality levels into "satisfactory" and "unsatisfactory," and the contract could stipulate one payment level for a satisfactory job and another payment level for an unsatisfactory job. Using the language of probability theory, refer to the elements of $P(\mathcal{A})$ as *events*.

The judge does not observe the agent's action, a, directly. Rather, she observes evidence that leads her to "conclude" that $A_n \in P(\mathcal{A})$ is the event that occurred. Let $\pi_n(a)$ denote the probability that she concludes the event was A_n when the agent took action a and let $\boldsymbol{\pi}(a) \stackrel{\text{def}}{=} \big(\pi_1(a), \dots, \pi_N(a)\big)$ denote the probability density vector over events (conclusions) conditional on the agent's having taken action a.

It is worth commenting on what is meant by the judge's "concluding" that event A_n occurred. If the judge were "sophisticated," in the sense of understanding the structure of the game and equilibrium play, then the judge should base her decisions on that knowledge. For instance, if the agent plays a pure strategy in equilibrium, $\hat{a}$, then a sophisticated judge might well determine that $A_{I(\hat{a})}$ occurred regardless of what the evidence might indicate.[6] Alternatively, if the agent plays a mixed strategy, the judge should employ Bayes's Rule to update from the evidence to reach a determination of what occurred.

[5]A partition of a set $\mathcal{Z}$ is a collection of non-overlapping subsets of $\mathcal{Z}$, $\{Z_1, \dots, Z_J\}$, that cover $\mathcal{Z}$; that is, such that $Z_j \cap Z_k = \emptyset$ for all j, k, $j \neq k$ and $\bigcup_{j=1}^{J} Z_k = \mathcal{Z}$.

[6]Although, given that the parties should never end up in court on the equilibrium path, a sophisticated judge would recognize that her involvement has followed some out-of-equilibrium play, in which case it is not clear that the most sensible belief for her to hold is that $\hat{a}$ was chosen.

Here, however, I will assume that the judge does not act in such a sophisticated manner; rather her conclusion is based solely on the evidence. There are two justifications for this assumption. First, because of the myriad cases she adjudicates, a judge may not understand the precise structure of the game or she otherwise lacks the knowledge necessary to determine what equilibrium play is. That is, a "naïve" judge is a better model of the judiciary than a sophisticated judge. Second, legal systems possess rules of evidence that restrict judges to making decisions based on the evidence.[7] Indeed, the Anglo-American legal system could be described as "anti-Bayesian" insofar as great effort is made to insure that judges hold as neutral a prior as possible (*e.g.,* attempts to exclude evidence pertaining to prior behavior or other "prejudicial" evidence). In particular, any arguments about how another party should have played the game would be ruled out as "speculative."

Indeed, it is arguably superior to have such a naïve judge rather than a sophisticated judge, because if a judge were sophisticated enough to determine what action, $\hat{a}$, is chosen in equilibrium and always rules, accordingly, that event $A_{I(\hat{a})}$ occurred, then it would become impossible to support pure-strategy equilibria (other than one in which the agent chooses a least-cost action): The judge would be like the *uninformed* principal in Fudenberg and Tirole [1990] and, consequently, the judge would serve to undermine incentives.[8]

Based on her conclusion, the judge issues a ruling. Specifically, if the judge concludes A_n, then she orders the principal to pay the agent $w(A_n)$ as stipulated by the contract.

[7]For instance, in California, juries are admonished not to draw inferences from the fact that an arrest was made or that the case is being brought to trial (author's personal experience as a member of two juries). Although not stated in such terms, these admonitions could be understood as instructions to ignore what can be inferred from the structure of the game (*e.g.,* incentives of the police to make an arrest, a plaintiff's incentives to bring suit, the implications of the apparent failure of the prosecution to obtain a plea bargain, or the implications of the apparent failure of the litigants to settle).

[8]Although one needs in modeling such a sophisticated judge to account for the fact that any court case is an out-of-equilibrium event (see footnote 6), which could cause even a sophisticated judge to doubt $\hat{a}$ was chosen.

3 Analysis of the Basic Model

3.1 The Principal's Payment Offer

Consider a path of the game tree in which the agent has chosen action $\tilde{a}$.

Suppose the agent rejects the principal's offer in stage 3, the agent sues, and the parties end up in court. At the time they proceed to court, both know which action the agent chose, $\tilde{a}$. The expected payment to the agent and his expected utility are, respectively,

$$\mathbb{E}w = \sum_{n=1}^{N} \pi_n(\tilde{a}) w(A_n) = \boldsymbol{\pi}(\tilde{a}) \cdot \mathbf{w} \text{ and } \mathbb{E}V = \sum_{n=1}^{N} \pi_n(\tilde{a}) V\big(w(A_n)\big) = \boldsymbol{\pi}(\tilde{a}) \cdot \mathbf{v}\,,$$

where $\mathbf{w} \stackrel{\text{def}}{=} \big(w(A_1), \dots, w(A_N)\big)$ and $\mathbf{v} \stackrel{\text{def}}{=} \Big(V\big(w(A_1)\big), \dots, V\big(w(A_N)\big)\Big)$.

It follows that is rational for the agent to accept the principal's stage-3 offer, w, if and only if

$$\boldsymbol{\pi}(\tilde{a}) \cdot \mathbf{v} \leq V(w)\,. \tag{2}$$

Clearly, if the principal wishes to make an offer at stage 3 that the agent will accept, she does best to offer the smallest value of w satisfying (2); that is,

$$\underline{w}(\tilde{a}) \stackrel{\text{def}}{=} V^{-1}\big(\boldsymbol{\pi}(\tilde{a}) \cdot \mathbf{v}\big)\,.$$

Given that the agent is risk neutral or risk averse, it follows that $V^{-1}(\cdot)$ is convex (at least weakly) and, thus, from Jensen's inequality that

$$\underline{w}(\tilde{a}) \leq \sum_{n=1}^{N} \pi_n(\tilde{a}) V^{-1}\Big(V\big(w(A_n)\big)\Big) = \sum_{n=1}^{N} \pi_n(\tilde{a}) w(A_n) = \mathbb{E}w\,.$$

This last expression states that the minimum payment the agent will accept rather than go to court is no greater than the principal's expected payment should they go to court. That is, the principal prefers to offer $\underline{w}(\tilde{a})$ rather than go to court.[9] This can be summarized as:

LEMMA 1 *Under the assumptions of the basic model, whatever action $a \in \mathcal{A}$ the agent takes, it is an equilibrium for the parties not go to court. Rather,*

[9] As derived, this is a weak preference. It is a strict preference if the agent is risk averse and the contract is not a constant-wage contract.

the principal will offer the agent a payment, $\underline{w}(a)$, just large enough to provide the agent the same utility as he would expect to receive from going to court under the prevailing contract and the agent will accept this offer.

Intuitively, going to court exposes the agent to risk. This can never improve efficiency (and will, in fact, reduce it if the agent is risk averse) insofar as there is a risk-neutral party, the principal, who can provide full insurance. Because actions are sunk at this stage of the game, there is no incentive problem to motivate exposing the agent to risk. Hence, the principal "sells" the agent full insurance at a price just equal to agent's willingness to pay for such insurance.

3.2 Initial Contract Design

Consider the situation in which the principal wishes to employ the agent to undertake task $\tilde{a}$. The initial contract must—recognizing the principal's subsequent offer as given by Lemma 1—provide the agent the appropriate incentives to undertake task $\tilde{a}$:

$$V\big(\underline{w}(\tilde{a})\big) - C(\tilde{a}) \geq V\big(\underline{w}(a)\big) - C(a)\ \forall a \in \mathcal{A}\,. \tag{IC}$$

Moreover, recognizing that his best response conditional on accepting the contract is $\tilde{a}$, the agent must prefer accepting the contract to rejecting it:

$$V\big(\underline{w}(\tilde{a})\big) - C(\tilde{a}) \geq 0\,. \tag{IR}$$

Because the principal offers the smallest w consistent with (2), $V\big(\underline{w}(a)\big) = \boldsymbol{\pi}(a)\cdot\mathbf{v}$. Hence, (IC) and (IR) can be rewritten as

$$\boldsymbol{\pi}(\tilde{a})\cdot\mathbf{v} - C(\tilde{a}) \geq \boldsymbol{\pi}(a)\cdot\mathbf{v} - C(a)\ \forall a \in \mathcal{A}\,;\ \text{and} \tag{IC$'$}$$

$$\boldsymbol{\pi}(\tilde{a})\cdot\mathbf{v} - C(\tilde{a}) \geq 0\,. \tag{IR$'$}$$

It follows that whether or not the principal can employ the agent to undertake task $\tilde{a}$—that is, whether action $\tilde{a}$ is *implementable*—depends on whether or not there exists an N-dimensional vector $\mathbf{v}$ that solves (IC$'$) and (IR$'$).

LEMMA 2 *Under the assumptions of the basic model, there exists a contract (an N-dimensional vector $\mathbf{v}$ that solves (IC$'$) and (IR$'$)) that both induces the agent to work for the principal and undertake task $\tilde{a}$ if and only if there is* no *strategy for the agent, $\sigma : \mathcal{A}\backslash\{\tilde{a}\} \to \Delta_{K-1}$, that induces the same* *density over events as $\tilde{a}$ (i.e., such that $\boldsymbol{\pi}(\tilde{a}) = \sum_{a\in\mathcal{A}\setminus\{\tilde{a}\}} \sigma(a)\boldsymbol{\pi}(a)$) and which costs less, in terms of expected disutility, than $\tilde{a}$ (i.e., such that $C(\tilde{a}) > \sum_{a\in\mathcal{A}\setminus\{\tilde{a}\}} \sigma(a)C(a)$).*[10]

PROOF This is essentially a restatement of Proposition 2 of HERMALIN AND KATZ [1991] and follows immediately from it. *Q.E.D.*

The necessity of Lemma 2 is obvious. Sufficiency follows because if the density over events induced by $\tilde{a}$ is distinct in the sense given above, then events must be informative as to whether $\tilde{a}$ or some other action were taken. Because, in the basic model, the range of $V(\cdot)$ is unbounded, even a small amount of information can be used to implement $\tilde{a}$ by appropriately rewarding the agent for certain events and punishing him for others. This concept of distinct can be restated as

COROLLARY 1 *Under the assumptions of the basic model, there exists a contract that both induces the agent to work for the principal and undertake task $\tilde{a}$ if $\boldsymbol{\pi}(\tilde{a})$ is* not *an element of the convex hull of $\{\boldsymbol{\pi}(a)|a \neq \tilde{a}\}$.*[11]

Because the principal's cost of implementing $\tilde{a}$ is $\underline{w}(\tilde{a}) = V^{-1}(\boldsymbol{\pi}(\tilde{a})\cdot\mathbf{v})$, it follows that the principal prefers, out of all the contracts that implement $\tilde{a}$, the one that minimizes $\boldsymbol{\pi}(\tilde{a})\cdot\mathbf{v}$ subject to (IC′) and (IR′). As is well known, if there is a vector, $\mathbf{v} = (v_1, \ldots, v_N)$, satisfying (IC′) and (IR′), such that (IR′) is a strict inequality, then there exists an $\varepsilon > 0$ such that the vector $(v_1 - \varepsilon, \ldots, v_N - \varepsilon)$ also satisfies (IC′) and (IR′). Hence, we may conclude that the principal will offer a contract such that (IR′) is an equality.

LEMMA 3 *Under the assumptions of the basic model, if $\tilde{a}$ is an implementable action that the principal wishes to implement in equilibrium, then the cost to her of implementing $\tilde{a}$ is $\Psi(\tilde{a})$, the first-best (full-information) cost.*

[10] Δ_{K-1} denotes the $K-1$-dimensional unit simplex; that is, in this case, the set of all possible mixed and pure strategies for the agent over actions other than $\tilde{a}$. (Recall $\mathcal{A}$ has K elements.)

[11] The convex hull of a set $\mathcal{X}$ with J elements and representative element $\mathbf{x}_j$ is the set of all convex combinations of the elements of $\mathcal{X}$; that is

$$\left\{ \sum_{j=1}^{J} \lambda_j \mathbf{x}_j \middle| (\lambda_1, \ldots, \lambda_J) \in \Delta_J \right\}.$$

PROOF Because (IR′) and, therefore, (IR) binds:

$$V\big(\underline{w}(\tilde{a})\big) - C(\tilde{a}) = 0\,.$$

Hence, $\underline{w}(\tilde{a}) = V^{-1}\big(C(\tilde{a})\big) = \Psi(\tilde{a})$. *Q.E.D.*

3.3 Partitions and Information

Heretofore, the partition of the action space, $\mathcal{A}$, into events and the probabilities of the judge's conclusions, the densities $\{\boldsymbol{\pi}(a)\}_{a\in\mathcal{A}}$, have been treated as essentially exogenous. But, of course, the partitions are part of contract design. Moreover, perhaps through better description of the events, it is possible that the densities are also part of the contract design. In this subsection, I consider the choice of a partition, or the choice of densities, or both.

From Lemma 3, if an action, $\tilde{a}$, is implementable, it is implementable at first-best cost, $\Psi(\tilde{a})$. Because $\Psi(\tilde{a})$ is determined solely by the agent's utility function (*i.e.,* $V(\cdot)$ and $C(\cdot)$), it follows that the choice of partition or densities affects the cost of implementing an action only insofar as it reduces it from infinite (*i.e.,* from being non-implementable) to the first-best cost (*i.e.,* to being implementable). Consequently, preferences over partitions or densities are, in this base model, guided solely by whether the partitions or densities permit a desired action to be implemented or not.

Partitions lack a complete order. One can place a partial order over partitions using the notion of *finer* (alternatively, *coarser*). Specifically, partition $P(\mathcal{A})$, with representative element A, is finer than partition $\hat{P}(\mathcal{A})$, with representative element $\hat{A}$, if for every $A \in P(\mathcal{A})$ there exists an $\hat{A} \in \hat{P}(\mathcal{A})$ such that $A \subseteq \hat{A}$ and there is at least one $A \in P(\mathcal{A})$ and $\hat{A} \in \hat{P}(\mathcal{A})$ such that $A \subset \hat{A}$ (*i.e.,* A is a proper subset of $\hat{A}$, so $\hat{A}$ contains elements not in A). Stating that $P(\mathcal{A})$ is finer than $\hat{P}(\mathcal{A})$ is equivalent to stating that $\hat{P}(\mathcal{A})$ is coarser than $P(\mathcal{A})$.

Although it is possible that writing a contract that partitions $\mathcal{A}$ more finely also affects the probabilities with which the judge reaches her conclusions, it is worth considering the case in which those probabilities are fixed in the following sense: If $P(\mathcal{A})$ is a finer partition than $\hat{P}(\mathcal{A})$ and $\boldsymbol{\pi}(\cdot) : \mathcal{A} \to \Delta_N$ and $\hat{\boldsymbol{\pi}}(\cdot) : \mathcal{A} \to \Delta_{\hat{N}}$ are the densities associated with $P(\mathcal{A})$ and $\hat{P}(\mathcal{A})$, respectively, then

- $A_n \subseteq \hat{A}_m$ implies $\pi_n(a) \leq \hat{\pi}_m(a)$ for all $a \in \mathcal{A}$; and
- $\hat{A}_m = A_{n_1} \cup \cdots \cup A_{n_J}$ implies $\hat{\pi}_m(a) = \sum_{j=1}^{J} \pi_{n_j}(a)$ for all $a \in \mathcal{A}$.

PROPOSITION 1 *Assume fixed probabilities. If action $\tilde{a}$ is implementable given partition $\hat{P}(\mathcal{A})$, then it is implementable under any finer partition.*

PROOF Let $\hat{w} : \hat{P}(\mathcal{A}) \to \mathbb{R}$ be a contract that implements $\tilde{a}$ given partition $\hat{P}(\mathcal{A})$. Let $\hat{\mathbf{v}} \stackrel{\text{def}}{=} \Big(V\big(\hat{w}(\hat{A}_1)\big), \ldots, V\big(\hat{w}(\hat{A}_{\hat{N}})\big)\Big)$. Let $P(\mathcal{A})$ be a finer partition. Observe its elements can be indexed by j, n such that $A_{j,n} \subseteq \hat{A}_n$ and $\bigcup_{j=1}^{J(n)} A_{j,n} = \hat{A}_n$. Construct a contract such that $w(A_{j,n}) = \hat{w}(\hat{A}_n)$ for all $n \in \{1, \ldots, \hat{N}\}$ and all $j \in \{1, \ldots, J(n)\}$. Let $v_{j,n} = V\big(w(A_{j,n})\big)$ and let $\mathbf{v}$ be the corresponding vector. This new contract implements $\tilde{a}$ under this finer partition if it satisfies (IC$'$) and (IR$'$). Observe, for any $a \in \mathcal{A}$,

$$\begin{aligned}\boldsymbol{\pi}(a) \cdot \mathbf{v} &= \sum_{n=1}^{\hat{N}} \sum_{j=1}^{J(n)} \pi_{j,n}(a) V\big(w(A_{j,n})\big) = \sum_{n=1}^{\hat{N}} \sum_{j=1}^{J(n)} \pi_{j,n}(a) V\big(\hat{w}(\hat{A}_n)\big) \\ &= \sum_{n=1}^{\hat{N}} V\big(\hat{w}(\hat{A}_n)\big) \sum_{j=1}^{J(n)} \pi_{j,n}(a) = \sum_{n=1}^{\hat{N}} V\big(\hat{w}(\hat{A}_n)\big) \hat{\pi}_n(a) = \hat{\boldsymbol{\pi}}(a) \cdot \hat{\mathbf{v}}\,.\end{aligned}$$

Given that $\hat{w}(\cdot)$ satisfies (IC$'$) and (IR$'$) when the partition is $\hat{P}(\mathcal{A})$, it follows from this last expression that the $w(\cdot)$ contract satisfies those conditions when the partition is $P(\mathcal{A})$. *Q.E.D.*

Intuitively, a finer partition simply provides greater contractual flexibility insofar as there are more events upon which compensation can be contingent. One does not, however, have to make use of that additional flexibility if a coarser partition is good enough; that is, one effectively offers the same action-implementing contract with a finer partition as one did with a coarser partition.

Another way to summarize this result is

COROLLARY 2 *If there exists a contract that implements that principal's desired action based on a given partition of the action space, then there is no* *gain for the principal (and thus to welfare) of using a contract that requires a finer partition of the action space.*[12]

As noted previously, there is the possibility that the principal—perhaps through her efforts to describe the contractual contingencies (events) more clearly—could change the densities. In particular, she could improve informativeness. In the context of contract theory, the standard definition of informativeness is Blackwell informativeness (see, *e.g.,* GROSSMAN AND HART, 1983).[13] Fix a partition and consider two different sets of densities $\{\boldsymbol{\pi}^1(a)\}_{a\in\mathcal{A}}$ and $\{\boldsymbol{\pi}^2(a)\}_{a\in\mathcal{A}}$ over that partition. If, for all $a \in \mathcal{A}$, $\boldsymbol{\pi}^2(a) = \mathbf{Q}\boldsymbol{\pi}^1(a)$, where $\mathbf{Q}$ is a constant stochastic transformation matrix (*i.e.,* a matrix with non-negative elements in which each column sums to one and at least one column has two positive elements), then events are more informative under the first set of densities (*i.e.,* those with superscript 1) than under the second set of densities.

PROPOSITION 2 *Consider a fixed partition and two different sets of densities over that partition,* $\{\boldsymbol{\pi}^1(a)\}_{a\in\mathcal{A}}$ *and* $\{\boldsymbol{\pi}^2(a)\}_{a\in\mathcal{A}}$*. Suppose events are more informative in the Blackwell sense given the first set of densities than given the second and that action* $\tilde{a}$ *is implementable under the second. Then* $\tilde{a}$ *is implementable under the first set of densities as well.*

PROOF Let $\mathbf{v}^2$ be a contract (in utility terms) that implements $\tilde{a}$ (*i.e.,* that satisfies conditions (IC′) and (IR′)) when the relevant densities are $\{\boldsymbol{\pi}^2(a)\}_{a\in\mathcal{A}}$. By assumption, there exists a stochastic transformation matrix $\mathbf{Q}$, with element $q_{m,n}$ in the mth row of the nth column, such that, for all $a \in \mathcal{A}$, $\boldsymbol{\pi}^2(a) = \mathbf{Q}\boldsymbol{\pi}^1(a)$. Let $\mathbf{q}_n$ be the nth column of $\mathbf{Q}$. Define the contract $\mathbf{v}^1$ such that $v_n^1 \stackrel{\text{def}}{=} \mathbf{q}_n \cdot \mathbf{v}^2$. The contract $\mathbf{v}^1$ implements $\tilde{a}$ if it satisfies (IC′) and (IR′) when the relevant set of densities is $\{\boldsymbol{\pi}^1(a)\}_{a\in\mathcal{A}}$. To see that it does,

[12]The "thus to welfare" follows because the agent's expected utility is always held to a constant, zero, so the principal's gains are equivalent to gains to total welfare.

[13]See KIM [1995] and HERMALIN AND KATZ [2000] for alternative notions of greater informativeness applicable to agency problems.

observe

$$\begin{aligned}\boldsymbol{\pi}^1(a)\cdot\mathbf{v}^1 = \sum_{n=1}^{N}\pi_n^1(a)v_n^1 &= \sum_{n=1}^{N}\pi_n^1(a)\sum_{m=1}^{N}q_{m,n}v_m^2 \\ &= \sum_{m=1}^{N}v_m^2\sum_{n=1}^{N}\pi_n^1(a)q_{m,n} = \sum_{m=1}^{N}v_m^2\pi_m^2(a) = \boldsymbol{\pi}^2(a)\cdot\mathbf{v}^2\,.\end{aligned}$$

Hence, because $\mathbf{v}^2$ satisfies (IC′) and (IR′) under $\{\boldsymbol{\pi}^2(a)\}_{a\in\mathcal{A}}$, so too must $\mathbf{v}^1$ under $\{\boldsymbol{\pi}^1(a)\}_{a\in\mathcal{A}}$. *Q.E.D.*

Another way to summarize this result is

COROLLARY 3 *Consider a fixed partition of the action space. If there exists a contract that implements that principal's desired action under a given set of densities, then there is no gain for the principal (and thus to welfare) of using a contract that utilizes a more informative (in the Blackwell sense) set of densities.*

3.4 Judicial Accuracy

One informativeness criterion that the courts could care about is *accuracy*: How likely is the judge to reach the "right" conclusion. If right is interpreted to mean concluding $A_{I(a)}$ when the agent's action was a, then the judge's probability of being right is $\pi_{I(a)}(a)$. Improvements in accuracy would, thus, correspond to increasing $\pi_{I(a)}(a)$ for all a. One way to capture such an increase is the following:

DEFINITION 1 *Fix a partition of the action space. Let*

$$\mathbf{e}_n \overset{def}{=} (0,\ldots,0,\underbrace{1}_{n\text{th element}},0,\ldots,0)\,.$$

The densities (information structure) $\{\boldsymbol{\pi}^1(a)\}_{a\in\mathcal{A}}$ *are (is)* more accurate *than* $\{\boldsymbol{\pi}^2(a)\}_{a\in\mathcal{A}}$ *if there exists a* $\lambda\in(0,1]$ *such that, for all* $a\in\mathcal{A}$,

$$\boldsymbol{\pi}^1(a) = \lambda\mathbf{e}_{I(a)} + (1-\lambda)\boldsymbol{\pi}^2(a)\,.$$

Observe, conditional on a given action, a, that a more accurate density has weight shifted to the event corresponding to that action (*i.e.,* to $A_{I(a)}$) and, correspondingly, weight shifted away from other events.

The notion of accuracy is distinct from that of Blackwell informativeness. For instance, suppose $\mathcal{A} = \{1, 2\}$ and $P(\mathcal{A}) = \big\{\{1\}, \{2\}\big\}$. Suppose $\pi^2_{I(a)}(a) = 2/5$ and $\pi^1_{I(a)}(a) = 4/7$. Then $\{\boldsymbol{\pi}^1(a)\}$ is a more accurate information structure than $\{\boldsymbol{\pi}^2(a)\}$, since

$$\boldsymbol{\pi}^1(a) = \frac{2}{7}\mathbf{e}_{I(a)} + \frac{5}{7}\boldsymbol{\pi}^2(a)\,.$$

The matrix, $\mathbf{Q}$, that maps the former into the latter is not, however, a stochastic transformation matrix because it has negative elements:

$$\mathbf{Q} = \begin{pmatrix} -\frac{1}{5} & \frac{6}{5} \\ \frac{6}{5} & -\frac{1}{5} \end{pmatrix}.$$

Hence, the more accurate information structure is not more informative in the Blackwell sense. In fact, in this example, the opposite is true—the more accurate structure is *less* informative: $\boldsymbol{\pi}^1(a) = \hat{\mathbf{Q}}\boldsymbol{\pi}^2(a)$, where

$$\hat{\mathbf{Q}} = \begin{pmatrix} \frac{1}{7} & \frac{6}{7} \\ \frac{6}{7} & \frac{1}{7} \end{pmatrix}$$

is a stochastic transformation matrix.

It is not surprising, therefore, that increasing accuracy can *harm* the principal insofar as it can render some actions non-implementable. For example, suppose that $\mathcal{A} = \{a_1, \ldots, a_4\}$ and $P(\mathcal{A}) = \big\{\{a_1\}, \{a_2\}, \{a_3, a_4\}\big\}$. Define $A_n = \{a_n\}$ if $n \leq 2$ and $A_3 = \{a_3, a_4\}$. Let the less accurate distributions be

$$\boldsymbol{\pi}^2(a_1) = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{4} \\ \frac{1}{4} \end{pmatrix}, \ \boldsymbol{\pi}^2(a_2) = \begin{pmatrix} \frac{3}{8} \\ \frac{3}{8} \\ \frac{1}{4} \end{pmatrix}, \ \boldsymbol{\pi}^2(a_3) = \begin{pmatrix} \frac{1}{4} \\ \frac{1}{4} \\ \frac{1}{2} \end{pmatrix}, \ \& \ \boldsymbol{\pi}^2(a_4) = \begin{pmatrix} \frac{3}{40} \\ \frac{1}{40} \\ \frac{9}{10} \end{pmatrix}.$$

As Figure 1 illustrates, action a_3 is implementable because it lies outside the convex hull formed by the other densities (Corollary 1).

Consider the more accurate information structure defined by

$$\boldsymbol{\pi}^1(a) = \frac{1}{2}\mathbf{e}_{I(a)} + \frac{1}{2}\boldsymbol{\pi}^2(a)\,.$$

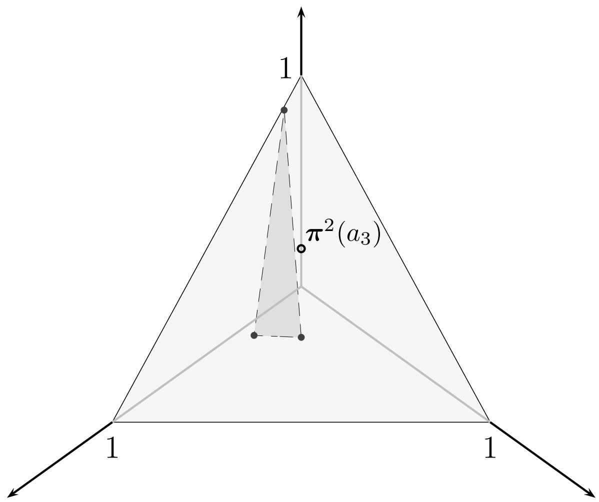

Figure 1: Action a_3 is implementable because $\boldsymbol{\pi}^2(a_3)$ lies *outside* the convex hull of the other densities. (The light gray triangle is the 3-dimensional unit simplex and the smaller darker triangle is the convex hull of the other densities.)

Hence,

$$\boldsymbol{\pi}^1(a_1) = \begin{pmatrix} \frac{3}{4} \\ \frac{1}{8} \\ \frac{1}{8} \end{pmatrix}, \ \boldsymbol{\pi}^1(a_2) = \begin{pmatrix} \frac{3}{16} \\ \frac{11}{16} \\ \frac{1}{8} \end{pmatrix}, \ \boldsymbol{\pi}^1(a_3) = \begin{pmatrix} \frac{1}{8} \\ \frac{1}{8} \\ \frac{3}{4} \end{pmatrix}, \ \& \ \boldsymbol{\pi}^1(a_4) = \begin{pmatrix} \frac{3}{80} \\ \frac{1}{80} \\ \frac{19}{20} \end{pmatrix}.$$

Under this information structure $\boldsymbol{\pi}^1(a_3)$ lies *in* the convex hull formed by the other densities. Specifically,

$$\boldsymbol{\pi}^1(a_3) = \frac{3}{33}\boldsymbol{\pi}^1(a_1) + \frac{5}{33}\boldsymbol{\pi}^1(a_2) + \frac{25}{33}\boldsymbol{\pi}^1(a_4)\,.$$

If

$$C(a_3) > \frac{3}{33}C(a_1) + \frac{5}{33}C(a_2) + \frac{25}{33}C(a_4)\,,$$

then action a_3 is not implementable under this more accurate information structure (Lemma 2). If a_3 were the welfare-maximizing action (*i.e.,* the solution to $\max_{a\in\mathcal{A}} B(a) - \Psi(a)$), then the consequence of greater accuracy is to reduce welfare.[14] To summarize:

Observation 1 *Greater judicial accuracy can be worse for welfare than less accuracy.*

In fact, if the judiciary is too accurate, then for a partition with N elements it could be that only N actions are implementable, which could represent a reduction in the number of implementable actions *vis-à-vis* less accuracy (*e.g.,* as in the above example, where, given the less accurate information structure all four actions are implementable by Corollary 1, but only three are implementable given the more accurate information structure). To understand this point better, consider the following definition and its implication for the implementability of actions.

Definition 2 *Fix a partition of the action space. An information structure is* accurate *if, for all $a \in \mathcal{A}$, $\pi_{I(a)}(a) > \pi_{I(a)}(a')$ for all $a' \notin A_{I(a)}$. That is, the information structure is accurate if, for any action, the probability that the judge concludes the event (element of $P(\mathcal{A})$) to which the action belongs occurred is greater when the agent takes that action than when he chooses an action not in that event.*

Proposition 3 *Fix a partition of the action space. If the information structure is accurate, then each event $A_n \in P(\mathcal{A})$ has at least one element, $\hat{a}_n$, that is implementable.*

Proof Let $A_n^* \stackrel{\text{def}}{=} \{a \in A_n | \pi_n(a) \geq \pi_n(a') \; \forall a' \in A_n\}$. Because $\mathcal{A}$ and, hence, A_n are finite, A_n^* is non-empty. By the definition of accurate, if $a \in A_n$ and $a' \notin A_n$, then $\pi_n(a) > \pi_n(a')$. Hence, if $a \in A_n^*$, then $\pi_n(a) \geq \pi_n(a') \; \forall a' \in \mathcal{A}$. Let $\hat{a}_n \stackrel{\text{def}}{=} \min_{a\in A_n^*} C(a)$. At least one $\hat{a}_n$ exists because A_n^* is finite. The claim is $\hat{a}_n$ is implementable. Observe $\boldsymbol{\pi}(\hat{a}_n)$ can lie in the convex hull of the densities of *other* actions only if those actions are *all* in A_n^*. But for any convex combination σ over A_n^*:

$$\sum_{a\in A_n^*\setminus\{\hat{a}_n\}} \sigma(a)C(a) \geq \sum_{a\in A_n^*\setminus\{\hat{a}_n\}} \sigma(a)C(\hat{a}_n) = C(\hat{a}_n)\,.$$

[14]As will become clear later, the principal might be able to respond to greater judicial accuracy by utilizing a finer partition (*i.e.,* separating a_3 and a_4). However, to the extent that a finer partition represents greater contract writing costs, it will still be the case that welfare has been reduced—the principal may succeed in implementing a_3, but at the expense of having to write a more detailed contract.

Hence, by Lemma 2, $\hat{a}_n$ is implementable. *Q.E.D.*

Another way to state Proposition 3 is that, given sufficient accuracy, the set of implementable actions has at least as many elements as the set of events. The two sets can have exactly the same cardinality if there is *complete* accuracy.

DEFINITION 3 *Fix a partition of the action space. An information structure is* completely accurate *if, for all $a \in \mathcal{A}$, $\pi_{I(a)}(a) = 1$. That is, the information structure is completely accurate if, for any action, the judge always concludes that the event that occurred is the one to which the action belongs.*

PROPOSITION 4 *Fix a partition of the action space. If the information structure is completely accurate and if for any two actions, a and a', $C(a) \neq C(a')$ (i.e., actions are distinct with respect to the agent's disutility of taking them), then the number of implementable actions equals the number of events in the partition.*

PROOF Suppose not. Then there exists an A_n with at least two elements, a and a', that are implementable. Because the information structure is completely accurate, $\boldsymbol{\pi}(a) = \boldsymbol{\pi}(a')$. By assumption, $C(a) \neq C(a')$. Without loss of generality assume $C(a) < C(a')$. But, then a' cannot satisfy (IC$'$):

$$\boldsymbol{\pi}(a') \cdot \mathbf{v} - C(a') = \boldsymbol{\pi}(a) \cdot \mathbf{v} - C(a') < \boldsymbol{\pi}(a) \cdot \mathbf{v} - C(a)$$

for any contract $\mathbf{v}$. The result follows by contradiction. *Q.E.D.*

As the example connected to Figure 1 demonstrates, the cardinality of the set of implementable actions can strictly exceed the cardinality of $P(\mathcal{A})$ when the information structure is less accurate.

While it is true that greater judicial accuracy can reduce the set of implementable actions, it remains true that the parties might be able to recover if they use a contract with different partitions or one with finer partitions. In particular, Proposition 3 implies any action $\tilde{a}$ can be made implementable if $\mathcal{A}$ is partitioned so that $A_{I(\tilde{a})} = \{\tilde{a}\}$ when there is judicial accuracy (the information structure is accurate in the sense of Definition 2). In fact, this is more extreme than necessary: As the proof of Proposition 3 shows, if $\mathcal{A}$ is partitioned in such a way that there is sufficiency accuracy about $\tilde{a}$ (specifically, if $\pi_{I(\tilde{a})}(\tilde{a}) > \pi_{I(\tilde{a})}(a)$ for all $a \neq \tilde{a}$), then $\tilde{a}$ is implementable.

Corollary 4 *Fix a partition of the action space. If* $\pi_{I(\tilde{a})}(\tilde{a}) > \pi_{I(\tilde{a})}(a)$ *for all* $a \neq \tilde{a}$*, then* $\tilde{a}$ *is implementable.*

Proof Observe that $A^*_{I(\tilde{a})} = \{\tilde{a}\}$. The proof then follows that of Proposition 3. *Q.E.D.*

Although this flexibility on the part of the contract designers arguably mitigates the case against judicial accuracy, it needs to be remembered that the partitions in a contract could reflect exogenous aspects of the contracting environment. For instance, although one could conceivably put a desired action into its own partition (*e.g.,* partition $\mathcal{A}$ into $\{\tilde{a}\}$ and $\mathcal{A}\backslash\{\tilde{a}\}$) this might be more costly to describe given the nature of language than other divisions of the action space. For example, recall the house-painting example in the Introduction—stating that the job met or failed to meet professional standards is a simpler description than trying to single out a specific quality level. In addition, there could be terms of art that are well understood. By partitioning the action space using those terms, the parties could economize on contracting costs; but this suggests that partitions are more exogenous than endogenous, which indicates that judicial accuracy could indeed prove costly. A related point is that there could be default or standard contracts or terms that also serve to promote economizing on contracting costs.

In sum, though, we reach a second observation:

Observation 2 *The consequence of greater judicial accuracy (as defined by Definition 1) is that parties could need to incur greater costs in detailing their partitions of the action space or settle for less optimal actions.*

Even if one accepted that the costs incurred by the parties in arriving at the proper partitions is *de minimis*, the preceding analysis clearly shows that there is no benefit to either greater judicial accuracy—indeed any "improvement" in the information structure—once a desired action becomes implementable. To the extent that greater judicial accuracy increases procedural costs, the previous analysis suggests that efficiency could dictate a relatively low level of judicial accuracy.

4 Extensions

4.1 Going to Court is Costly

Heretofore, I have ignored the costs involved in going to court. Although the players never go to court on the equilibrium path, the possibility of going to court helps to determine the nature of equilibrium play. Consequently, court costs could matter insofar as they alter the credibility of going to court and, thus, the nature of equilibrium play.

Let k_A and k_P be the costs borne by the agent and principal, respectively, if they go to court. Assume each side must pay its own costs regardless of the outcome of the trial (*i.e.,* unlike certain types of litigation in certain countries, a judgment never involves shifting costs across the litigants).

Let $\boldsymbol{\nu} \stackrel{\text{def}}{=} \Big(V\big(w(A_1) - k_A\big), \ldots, V\big(w(A_N) - k_A\big)\Big)$ denote the vector of post-judgement utilities for the agent. By the same reasoning as before (*i.e.,* expression (2)), conditional on the agent choosing action $\tilde{a}$, the principal will offer the smallest w such that

$$\boldsymbol{\pi}(\tilde{a}) \cdot \boldsymbol{\nu} \leq V(w)\,. \tag{2'}$$

The rest of the analysis likewise proceeds analogously to before.

Observe that if $w(\cdot)$ were a contract that would implement $\tilde{a}$ at first-best cost in a regime in which $k_A \equiv 0$, then the contract that promises the agent $w(\cdot) + k_A$ will do the same in a regime in which $k_A > 0$ and parties pay their own costs. This can be summarized as

Proposition 5 *Even if court costs are positive, as long as each party pays its own costs, if an action $\tilde{a}$ is implementable (in the sense of Lemma 2), then the principal can implement it at first-best (full-information) cost, $\Psi(\tilde{a})$.*

As noted, in many legal systems there is cost shifting across litigants as part of the judgment (*e.g.,* as is often the case in the US, a winning plaintiff may recover court costs from a losing defendant). Modeling such cost shifting requires having a notion of "winning" at trial, which the current model does not. This, in turn, would require an alternative model of how the court functions, a topic I turn to next.

4.2 An Adversarial Court Process

To this point, if the parties were to go to court, the judge would simply reach a conclusion as to which event, A_n, had occurred and order the principal to pay the agent $w(A_n)$. In many actual court cases, the burden of proof falls to the plaintiff (here, the agent) and the judge rules in his favor if the evidence supports the plaintiff's claim. Specifically, assume:

1. Following the agent's rejection of the principal's stage-3 offer, the agent sues the principal for payment and claims to be entitled to $\hat{w}$.

2. The judge concludes event A_n occurred.

3. If $\hat{w} \leq w(A_n)$, then the judge rules in favor of the agent (the plaintiff) and awards $\hat{w}$ to the agent. If $\hat{w} > w(A_n)$ then the judge rules against the agent and awards $w(A_n)$.

Absent any cost shifting, it is a dominant strategy for the agent to claim $\hat{w} = \max_n w(A_n)$ (*i.e.,* the largest possible payment). It follows that if the parties proceed to court, the agent's payoff (in expected utility) is given by the left-hand side of (2) or (2′) (depending on whether there are court costs). The earlier analysis, therefore, continues to apply:

Proposition 6 *Assume no cost shifting and the trial process is as set forth in this subsection. If an action $\tilde{a}$ is implementable (in the sense of Lemma 2), then the principal can implement it at first-best (full-information) cost, $\Psi(\tilde{a})$.*

The analysis with cost shifting is far more difficult and is left for future work.

4.3 No Penalties

A common restriction in many legal systems is a limitation on contractual penalties; that is, the courts will not enforce a contingency that requires one party to pay an "excessive" penalty to the other. For example, the courts could be unwilling to enforce a contract contingency that "paid" the agent less than w_0 (one can think of $w_0 < 0$, so it is actually a payment from agent to principal). Let $v_0 \stackrel{\text{def}}{=} V(w_0)$.[15]

[15] If w_0 is not in the domain of $V(\cdot)$ (*i.e.,* $w_0 < \underline{t}$), then the restriction can never be binding and can, thus, be ignored.

Suppose the restriction were *not* in force and let $\mathbf{v}$ be a contract (in utility terms) that implements action $\tilde{a}$ at first-best cost. That is,

$$\boldsymbol{\pi}(\tilde{a}) \cdot \mathbf{v} = C(\tilde{a})\,.$$

Let $\underline{v}$ be the smallest element of $\mathbf{v}$. If $\underline{v} < v_0$, then this contract violates the no-penalty restriction.

As a preliminary result, observe that $\tilde{a}$ is still implementable given the non-penalty restriction. To see this, form a new contract, $\hat{\mathbf{v}}$, such that $\hat{v}_n = v_n + (v_0 - \underline{v})$. Because

$$\boldsymbol{\pi}(a) \cdot \hat{\mathbf{v}} = \boldsymbol{\pi}(a) \cdot \mathbf{v} + (v_0 - \underline{v})$$

for all a, it is readily seen that this new contract satisfies (IC′) and (IR′), so $\tilde{a}$ is indeed implementable. To summarize:

PROPOSITION 7 *The set of implementable actions is unaffected by a no-penalty restriction.*

Given that the cost of this newly constructed contract to the principal is $V^{-1}\big(\boldsymbol{\pi}(\tilde{a}) \cdot \mathbf{v} + (v_0 - \underline{v})\big)$, we see that while the restriction has no impact on implementability, it can affect the principal's cost. For this reason, unlike the situation in Corollary 3, improved information in the Blackwell sense can be of value when there is a no-penalty restriction.

PROPOSITION 8 *Consider a fixed partition and two different sets of densities over that partition, $\{\boldsymbol{\pi}^1(a)\}_{a\in\mathcal{A}}$ and $\{\boldsymbol{\pi}^2(a)\}_{a\in\mathcal{A}}$. Suppose events are more informative in the Blackwell sense given the first than given the second and that action $\tilde{a}$, which is not a least-cost action for the agent, is implementable under the second. Suppose the no-penalty restriction binds given the second. Then the principal's cost of implementing $\tilde{a}$ is no greater under the first set of densities than the second; and strictly less if the stochastic transformation matrix from the first to the second, $\mathbf{Q}$, has all positive elements.*

PROOF Consider the proof of Proposition 2 and define terms as done there. Let $\mathbf{v}^2$ be the cost-minimizing contract under the no-penalty restriction given the second information structure. The "no greater" result follows immediately from that proof, since the cost of $\mathbf{v}^1$—the constructed contract that implements $\tilde{a}$ under the first information structure—costs the principal the same amount as $\mathbf{v}^2$ (recall the cost of $\mathbf{v}$ is $V^{-1}\big(\boldsymbol{\pi}(\tilde{a}) \cdot \mathbf{v}\big)$). To see the "strictly less" result, note that at least one element of $\mathbf{v}^2$ is v_0 (as otherwise the no-penalty restriction would not be binding). Because $\tilde{a}$ is not a least-cost action for the agent, not all elements of $\mathbf{v}^2$ can be v_0 (as otherwise the contract would fail (IC′)); that is, some elements of $\mathbf{v}^2$ are greater than v_0. Recall, for each n, that $v_n^1 = \mathbf{q}_n \cdot \mathbf{v}^2$. Because no element of $\mathbf{q}_n$ is 0, it follows that $v_n^1 > v_0$ for all n. In particular, the smallest element of $\mathbf{v}^1$, $\underline{v}^1$, is greater than v_0. Construct a new contract $\hat{v}_n^1 = v_n^1 - (\underline{v}^1 - v_0)$. This contract implements $\tilde{a}$, satisfies the no-penalty restriction, and costs strictly less. *Q.E.D.*

Proposition 8 tempers the conclusion that improvements in the information structure are unnecessary if an action is implementable. As such, it might serve to temper the conclusions about judicial accuracy and the courts' distaste for vague contracts. On the other hand, it is worth noting that restrictions on penalties are themselves imposed by the courts and, at least in this context, cannot be justified as welfare enhancing (especially as they would never be imposed on the equilibrium path). Moreover, even under these restrictions, if the parties choose to write the contracts they do, there is no rationale in this model for the courts to dismiss them as too vague.[16]

5 *Conclusions*

This article has shown that contractual incompleteness due to the difficulties of describing relevant contingencies precisely is not necessarily an impediment to efficient contracting. In particular, if the contingencies can be described in such a way that a judge's ruling as to what has occurred is informative with respect to the agent's actions, then even a vague contract can achieve an efficient solution.

The analysis further shows that once a contract is good enough to achieve efficiency, there is no gain to improving it in terms of precision or information structure; hence, there could be little justification on efficiency grounds for the courts' distaste for vagueness. In fact, and contrary to the general perception in the literature, improvements in judicial accuracy can even be counter-productive by making it harder to achieve efficiency with simple, standard, or default contracts.

The article also provides further grounds upon which to argue for freedom of contract. In particular, as noted, it argues against the courts' refusing to

[16]See HERMALIN ET AL. [in press] for a survey on issues connected to freedom of contract.

enforce contracts deemed vague. It also provides a further argument against the penalty doctrine. Finally, although not formally analyzed, it should be clear from Section 4 that the shifting of court costs from one party to another cannot improve efficiency in matters involving contractual disputes.

Further work remains to be done. As discussed, the issue of cost shifting has not been investigated. Moreover, the article has only scratched the surface with respect to incorporating actual procedure into the analysis and in fully understanding the implications of the penalty doctrine.

Finally, two fundamental issues have arguably been ignored. First, as mentioned in the Introduction, if it is difficult to describe contractual contingencies, then how do the parties to the contract necessarily communicate their intentions to each other? Second, as noted, going to court is always an out-of-equilibrium event in this analysis. Although it is true that an overwhelming majority of contractual relations do not lead to court cases, some do. Clearly, there is some aspect of real-life contracting missing from this analysis insofar as actual court cases are inexplicable. Whether capturing that missing aspect materially affects the analysis presented above is, thus, one more topic for future work.

References

Anderlini, L. and L. Felli [1994], "Incomplete written contracts: Undescribable states of nature," *Quarterly Journal of Economics*, 109, 1085–1124.

Dye, R. A. [1985], "Costly contract contingencies," *International Economic Review*, 26, 233–250.

Fudenberg, D. and J. Tirole [1990], "Moral hazard and renegotiation in agency contracts," *Econometrica*, 58, 1279–1319.

Grossman, S. J. and O. D. Hart [1983], "An analysis of the principal-agent problem," *Econometrica*, 51, 7–46.

Hermalin, B. E., A. W. Katz, and R. Craswell [in press], "Contract law," in: A. Polinsky and S. Shavell (eds.), "Handbook of Law and Economics," Handbooks in Economics, Elsevier, Amsterdam, forthcoming.

Hermalin, B. E. and M. L. Katz [1991], "Moral hazard and verifiability: The effects of renegotiation in agency," *Econometrica*, 59, 1735–1753.

— [1993], "Defense procurement with unverifiable performance," in: J. Leitzel and J. Tirole (eds.), "Incentives in Procurement Contracting," Pew Studies in Economics and Society, Westview Press, Boulder, CO.

— [2000], "Corporate diversification and agency," in: P. J. Hammond and G. D. Myles (eds.), "Incentives, Organization, and Public Economics: Papers in Honour of Sir James Mirrlees," Oxford University Press, Oxford.

Holmstrom, B. [1979], "Moral hazard and observability," *Bell Journal of Economics*, 10, 74–91.

Katz, A. W. [1990], "Your terms or mine? The duty to read the fine print in contracts," *RAND Journal of Economics*, 21, 518–537.

Kim, S. K. [1995], "Efficiency of an information system in an agency model," *Econometrica*, 63, 89–102.

Posner, R. A. [1973], "An economic approach to legal procedure and judicial administration," *Journal of Legal Studies*, 2, 399–458.

Rasmusen, E. B. [2001], "Explaining incomplete contracts as the result of contract-reading costs," *Advances in Economic Analysis & Policy*, 1, Article 2.

Segal, I. R. [1999], "Complexity and renegotiation: A foundation for incomplete contracts," *Review of Economic Studies*, 66, 57–82.

Spier, K. E. [1992], "Incomplete contracts and signalling," *RAND Journal of Economics*, 23, 432–443.

Townsend, R. M. [1979], "Optimal contracts and competitive markets with costly state verification," *Journal of Economic Theory*, 20, 265–293.

Walter A. Haas School of Business & Department of Economics
University of California at Berkeley
S545 Student Services Building #1900
Berkeley, California 94720–1900
United States
E-mail:
hermalin@berkeley.edu