Courts of Law and Unforeseen Contingencies

Luca Anderlini
Georgetown University

Leonardo Felli*
London School of Economics

Andrew Postlewaite
University of Pennsylvania

We study a contracting model with unforeseen contingencies in which the court is an active player. Ex ante, the contracting parties cannot include the risky unforeseen contingencies in the contract they draw up. Ex post, the court observes whether an unforeseen contingency occurred and decides whether to void or uphold the contract. If the contract is voided by the court, the parties can renegotiate a new agreement ex post. There are two effects of a court that voids contracts. The parties’ incentives to undertake relationship-specific investment are reduced, and the parties enjoy greater insurance against the unforeseen contingencies that the ex ante contract cannot account for. In this context, we fully characterize the optimal decision rule for the court. The behavior of the optimal court is determined by the trade-off between the need for incentives and the gains from insurance that voiding in some circumstances offers to the agents.

1. Introduction

1.1 Motivation

Courts regularly intervene in contracts at the behest of one of the contracting parties to void or otherwise modify an agreement the parties have signed. One

---

*Department of Economics, London School of Economics, Houghton Street, London WC2A 2AE, United Kingdom. Email: lfelli@econ.lse.ac.uk.

The authors thank Cristopher Harris, Oliver Hart, Jean-Jacques Laffont, Gaoquan Liu, Eric Maskin, Meg Meyer, Kevin Roberts, Alan Schwartz, Ilya Segal, Kathy Spier, Jean Tirole, and seminar participants at Certosa di Pontignano (Siena), Nuffield College (Oxford), Institut d’Économie Industrielle (Toulouse), European Summer Symposium in Economic Theory 2002 (Gerzensee), Econometric Society European Meeting 2002 (Venice), the Economic and Social Research Council Seminar in Game Theory (Kenilworth), Harvard, LSE, Penn Law School, USC, and Yale Law School for very helpful discussions and comments. Luca Anderlini and Leonardo Felli are grateful to the Economic and Social Research Council (U. K.) for financial support (Grant R000237825). Andrew Postlewaite acknowledges financial support from the National Science Foundation. This article was started while Leonardo Felli was visiting the Department of Economics at the University of Pennsylvania. He is grateful for their generous hospitality.

Advance Access publication September 13, 2006
© The Author 2006. Published by Oxford University Press on behalf of Yale University. All rights reserved. For permissions, please email: journals.permissions@oxfordjournals.org
justification for courts overriding voluntary agreements is to insure the parties against changes in the environment between the time the agreement was made and the time when it is to be consummated. Changes in the environment can lead to changes in the costs and benefits to the parties involved that expose them to risks that they prefer to avoid.\(^1\) The possibility of renegotiation protects the parties from carrying out Pareto-dominated transactions but not from the fluctuations in utility that stem from the uncertainty in the underlying environment.

If the parties foresee all relevant contingencies and agree on the optimal transactions given them, these can be included in the contract, thus providing protection from these risks. Both common sense and court decisions suggest that such foresight is unreasonable, however. Regardless of the parties’ experience and care in designing their contract, there will always be residual risk they face due to “unforeseen contingencies.”

There is considerable ambiguity about the meaning of unforeseen contingencies; we will discuss the term briefly before proceeding further. We take the position that parties can perfectly foresee the possibility of various contingencies but are unable to describe the circumstances in sufficient detail to include all relevant contingencies in their contract.\(^2\) When the contracting parties understand that they are unaware of all significant potential events, the question arises as to how they can protect themselves against the risks they face when committing to a necessarily incomplete contract.

In this article, we take the view that although contracting parties are unable to identify all relevant contingencies ex ante, it may be clear both to the parties and outsiders that the circumstances at the time the contract calls upon one of the parties to act differ materially from those envisioned at the time the contract was written. In this event, a court can make such a determination and void the contract in order to provide insurance the parties arguably desire but cannot effect on their own. A primary goal of this article is to model formally the effects of ex ante unforeseen events that might be recognized as such ex post.

A court that voids contracts in this way may provide desirable insurance but not without cost. A central benefit of a contract is a guarantee that parties will receive a return for investments that have specific value in their relationship. Without a guarantee, an individual has a diminished incentive to invest because he or she may obtain only a portion of the benefits stemming from investment under an ex post (re)negotiated outcome. Courts that void contracts to provide insurance do so at the cost of reducing the ability to provide incentives for an efficient level of ex ante investment.

We develop and analyze a model of a buyer and a seller who contract in an environment that includes an active court whose role is to determine which

---

\(^1\) See Kaplow and Shavell (2002, sec. 4) for a general discussion of incomplete contracts and enforcement.

\(^2\) Al-Najjar et al. (2006) provide a formal model that fits this view of what an unforeseen contingency is. They use the term “undescribable” to label these contingencies. Concerning the possible meanings of the term unforeseen contingencies that have been discussed in previous literature, see also Tirole (1999) and the survey by Dekel et al. (1998).
contracts to void and which to uphold. In our model, a court can coordinate and guide contracting parties by means of precedents that shape the contracting parties’ expectations about future rulings. We assume that the court maximizes ex ante expected gains from contracting and characterize the optimal policy, which is to void contracts in events that are deemed ex post to impose a high level of uninsurable risk on the contracting parties.

In the simple setup that we analyze, the interests of all participants are aligned. Ex ante, the objective function of the court is not in conflict with the expected utility of either of the trading parties. This, in turn, implies that the parties could attempt to replicate the behavior of the optimal court using private means. However, this will only be true in the simplified setup that we deal with here. For example, if there is any asymmetric information between the contracting parties, informational externalities would arise and this conclusion would not necessarily hold.

1.2 Relation to the Literature

The seminal works on incomplete contracts by Grossman and Hart (1986) and Hart and Moore (1990) took as given the existence of contingencies that may occur after the signing of a contract but cannot be described at the time the parties contract. The inability to describe all relevant contingencies, and make contract terms a function of them, affects agents’ incentives. When contracts are incomplete, the contracting parties may find it optimal to renegotiate the terms of trade in the event that certain contingencies arise. Agents whose investments are sunk at this time will not receive the full benefits of those investments. This holdup problem leads to inefficient initial investments. In summary, incomplete contracts may make it impossible to avoid inefficient outcomes.

A number of articles have shown that the amount of inefficiency, however, is not fixed. Grossman and Hart (1986) and Hart and Moore (1990) show that the ownership structure of physical assets can affect investment incentives and, hence, efficiency; Bernheim and Whinston (1998) show that if it is impossible to contract over some part of a relationship, it may be optimal to be less specific than is possible in other parts of that relationship; Aghion and Tirole (1997) and Rajan and Zingales (1998) show that the distribution of authority and power in a firm can affect efficiency when complete contracts are impossible.

Both the original work, illustrating how incomplete contracts can precipitate inefficiency, and the subsequent work, demonstrating how institutional design can ameliorate that inefficiency, essentially ignore the role of a court in

3. Throughout the rest of the article, we use the terms uphold and enforce (a contract) in a completely interchangeable way. Our court does not engage in “gap filling” in that it only rules on whether a contract should be voided or not, rather than attempting to impose new terms on the parties. This seems to be the predominant view of how actual courts behave and of how they should in fact behave (Kull 1991).

4. In Anderlini et al. (2005), we analyze a model where there is an informational externality between different types of contracting parties.
adjudicating and enforcing contracts that are written.\textsuperscript{5} The inefficiencies analyzed in the articles discussed above might be diminished by a court that can ameliorate them through various forms of intervention.\textsuperscript{6} Stated more strongly, the work on incomplete contracts is “partial equilibrium,” analyzing a subset of agents’ behavior taking as fixed the behavior of agents outside the model (the courts), without investigating whether the assumed fixed behavior of the outside agents is in fact optimal. Maskin and Tirole (1999) make this point most forcefully by showing that in a standard incomplete contracting model, the existence of undescribable (unforeseen) contingencies does not affect the set of payoff outcomes that can be achieved through contracting, if one allows a court with large discretionary authority. This is shown by exhibiting a mechanism capable of generating as equilibrium any payoffs that could be achieved with complete contracts. This mechanism does not mean that contract incompleteness is irrelevant, however. Maskin and Tirole (1999) expand the scope of interaction between the contracting parties and the court to include arbitrary systems of communication and clearly go beyond anything seen in practice.

Our article incorporates an active court, the scope of which is limited, in contrast to both the traditional literature and Maskin and Tirole (1999). We provide a detailed specification of undescribable contingencies, including the information available to a court at the time performance is called for. The contracts that parties write differ from those they would write if courts did nothing more than passively enforce the contracts that are written. Despite the inclusion of a more active court, the basic message of the incomplete contracting literature remains: contracts will still be incomplete and the incompleteness causes inefficiency.

There is a relatively large literature on the effect of the rules courts use on the actions of those governed by the rules. For example, there is a substantial body of analysis comparing the incentive effects of strict liability with the incentive effects of a negligence rule in tort theory and comparisons of different remedies for breach in contract theory.\textsuperscript{7} Our analysis differs from this work in two ways. First, these literatures focus largely on particular rules that are used in practice and compare the incentive effects of these rules in different environments. In contrast, we consider a richer set of rules, with courts optimizing across that set; our framework admits more easily the formulation of alternative rules to those already in existence. The second difference is that earlier work is typically concerned with comparisons between qualitatively different rules, whereas our court must make quantitative decisions, such as the threshold for which unforeseen contingencies will change the court’s decision of whether or not to void the contract.\textsuperscript{8}

\textsuperscript{5} A “minimal” court is assumed to exist to force the parties to perform according to the contract as originally written.
\textsuperscript{6} See Eggleston et al. (2000) for a discussion on the role of courts in interpreting and enforcing contracts.
\textsuperscript{7} See Kaplow and Shavell (2002) for a discussion of these literatures.
\textsuperscript{8} See Kaplow (2000) and Kaplow and Shavell (2002).
A major benefit of formally incorporating the court is that it allows a richer analysis of contracting. In addition, it provides the structure for a serious examination of what precisely a court might do. In this article, we restrict attention to particularly simple rules a court can follow, namely, to determine the circumstances under which a contract will be voided.

1.3 Court Practices

We discussed in Section 1.2 the relation of our work to previous literature. Before proceeding to our formal model, it is useful to also discuss the relation between our work and actual court practices to illustrate that courts insure contracting parties along the lines that we argue are optimal. They will discharge a party’s obligation to perform under a contract based on the emergence of risks that were not foreseen at the time the contract was entered into under some conditions. There are several categories of intervening events that might be the basis for excusing performance, two of which are similar to the unforeseen contingencies that are the focus of this article. The first is impracticability of performance; this occurs when unanticipated events subsequent to contracting make the promised performance extremely burdensome economically. The second category is termed frustration of purpose. One view of the frustration doctrine is that it will “… excuse performance where performance remains possible, but the value of the performance to at least one of the parties and the basic reason recognized by both parties for entering into the contract have been destroyed by a supervening and unforeseen event.”

The court intervention proposed in this article that voids contracts under some circumstances can be considered to be of either of these two types. Performance is clearly not impossible since ultimately the contracted transaction is consummated; the voiding of the contract serves only to relieve one or the other of the parties from an abnormally negative consequence resulting from supervening events. Frustration of purpose has been applied in a manner very similar to that proposed in this article. Small risks will not be the cause for voiding the contract, but sufficiently large risks will be.

It is, of course, the very essence of contract that it is directed at the elimination of some risks for each party in exchange for others. Each receives the certainty of price, quantity, and time, and assumes the risk of changing market prices, superior opportunity, or added costs. It is implicit in the doctrine of impossibility (and the companion rule of ‘frustration of purpose’) that certain risks are so unusual and have such severe consequences that they must have been beyond the scope of the assignment of risks inherent in the contract, that is, beyond the agreement made by the parties. To require performance in that case would be to grant the

---

promisee an advantage for which he could not be said to have bargained in making the contract. . . . The question is, given the commercial circumstances in which the parties dealt: Was the contingency which developed one which the parties could reasonably be thought to have foreseen as a real possibility which could affect performance? Was it one of that variety of risks which the parties were tacitly assigning to the promisor by their failure to provide for it explicitly? If it was, performance will be required. If it could not be so considered, performance is excused.10

This case is not an isolated instance. Williston (1938) on Contracts, a standard reference to the interpretation of contracts, has this to say:

The important question is whether an unanticipated circumstance has made performance of the promise vitally different from what should reasonably have been within the contemplation of both parties when they entered into the contract. If so, the risk should not fairly be thrown upon the promisor.

We focus on the case in which the events that alter the costs and benefits to the parties of performance as specified in the contract were unforeseeable, but courts have typically taken a less stringent attitude of the circumstances in which supervening events could warrant excusing performance. Specifically, it is not necessary that a supervening event be literally unforeseeable but, rather, that it was in fact unforeseen; this is illustrated by the following two cases.

The question we answer here is not whether the destruction of the forest’s regenerative capacity should have been considered at the time of contracting but, rather, whether it was considered. There is nothing in the contract nor in the parties’ dealings to suggest that the parties ever presumed more than a continuance of the conditions necessary to give purpose to a selective cut contract. In short, the contract did not address the conditions that arose; hence, further performance under the contract is excused.11

. . . it would be untenable to conclude that the parties intended that the [plaintiffs] should assume the risk of an adverse tax ruling simply because such a ruling was, in a sense, ‘foreseeable’ and because the


11. See footnote 9. Spalding had a contract to harvest timber on U.S. government land that the Bureau of Land Management (BLM) cancelled after a fire on the adjacent property required unforeseen remedial action. The court upheld BLM’s right to cancel.
contract did not expressly excuse performance in the event of its occurrence.12

The main point of this article is to demonstrate how a court can increase welfare by excusing performance in some situations where unforeseen events have dramatically changed the consequences of performance for one of the parties to a contract. We point out that the role for courts that we advocate fits within the U.C.C.: “Delay in delivery or non-delivery in whole or in part by a seller . . . is not a breach of his duty under a contract for sale if performance as agreed has been made impracticable by the occurrence of a contingency the non-occurrence of which was a basic assumption on which the contract was made . . .” U.C.C. 2-615(a). Comment 4 to this provision provides more substance:

Increased cost alone does not excuse performance unless the rise in cost is due to some unforeseen contingency which alters the essential nature of the performance. Neither is a rise or a collapse in the market in itself a justification, for that is exactly the type of business risk which business contracts made at fixed prices are intended to cover. But a severe shortage of raw materials or of supplies due to a contingency such as war, embargo, local crop failure, unforeseen shutdown of major sources of supply, which either causes a marked increase in cost or altogether prevents seller from securing supply necessary for his performance, is within contemplation of this section.

The optimal court that we derive subsequently is entirely consistent with this. The code allows for excuse when there is a “marked” increase in the cost to the seller. Our model will call for excuse when there is an unforeseen contingency that results in a significant difference between the actual and the expected cost. The value of the insurance that results from excusing performance outweighs the diminished incentives to invest that accompany voiding the contract. In sum, contract law is amenable to the rule that we will derive.13

1.4 Outline

The plan of the rest of the article is as follows: In Section 2 we describe the model in full detail, and we comment on the assumptions we make. We

12. West Los Angeles Institute for Cancer Research, Appellant, v. Ward Mayer et al., Appellees, No. 19551; U.S. Court of Appeals for the Ninth Circuit; 366 F.2d 220; 1966 U.S. App. LEXIS 5088. In August 1951, Ward Mayer and his wife and son contracted to sell the business to the West Los Angeles Institute for Cancer Research, a tax-exempt entity. The transaction was patterned after the sale and leaseback agreements previously approved by the Internal Revenue Service (IRS). The IRS rejected the tax premises upon which the transaction was based, and the Mayers sued to recover the property. The district court granted the relief sought on the ground that the sale and leaseback arrangement was frustrated by the revenue ruling.

13. We are grateful to a referee for pointing this out.
characterize in Section 3 the equilibrium contract that the parties to the trade will choose for the general court’s decision rules, whereas in Section 4 we present the main result of the article: the characterization of the optimal decision rule for the court. In Section 5 we discuss several leading cases dealing with frustration and impracticability in the context of our optimal court. Section 6 concludes the article. For ease of exposition, we have relegated all proofs to the Appendix.

2. The Model

As mentioned in Section 1, we are interested in courts that have a role in trading off parties’ incentive to invest with their desire for insurance in the event of unforeseen contingencies. To investigate this trade-off, we consider a simple buyer and seller model.

For insurance to have any benefit, at least one of the parties must be risk averse; we assume a risk-neutral buyer and risk-averse seller. The buyer and seller trade a widget; the risk they face is that the cost and benefit of the widget are uncertain at the time they contract. The uncertainty about costs and benefits captures the idea that there is a “normal” cost and benefit, $c_N$ and $v_N$, but that both parties are aware that there is a possibility that an unforeseen contingency could give rise to high levels of costs and benefits: $c_H$ and $v_H$. For simplicity, we assume that the gains from trade are constant, that is,

$$\Delta = v_H - c_H = v_N - c_N.$$ 

Hence, it is efficient to trade whether the costs and benefits are normal or high. This assumption is made for tractability. Our results would not qualitatively change if the costs and benefits were not perfectly correlated or if the magnitude of the gains from trade was variable. We assume that $c_H \geq c_N$.

Before going on, we will illustrate the components of the model with reference to Transatlantic Financing Corp. v. United States (1966), a case involving commercial impracticability. In this case, the defendant chartered a ship operated by Transatlantic to carry a cargo of wheat from the United States to Iran. Six days after the ship left port, the Suez canal was closed by the Egyptian government, forcing the ship to reroute around the Cape of Good Hope. Transatlantic sued for additional compensation for its increased expenses. Put into our model, the United States is the risk-neutral buyer and Transatlantic the risk-averse seller. The normal cost is the cost of transporting the wheat via the Suez canal, whereas the high cost is the cost of transporting via the longer route.

We assume that the buyer has all the bargaining power ex ante when a contract is proposed. In other words, the equilibrium contract is the result of a take-it-or-leave-it offer from the buyer to the seller. Ex post, in some instances, renegotiation will take place. We assume that the seller has all the bargaining power in the ex post renegotiation: if renegotiation occurs, the seller makes

a take-it-or-leave-it offer to the buyer. The assumption that both ex ante and ex post, one or the other of the parties has all the bargaining power is for expository ease; none of our results depends qualitatively on bargaining power being absolute for one or the other. Our results would not hold, however, if the buyer has all the bargaining power ex post.

A central issue in this article is how unforeseen contingencies are modeled, and we will discuss verbally our approach before describing the formal modeling. We assume that ex post, the court, as well as the parties, can recognize some events that are out of the ordinary. For example, all parties recognize and agree that the events of 9/11 were, in some sense, unforeseen. However, it is likely that for every possible unfolding of events, one could claim that there is some unforeseen component, so excusing performance whenever there has been an unforeseen event cannot be a useful rule. We assume that the court can “categorize” events ex post in the following sense. For any given realized event, the court will understand that if performance were excused in that instance, consistency (i.e., following precedent) would lead it to excuse performance in similar circumstances in the future. Assuming that the court can categorize events ex post essentially means that the court understands the consequences of excusing performance in the present contractual arrangement on future contracting parties, if the court wishes to be consistent.

In addition to the court’s categorizing events, we assume that the court, implicitly or explicitly, assigns a probability to the category of events that are similar to the events at hand. That is, the court understands that if it desires to be consistent, excusing performance in the present contract will result in excusing performance in future contracts with the probability the court assigns to the category of events similar to the case before it. The basic notion, then, is that courts make decisions at the ex post stage but understand that, based on the court’s decision, future contracting parties will make inferences about the probability that performance will be excused.

In our model, the presumption would be that the closing of the Suez canal was unforeseen by both Transatlantic and the United States and recognized as so by the courts after the fact. Our assumption is that if courts are consistent, however they will treat the suit between Transatlantic and the United States, they will treat “similar” future cases in the same way. This leaves open what cases would be similar—future cases in which the Suez is again closed? Future cases in which some canal is closed? Future cases in which some unforeseen event results in increased transportation costs? In effect, our assumption that the court can categorize the event “Suez closed by Egyptian government” is an assumption that the court can assign a probability that the decision in the case at hand will affect future cases. Although we do not include it in our model, the written opinion accompanying the court’s decision will determine to a large extent what future cases would be deemed similar in practice.

We formalize these ideas next. With probability \((1 - q)\), we assume that the world is in a “normal” state. In this case, the cost of the widget to the seller is \(c_N\), whereas the value of the widget to the buyer is \(v_N = c_N + \Delta\). With the complementary probability \(q\), the world is in a state that will be deemed to
be “exceptional,” meaning that, ex post, it will be deemed to have been unforeseen. In the case of an unforeseen state, the cost of a widget to the seller and the benefit of the widget to the buyer are uncertain.

Our aim is to model a court that trades off the diminished incentive effects resulting from voiding contracts with the insurance gains such voiding generates. Categorization of an unforeseen event and assigning that category a probability allow the court to measure the incentive costs of excusing performance. What remains is a specification of the information the court would need to gauge the insurance benefits of voiding. There cannot be a role for a court that excuses performance if the court can precisely observe the payoffs to the parties; in such a world, the parties could simply specify a contract price for any change in payoffs resulting from unforeseen contingencies, thereby providing full insurance within the contract itself. Thus, a necessary condition for a court to have a role that includes excusing performance in some unforeseen events, but not in all, is that the court must have some idea of the magnitude of the effect of the unforeseen contingency on payoffs but not observe precisely (and hence condition on) these payoffs. For example, the increased costs to Transatlantic due to the Suez closure include the opportunity cost of the vessel for the increased time, which the court might be unable to determine with more precision than that they were very large. We model the court’s information in the simplest way to capture this: we will assume that, although the court does not observe whether the state of the world is normal or exceptional, the court can assess the magnitude of the impact that this unforeseen contingency has on the parties’ payoffs. Specifically, in an exceptional state, the cost of the widget to the seller is \( c_H(\theta) \) (and hence, from the assumption that the gains from trade are constant, the buyer’s valuation is \( v_H(\theta) = c_H(\theta) + \Delta \)), where \( \theta \) parameterizes the magnitude of the effect that an unforeseen state has on the cost and benefit. We further assume that \( \theta \) is independent of whether the world is in a normal state or in an exceptional one, and it is uniformly distributed in the interval \([0, 1]\). The court does observe the realization of \( \theta \) but does not observe whether the world is in a normal state or in an exceptional one. The value of \( \theta \) reveals to the court the magnitude of the impact of unforeseen contingencies.

If we denote by \( g(\theta) \) the difference between \( c_H \) and \( c_N \) for a given \( \theta \), we have that

\[
c_H(\theta) = c_N + g(\theta).
\]

We also take \( g \) to be differentiable and to satisfy \( g(\theta) = 0 \) for every \( \theta \in [1/2, 1] \) and \( \lim_{\theta \to 0} g(\theta) = \infty \). Thus, for \( \theta \in [1/2, 1] \), there is no risk associated with the cost. This risk is present for \( \theta \in [0, 1/2] \) and increases without bound as \( \theta \) approaches zero.

To summarize, the parties face a risk at the time they contract that as a consequence of an unforeseen contingency, the cost and value of the widget will be abnormally high at the time production and delivery are to take place. Ex post, unforeseen contingencies will only be recognized by the contracting parties. The court will know the variance of costs associated with the unforeseen contingency but not the actual payoffs to the parties. We assume that the
parties cannot contract on $\theta$, the effect that an unforeseen contingency has on the parties’ payoffs. They can only rely on the court to be protected against the uncertainty associated with unforeseen contingencies (if this is what the court finds optimal to do).

This risk can be avoided by not contracting ex ante and simply contracting after the state is realized. So that there is a benefit to contracting ex ante, we assume that the buyer can undertake an ex ante, noncontractible, investment $e \in [0, 1]$ at a cost $\psi'(e)$, where we assume that $\psi$ is twice differentiable, convex, and satisfies $\psi'(0) = 0$ and $\lim_{e \to 1} \psi'(e) = +\infty$. A buyer’s investment of $e$ increases the value to him or her of the widget of an amount $eR$. Consequently, if the buyer chooses the level of relationship-specific investment $e$, his or her value of the widget is $eR + \Delta + c_i$, where $i \in \{N, H\}$.

Since the buyer is risk neutral, he or she maximizes expected profit, minus the convex cost of investment as above. The risk-averse seller maximizes the expected value of a strictly increasing twice differentiable $V : \mathbb{R} \to \mathbb{R}$. To embody risk aversion, we also take $V$ to be strictly concave so that $V'' < 0$.

The timing of the model can be specified as follows: The parties form beliefs about the court’s rule for enforcing or excusing performance, based on the court’s past record (i.e., based on the precedents). Negotiation then takes place between the contracting parties. Recall that the buyer has all the bargaining power at this stage; hence, negotiation is a simple take-it-or-leave-it offer of a contract from the buyer to the seller. A contract may specify an ex ante transfer; if it does, the transfer is made immediately after a contract is agreed on. After the negotiation of an ex ante contract, the buyer chooses the level of specific investment $e$ that increases the value of the widget to him or her by $eR$.

The state of the world—whether the parties trade in a normal or in an exceptional state—is then realized and is observed by both parties to the contract. Moreover, we also assume that the parties to a contract observe the exact value of the cost $c_i$, $i \in \{N, H\}$. Should the court become involved, as we discussed above, it does not observe whether the parties operate in a normal or in an exceptional state but does know the magnitude of the impact that an unforeseen contingency might have on the parties’ welfare. In other words, the court observes the realization of $\theta$. Either party can bring the other side to court, and if this occurs, the court is assumed to mandate or excuse performance consistent with past rulings.

In the case in which the court decides to void the existing contract, renegotiation takes place between the buyer and the seller. Renegotiation is modeled as a take-it-or-leave-it offer from the seller to the buyer of a price at which

15. Notice that if the transfer were “refundable” if the contract is voided, then we could simply incorporate it in the trade price that the contract specifies. Hence, a nonrefundable ex ante transfer like the one we consider allows for a richer set of possible contracts. With respect to the actual behavior of courts, it is argued that when courts determine that contracts should not be enforced as written, “... parties will be permitted to walk away from their bargain, without damages for reliance or restitution for benefits conferred” (Kull 1991).
to trade. When renegotiation occurs, following the court’s decision to void the contract, the parties’ outside options are represented by the payoffs associated with no trade. These payoffs are normalized to zero.

Finally, trade occurs according to the terms of the original contract, if the court decides to enforce it, or according to the terms of the renegotiated agreement, if the court decides to void the original ex ante contract.

3. The Optimal Ex Ante Contract

Given our assumptions above, the parties to a contract can only specify in an ex ante contract a constant price at which to trade, $p$, and an ex ante transfer from the buyer to the seller, $t$. If the parties decide to draw up such an ex ante contract, it is then left to the court to determine whether or not to protect them against the possibly very large risk associated with the unforeseen contingencies.

We identify the optimal court’s ruling solving the model backward from the last stage. We begin with the renegotiation that follows the court’s decision to void the contract. Denote as $\hat{e}$ the given level of investment chosen by the buyer. Since the seller has all the bargaining power at the renegotiation stage, he will receive all the gains from trade available to the parties; these of course total $\hat{e}R + \Delta$.

Consider now the court’s decision if one of the two parties brings the other to court. Without loss of generality, we can specify the court’s decision rule to be a set $\mathcal{E} \subseteq [0,1]$. The court enforces all contracts when $\theta \in \mathcal{E}$ and voids all contracts otherwise. In other words, when the impact of the unforeseen contingency on the parties’ welfare is too high, the court provides the parties with insurance by voiding the existing contract.

The court determines $\mathcal{E}$ prior to the parties’ negotiation of the ex ante contract. In other words, the parties infer the court’s decision rule from precedents when they decide which ex ante contract to draw up.

Before we analyze the parties’ negotiation of the ex ante contract, we need to specify the seller’s and buyer’s outside options if the ex ante negotiation breaks down. Notice that even in the absence of an ex ante contract the parties can still trade the widget ex post. Recall that in any ex post negotiation the seller has all the bargaining power. Hence, in any ex post agreement, he or she appropriates all the gains from trade and receives utility $V(\hat{e}R + \Delta)$, where $\hat{e}$ is the level of specific investment chosen by the buyer in the absence of any ex ante contract. The buyer receives a zero share of the gains from trade.

Notice that the advantage for the parties to trade ex post is that they do not face any uncertainty, and therefore, the seller is provided with full insurance. However, since the returns to the buyer from his ex ante investment are zero, he will choose an investment level such that $\psi'(\hat{e}) = 0$. In other words, when trade takes place ex post, because there is no ex ante contract, the buyer has no incentive to invest: $\hat{e} = 0$. We can then conclude that in the absence of an

---

16. Of course, $\mathcal{E}$ is assumed to be a Lebesgue-measurable set. As we will see in Lemma 2, it will never be optimal for the court to void a contract if it observes $\theta \in [1/2, 1]$. However, the general specification of the court’s decision rule must allow for this possibility.
ex ante contract the buyer’s payoff is zero whereas the seller’s level of utility is \( V(\Delta) \). The seller is fully insured, but no relationship-specific investment is undertaken by the buyer. The buyer’s outside option when the ex ante contract is negotiated is zero, whereas the seller’s outside option is \( V(D) \).

Next, we turn to the parties’ negotiation of the ex ante contract. Recall that ex ante the buyer makes a take-it-or-leave-it offer to the seller of a contract \((p, t)\). Given the court’s decision rule \( \mathcal{E} \) and a level of investment \( \hat{e} \), the seller’s expected utility associated with \((p, t)\) can now be written as follows:

\[
V_{\mathcal{E}}(p, t, \hat{e}) = \int_{\mathcal{E}} [qV(p + t - c_H(\theta)) + (1 - q)V(p + t - c_N)]d\theta + \int_{([0,1]\setminus\mathcal{E})} V(\hat{e}R + \Delta + t)d\theta.
\]  

(2)

Notice that the first integral in equation (2) refers to the case in which the contract is upheld by the court. The second integral in equation (2) captures those cases in which the court voids the ex ante contract.

Taking again as given the court’s decision rule \( \mathcal{E} \) and a level of investment \( \hat{e} \), the buyer’s expected profit associated with \((p, t)\) can be computed as follows:

\[
B_{\mathcal{E}}(p, t, \hat{e}) = \int_{\mathcal{E}} [q(\hat{e}R + \Delta + c_H(\theta) - p) + (1 - q)\hat{e}R + \Delta + c_N - p]d\theta - t - \psi(\hat{e}).
\]  

(3)

If we set \( \theta_{\mathcal{E}} = \int_{\mathcal{E}} d\theta \), recalling that \( c_H(\theta) = c_N + g(\theta) \), the payoffs in equations (2) and (3) can be rewritten more simply as

\[
V_{\mathcal{E}}(p, t, \hat{e}) = \int_{\mathcal{E}} [qV(p + t - c_N - g(\theta)) + (1 - q)V(p + t - c_N)]d\theta + (1 - \theta_{\mathcal{E}})V(\hat{e}R + \Delta + t)
\]  

(4)

and

\[
B_{\mathcal{E}}(p, t, \hat{e}) = \theta_{\mathcal{E}}[\hat{e}R + \Delta + c_N - p] + q \int_{\mathcal{E}} g(\theta)d\theta - t - \psi(\hat{e}).
\]  

(5)

From equation (5), it is immediate that given \((p, t)\) and the court’s decision rule \( \mathcal{E} \), the buyer will select a level of relationship-specific investment \( \hat{e} \) such that

\[
\psi'(\hat{e}) = \theta_{\mathcal{E}}R.
\]  

(6)

We can now state the buyer’s optimization problem for choosing an ex ante contract. Given the court’s decision rule \( \mathcal{E} \), the buyer’s take-it-or-leave-it offer to the seller is the solution, if it exists, to the following problem.

\[
\max_{p, t, \hat{e}} B_{\mathcal{E}}(p, t, \hat{e})
\]

s.t. \( V_{\mathcal{E}}(p, t, \hat{e}) \geq V(\Delta) \),

\[
B_{\mathcal{E}}(p, t, \hat{e}) \geq 0,
\]

\[
\psi'(\hat{e}) = \theta_{\mathcal{E}}R,
\]  

(7)
where the first two constraints guarantee that it is optimal for both the seller and the buyer to sign an ex ante contract rather than to trade ex post. If the feasible set of problem (7) is in fact empty, then no ex ante contract will be signed and trade will take place ex post. However, when the court’s decision rule is chosen so as to maximize the parties’ welfare, an ex ante contract will be signed. We state the following without formal proof.

Remark 1. For some specifications of the court’s decision rule, the feasible set of problem (7) is clearly not empty, and the maximized value of the objective function is strictly positive.

For example, suppose that the court never voids the contract if $\theta \in [1/2, 1]$ and always voids the contract if $\theta \in [0, 1/2)$ so that $\mathcal{E}=[1/2,1]$. In this case, the agents do not face any uninsurable risk from unforeseen contingencies and can take advantage of a fixed price for the case $\theta \in \mathcal{E}$ so that the buyer will undertake a positive amount of relationship-specific investment $\hat{e}$ such that $\psi'(\hat{e}) = R/2$. It is clear that in this case there is an ex ante contract that is preferred to no contract by both the buyer and seller.\footnote{When an ex ante contract is preferred to trading ex post, it is immediate by standard arguments that the solution to problem (7) is in fact unique.}

Notice that if the court’s decision rule is such that $\theta \in \mathcal{E} = 0$ we obtain a trivial special case, in which the court always voids the contract, the expected profit of the buyer is zero, and the expected utility of the seller is $V(\Delta)$, whatever the contract $(p, t)$. In this case, since both parties are indifferent, we assume that they prefer to implement the same outcome by having no contract at all.

Our characterization of the optimal contract given the court’s decision rule can now be summarized as follows:

Proposition 1. Let a decision rule $\mathcal{E}$ for the court be given and assume that it is such that it is optimal for the parties to draw up an ex ante contract. Let the optimal ex ante contract given $\mathcal{E}$—the solution to problem (7)—be denoted by $(p^*_E, t^*_E)$, with $\hat{e}_E$ the associated level of investment. Then $p^*_E, t^*_E,$ and $\hat{e}_E$ satisfy

\[
\begin{align*}
\int_{\mathcal{E}} [qV'(p^*_E + t^*_E - c_H(\theta)) + (1 - q)V'(p^*_E + t^*_E - c_N)]d\theta \\
= \theta EV'(\hat{e}_E R + \Delta + t^*_E)
\end{align*}
\]  

and hence

\[ p^*_E - c_N \geq \hat{e}_E R + \Delta. \tag{9} \]

Moreover, the transfer $t^*_E$ is such that

\[ V(\mathcal{E}(p^*_E, t^*_E, \hat{e}_E)) = V(\Delta). \tag{10} \]

Equality (10) of Proposition 1 is a simple consequence of the fact that the seller’s expected utility is increasing in $t$, whereas the buyer’s expected profit is a decreasing function of $t$.\footnote{When an ex ante contract is preferred to trading ex post, it is immediate by standard arguments that the solution to problem (7) is in fact unique.}
The intuition behind equations (8) and (9) of Proposition 1 is not hard to explain. In those states in which the contract is renegotiated, the seller necessarily gets a payoff (on top of the transfer) of \( \hat{e}_E R + \Delta \). The price \( p^*_E \) is chosen so as to provide the seller with the optimal partial insurance against the fluctuations of cost between \( c_N \) and \( c_H(\theta) \) that occur when the court upholds the contract. This means equating the seller’s expected marginal utility in this eventuality with the seller’s marginal utility that he or she achieves when the contract is voided by the court. Since the seller’s marginal utility is decreasing, this implies that the price \( p^*_E \) minus the lowest cost \( c_N \) must be above \( \hat{e}_E R + \Delta \).

4. The Court’s Optimal Decision Rule

We are now equipped with the characterization (Proposition 1) of the optimal contract \((p^*_E, t^*_E)\) given an arbitrary decision rule \( E \) for the court. This is enough to proceed to characterize the court’s optimal decision rule.

Recall that our court is a “Stackelberg leader.” Through precedents, its decision rule is effectively announced to the parties. Taking into account the effect of its choice of rule on the parties’ behavior, the court then acts so as to maximize their welfare. From Proposition 1 we know that as a result of the fact that the buyer makes a take-it-or-leave-it offer of an ex ante contract to the seller, the seller’s expected utility will be \( V(\Delta) \), regardless of the court’s decision rule. Therefore, the court’s decision rule can be characterized as the solution to the problem of maximizing the buyer’s expected profit subject to appropriate constraints.

The court’s maximization problem can be written as follows: Choose the set \( E \) of \( \theta \)’s in which the contract is upheld so as to solve

\[
\max_{E} B_{E}(p^*_E, t^*_E, \hat{e}_E)
\]

s.t. \( V_{E}(p^*_E, t^*_E, \hat{e}_E) \geq V(\Delta) \),
\[
B_{E}(p^*_E, t^*_E, \hat{e}_E) \geq 0,
\]

where \((p^*_E, t^*_E)\) is the optimal ex ante contract characterized in Proposition 1 and \( \hat{e}_E \) is the associated level of investment.

We begin with two partial characterizations of the court’s optimal decision rule. Our first claim asserts that provided a solution to problem (11) exists, it will be such that the court never voids the parties’ ex ante contract when \( \theta \in [1/2, 1] \); it is never optimal for the court to void the contract if, given \( \theta \), the parties face no risk.

**Remark 2.** It is optimal for the court to enforce the contract whenever \( \theta \in [1/2, 1] \). More formally, assume that a solution to problem (11) exists. Then any solution \( E^* \) to this problem satisfies

\([1/2, 1] \subseteq E^* \)

up to a set of \( \theta \)’s of Lebesgue measure zero.

The intuition behind Remark 2 is simple to outline. The court’s decision to void the contract provides the parties with insurance against unforeseen
contingencies. Whenever $\theta \in [1/2, 1]$, the cost to the seller is $c_N$ with probability one. It is therefore optimal for the court to enhance the buyer’s incentives to undertake the relationship-specific investment by enforcing the ex ante contract.

We now turn to a further partial characterization of the court’s optimal decision rule. We are concerned with the “shape” of the court’s optimal decision rule for those $\theta$’s that are in $[0, 1/2]$. We first assert that this part of the court’s optimal decision rule consists of a threshold level $\theta^*$. The court will void the ex ante contract when $\theta < \theta^*$ is observed and will uphold the ex ante contract otherwise.

Remark 3. Assume that a solution to problem (11) exists. Then, up to a set of $\theta$’s of Lebesgue measure zero, any solution to this problem has the form $E^*=[\theta^*,1]$ with $\theta^* \in [0, 1/2]$. In other words, the court will enforce the ex ante contract if $\theta \geq \theta^*$ and will void it if $\theta < \theta^*$.

The intuition behind this second partial characterization of the optimal court decision rule can be described as follows: The court is trading off the insurance it provides to the parties when it voids the contract with the decrease in incentives to invest that results from voiding. Incentives are adversely affected because when the court voids, at the margin, the buyer will not receive a full return from his or her investment. Hence, the higher the probability that the court voids, the lower is its incentive to invest. This negative effect on investment depends only on the probability that the court will void the contract. On the other hand, the value of the insurance to the parties from voiding is greater when $\theta$ is smaller since, by assumption, the spread between $c_N$ and $c_H(\theta)$ becomes higher as $\theta$ becomes smaller. Hence, whatever decrease in incentives is accepted, the optimal thing for the court to do is to void for the smallest values of $\theta$. In other words, whatever the overall probability that the court voids the ex ante contract, the set of values of $\theta$ for which the contract is in fact voided must take the threshold form described in Remark 3.

We now have all the elements to complete the characterization of the court’s optimal decision rule. We do so in Proposition 2. Aside from incorporating the content of Remarks 2 and 3, Proposition 2 asserts that an optimal decision rule for the court does in fact exist, that it is unique up to a set of $\theta$’s of Lebesgue measure zero, and that the threshold $\theta^*$ used by the court is interior in the sense that $0 < \theta^* < 1/2$.

Proposition 2. An optimal decision rule for the court exists, and it is unique up to a set of $\theta$’s of Lebesgue measure zero.

The court’s unique optimal decision rule has the form $E^*=[\theta^*,1]$ with $\theta^* \in (0, 1/2)$. In other words, given $\theta$, the court upholds the contract when the parties face no risk and when the risk they face is sufficiently low ($\theta \geq \theta^*$). It voids the contract otherwise.

We have already outlined the intuition behind part of the characterization of the court’s optimal decision rule presented in Proposition 2. To understand
why the threshold $\theta^*$ used by the court cannot be either 0 or 1/2, it is enough to refer back to the specification of the risk that the unforeseen contingencies entail, described in Section 2. Recall that as $\theta$ approaches 1/2, the risk faced by the parties becomes negligible ($c_H(\theta)$ approaches $c_N$). Therefore, as $\theta$ approaches 1/2, the value of the insurance that voiding provides shrinks to zero. On the other hand, the costs of voiding the ex ante contract do not vanish. The marginal cost (in terms of diminished incentives for the buyer to undertake relationship-specific investment) of increasing $\theta^*$ does not become zero as this threshold gets closer to 1/2. Therefore, the optimal $\theta^*$ is below 1/2.

Consider now the nature of the risk associated with the unforeseen contingencies for small $\theta$, approaching 0. In this case, the difference between $c_N$ and $c_H(\theta)$ becomes unboundedly large. The gain in incentives from upholding the ex ante contract is bounded above (it can never exceed $R$), although upholding the ex ante contract becomes more and more costly as the parties are faced with an ever-increasing amount of uninsurable risk. Therefore, the optimal $\theta^*$ is above zero.

5. Frustration, Impracticability, and Optimal Courts

The analysis of the optimal court involves the trade-off between the protection afforded a risk-averse party when performance is excused in the face of unforeseen events and the consequent negative effect excuse has on optimal investment by the contracting parties. Many of the leading cases involving frustration of purpose seem to have no significant investment, hence no trade-off. The classic frustration case is *Krell v. Henry* (1903).18 The contract was to rent for 2 days an apartment overlooking the coronation route for the coronation of King Edward VII. The coronation was canceled due to the king’s illness, which was deemed a frustrating event, and the contract to rent the apartment was voided. It is difficult to see a significant investment by either party, and voiding such a contract would seem simply to entail transferring the risk associated with the cancellation of the coronation from one party to the other. Absent any particular reason to believe one party was inherently more risk averse than the other, there is little reason for voiding (or not) on efficiency grounds. It would seem that “fairness” rather than a concern for the efficiency of investment is at the heart of this case.

In *Lloyd v. Murphy* (1944),19 the court was again faced with a frustration case. The plaintiff leased land to the defendant for 5 years solely to sell cars and gasoline shortly before World War II. After the United States entered the war, the government ordered the sale of most new cars discontinued. The defendant repudiated the contract and left the premises, whereupon the plaintiff sued for unpaid rent. The court ruled that both parties knew that the war was coming and that the possibility that car sales would be curtailed was possible; furthermore, car sales were restricted but not completely eliminated. The fact that car sales were only “severely restricted” rather than eliminated would

---

play little role in our model. The issue is rather the risk that is faced by the intervening event—World War II. There is an important point at which this case deviates from our model. The costs voiding contracts in our model stem from the decrease in investments that will be made prior to fulfilling contractual obligations when contacts may be voided. In *Lloyd v. Murphy* (1944), it is likely that such costs were incurred by the lessee and not the lessor. Thus, the logic of our analysis would suggest that there is little cost of excusing performance on the part of the lessee.20

Our analysis of an optimal court does bear on *Transatlantic Financing Corp v. United States* (1966). The court ruled against Transatlantic, saying that the injured party cannot proceed with performance, recover the contracted price, and then recover its extra costs in addition. Whereas our analysis deals only with a court voiding or enforcing a contract prior to performance, one expects that the logic carries over to a case in which performance has begun prior to the intervening event, and it is clear that it is efficient to complete performance. We emphasize, however, that allowing courts to go beyond voiding or enforcing contracts by revising the terms of a contract is outside the scope of this article.21

In *Selland Pontiac-GMC v. King* (1986),22 the plaintiff entered into a contract with the defendant to supply four chasses for buses. The defendant was to get bodies from a third party, specified in the contract, that would be assembled on the chasses. The third party went out of business and the defendant could not get the bodies, following which the defendant tried to cancel the order for the chasses. The court held for the defendant saying that the supply of the bodies was a basic assumption of the contract. This case fits well within our model: the plaintiff likely incurred nontrivial costs between the time the contract is signed and the time that the defendant cancels. If courts excuse performance in similar cases, sellers will decrease the investments they make due to the risk that contracts may be voided.

### 6. Concluding Remarks

#### 6.1 Modeling decisions

We have taken a particularly simple specification of the court’s strategy set and of its preferences. We will discuss each of these and how it relates to our analysis above.

There is a sense in which any restrictions (except for strictly physical ones) on the court’s strategy set take us back into a partial equilibrium approach. If

20. If the sale of cars entailed a significant investment on the part of the lessee, there would be a nontrivial trade-off had it been the lessor who asked that the contract be voided.

21. *American Trading & Production v. Shell International Marine*, 453 F.2d 939 (2nd Cir. 1972) is similar to *Transatlantic Financing Corp. v. United States* (1966) in that American Trading sued Shell for extra compensation that resulted from the Suez closing. It differed in that the amounts were approximately double those in the *Transatlantic Financing Corp. v. United States* (1966) case, but the court’s decision was the same, namely, to deny the extra compensation.

there are restrictions on the court’s strategy set, who put them there if the model is truly a closed one? This article is but one step in the direction of a model that is truly closed in this sense.23

Once we take the view that some external considerations must be taken as given, it is easy to see why our modeling choice of a “simple” strategy set for the court is plausible. Courts typically face a large pool of possible disputes and have very little prior specific knowledge about each case. It is clearly efficient to develop court procedures that are “detail free” wherever possible in the sense of being robust to even large variations in the parameters characterizing the situations to which they apply. Our courts that can only void or uphold contracts rather than dictate new terms of trade are a simple way to capture some of these considerations.

The restricted strategy space for the court that we have worked with in this article can also be interpreted as a crude way to model the effects of a richer domain for the preferences of the court. In particular, it is clear that in a dynamic world, courts must care about the reputation they accumulate about their rulings. In the static analysis above, precedents are assumed to be equivalent to the court announcing to the parties the rule that it will use in case of a dispute. In a richer dynamic model, this would be substituted by the reputation that the court has. At this point, the rationale for simple behavior becomes, again, apparent. In practice, simple rules will have greater “penetration” as the reputation of the court among the pool of (possibly simple minded) contracting parties who might take their disputes before the court.

6.2 Are courts necessary?

One might ask whether courts are necessary to insure against unforeseen contingencies. Should it not be possible for the parties to specify within the contract the nature of the events in which performance is to be excused? It is possible, and in fact common, for parties to specify within a contract that performance is to be excused in particular circumstances, for example, a force majeure clause. Such clauses typically excuse one or both parties from their obligations in the event of war, natural disaster, or some other event outside their control. An example of such a clause found on the Web is as follows:

Neither party shall be liable in damages or have the right to terminate this Agreement for any delay or default in performing hereunder if such delay or default is caused by conditions beyond its control including, but not limited to Acts of God, Government restrictions (including the denial or cancellation of any export or other necessary license), wars,

23. In a different context—the design of a “legal system” for society as a whole—Mailath et al. (2000) explore a model in which all “laws” are cheap talk. They find that the role of the legal system in this case is limited to selecting among the multiple equilibria of the game determined by the physical description of the environment. See also the discussion in Schwartz and Watson (2004).
insurrections and/or any other cause beyond the reasonable control of the party whose performance is affected.24

Such a clause would not eliminate the need for an analysis as in this article. A half inch of snow is certainly outside the parties’ control, and it might slow down slightly a truck making a contractually agreed on delivery, but a court would rule against a party seeking to be excused from performing on account of the snow. Unless courts implicitly or explicitly set a threshold for excusing performance, any party that would like a contract voided can always find some event that may technically fall within the force majeure clause. A court then must determine whether a contractual dispute actually does fall within a force majeure clause.

Our court increases efficiency by excusing performance in circumstances that are deemed to have been unforeseen and, further, that expose at least one of the parties to risk of substantial magnitude. The court, of course, is acting after all uncertainty has been resolved (even though the court does not fully know the realization). Hence, excusing performance at that date is simply a transfer from one party to the second. The increases in efficiency that stem from voiding a contract are a consequence of superior risk sharing between future contracting parties. It follows, then, that it is not simply the voiding or enforcement of a contract that determines the efficiency gains but the expectations induced in future contracting parties due to the court’s decision. If the court determines that performance is to be excused, there is still substantial scope for the court to affect expectations through its written decision. Future parties’ expectations will be quite different following narrowly written decisions than following broadly written ones.

It is worth pointing out that contracting parties will often have information ex ante that courts will not have, even ex post, including the likelihood of certain kinds of events, the risk aversion of the parties, and the importance of ex ante investments. The optimal court decision rule will be different for different contracting pairs. The existence of different venues for adjudicating disputes can allow for different levels of insurance by applying different thresholds for excusing performance. Thus, the selection of, say, New York or Delaware law to govern a contract may be about the level of insurance that is optimal for a contracting pair as about expertise in commercial law.

Appendix

Proof of Proposition 1. Consider the first-order conditions associated with problem (7). After elementary manipulations, we obtain that the following must hold:

\[
\int \theta E \left[ q V' \left( p^* + t^*_\theta - cH(\theta) \right) + (1 - q) V' \left( p^* + t^*_N - cN \right) \right] d\theta = \theta \theta V' \left( \hat{\theta} R + \Delta + t^*_\theta \right),
\]

(A1)

which of course proves equation (8). Since $V'' < 0$ and $c_H(\theta) \geq c_N$ for every $\theta$,
\[
q V'(p_\theta^* + t_\theta^* - c_H(\theta)) + (1 - q) V'(p_\theta^* + t_\theta^* - c_N) 
\geq V'(p_\theta^* + t_\theta^* - c_N).
\] (A2)

Substituting equation (A2) into equation (A1) yields
\[
V'(\hat{\epsilon}R + \Delta + t_\theta^*) \geq V'(p_\theta^* + t_\theta^* - c_N),
\]
which together with the fact that $V'' < 0$ implies equation (9).

The fact that equation (10) holds follows from the fact that the seller’s expected utility $V_\epsilon(p, t, \hat{\epsilon})$ is monotonic increasing in $t$, whereas the buyer’s expected surplus $B_\epsilon(p, t, \hat{\epsilon})$ is monotonic decreasing in $t$.

\textit{Lemma A1.} Let $E^*$ be any solution to problem (11), with associated $p^*, t^*$, and $\hat{\epsilon}$. Then up to a set of Lebesgue measure zero, $E^*$ must have the following property.

Let $\hat{\epsilon}$ be any point in $[0, 1]$. Then if the quantity
\[
(\hat{\epsilon}R + \Delta + c_N - p^*) + qg(\hat{\theta}) + (1 - \theta) \frac{R^2}{\Psi''(\hat{\theta})} + \lambda[q V(p^* + t^* - c_N - g(\hat{\theta})) + (1 - q) V(p^* + t^* - c_N) - V(\hat{\epsilon}R + \Delta + t^*)]
\] (A3)
is strictly less than zero, it must be that $\hat{\epsilon} \not\in E^*$. Conversely, if the quantity in equation (A3) is strictly greater than zero, then it must be that $\hat{\epsilon} \in E^*$.

\textit{Proof.} Consider the total change, as a function of $\gamma$, in the Lagrangean of problem (7) when we subtract from $E^*$ the arbitrarily small interval $[\hat{\theta}, \hat{\theta} + \gamma]$. After some manipulations, at $\gamma = 0$, the total marginal change in the Lagrangean can be seen to equal $-1$ times the quantity in equation (A3).

Therefore, if the quantity in equation (A3) is negative, the value of the Lagrangean can be increased by subtracting from $E^*$ the interval $[\hat{\theta}, \hat{\theta} + \gamma]$, for $\gamma$ appropriately small. This contradicts the fact that $E^*$ is the solution to problem (11). Clearly, this proves our first claim.

The proof of our second claim involves a completely symmetric argument, and the details are omitted.

\textit{Lemma A2.} Let any $E$ be given, and assume it is such that $E \cap [0, 1/2]$ has positive Lebesgue measure. Then the quantity in equation (A3) is strictly increasing in $\hat{\theta}$ for all $\hat{\theta} \in [0, \hat{\theta})$ with $< \hat{\theta} < \hat{\theta} < 1/2$. It is strictly decreasing in $\hat{\theta}$ for all $\hat{\theta} \in [\hat{\theta}, 1/2)$, and it is constant over the interval $\hat{\theta} \in [1/2, 1]$.

\textit{Proof.} Differentiating equation (A3) with respect to $\hat{\theta}$ and using the first-order conditions of problem (7) yields
\[
g'(\hat{\theta}) \left[ 1 - \frac{V'(p^* + t^* + c_N - g(\hat{\theta}))}{V'(\hat{\epsilon}R + \Delta + t^*)} \right].
\] (A4)
Our claim is now easily verified using (A4) if we let \( \tilde{\theta} \) be such that 
\[
p^* - c_N - g(\tilde{\theta}) = \hat{e}R + \Delta
\]
and recalling that \( g'(\tilde{\theta}) \) is negative over \([0, 1/2]\) and zero otherwise.

**Lemma A3.** Let \( \mathcal{E}^* \) be any solution to problem (11), with associated \( p^*, t^*, \) and \( \hat{e}. \) Then the value of the quantity in equation (A3) is strictly greater than zero for every \( \hat{\theta} \in [1/2, 1]. \)

**Proof.** Assume by contradiction that this quantity is nonpositive. Then using Lemma A2, it must be that, without loss of generality, either \( \mathcal{E}^* = \emptyset \) or \( \mathcal{E}^* \subseteq [0, 1/2]. \) This first possibility is ruled out by Remark 1, so our contradiction hypothesis is \( \mathcal{E}^* \subseteq [0, 1/2]. \)

Now consider an alternative enforcement set \( \mathcal{E}' \) with \( \theta_{\mathcal{E}'} = \theta_{\mathcal{E}}^* \) and \( \mathcal{E}' \subseteq [1/2, 1]. \) Given \( \mathcal{E}' \), the solution to problem (7), \( p_{\mathcal{E}'}^*, t_{\mathcal{E}'}^*, \hat{e}_{\mathcal{E}'} \), is easily seen to have the following properties. First of all, \( \hat{e}_{\mathcal{E}'} = \hat{e}. \) Moreover, \( p_{\mathcal{E}'}^* = \hat{e}R + \Delta + c_N \) and \( t_{\mathcal{E}'} = -\hat{e}R. \)

Therefore, the buyer’s payoff in the solution to problem (7) given \( \mathcal{E}' \) is equal to
\[
\hat{e}R - \psi(\hat{e}). \tag{A5}
\]
After elementary manipulations, the payoff to the buyer in the solution to problem (7) given \( \mathcal{E}^* \) can be written as
\[
\hat{e}R - \psi(\hat{e}) + \Delta + (1 - \theta_{\mathcal{E}})(\hat{e}R + \Delta + t^*)
- \int_{\mathcal{E}} [g(p^* + t^* - c_N - g(\theta)) + (1 - q)(p^* + t^* - c_N)]d\theta. \tag{A6}
\]
Using equation (10) and the concavity of \( V, \) it is immediate to show that the quantity in equation (A6) is strictly smaller than the payoff in equation (A5). Since this contradicts the fact that \( \mathcal{E}^* \) is a solution to problem (11), it suffices to prove our claim.

**Lemma A4.** The quantity in equation (A3) becomes negative as \( \hat{\theta} \) approaches zero.

**Proof.** From equation (A3), it is sufficient to show that the quantity
\[
g(\hat{\theta}) + \lambda V(p^* + t^* - c_N - g(\hat{\theta})) \tag{A7}
\]
diverges to \(-\infty\) as \( \hat{\theta} \) approaches zero. This can easily be verified dividing through by \( g(\hat{\theta}), \) using l’Hôpital’s rule, and recalling that \( V' \) is decreasing and that, by assumption, \( \lim_{\hat{\theta} \to 0} g(\hat{\theta}) = \infty. \) The details are omitted.

**Proof of Remark 2.** The claim is a direct consequence of Lemmas A1 and A3.

**Proof of Remark 3.** The claim is a direct consequence of Lemmas A1, A2, and A3.
Proof of Proposition 2. The claim is a direct consequence of Lemmas A1, A2, A3, and A4.

References