Legal Evolution, Contract Evolution and Standardization

Nicola Gennaioli
Enrico Perotti
Giacomo Ponzetto

December 2013
This version July 2014

Barcelona GSE Working Paper Series

Working Paper n° 747
Legal Evolution, Contract Evolution and Standardization*

Nicola Gennaioli          Enrico Perotti
U. Bocconi and IGIER      U. Amsterdam

Giacomo A. M. Ponzetto
CREI, U. Pompeu Fabra, and Barcelona GSE

July 2014

Abstract

In a model where biased judges can distort contract enforcement, we uncover positive feedback effects between the use of innovative contracts and legal evolution that improve verifiability and contracting over time. We find, however, that the cost of judicial bias also grows over time because the unpredictable application of precedents becomes costlier as the law matures. Contract standardization avoids this cost, statically improving enforcement; but it crowds out innovative contracts, hindering legal evolution. We shed light on the large-scale commercial codification undertaken in the nineteenth century in many common-law countries during a period of booming long-distance trade.

Keywords: Contracts, Imperfect enforcement, Legal evolution, Precedents, Standardization

JEL codes: D86, K12, K40, K41

*We are grateful to John Armour, Giuseppe Dari Mattiacci, Michael Fishman, Ronald Gilson, Raj Iyer, Canice Prendergast, Josh Schwartzstein, Andrei Shleifer, Dimitri Vayanos, Luigi Zingales, and seminar participants at Chicago Booth, Columbia Law School, Northwestern and Oxford for their useful comments. Chantal Mak, Stefano Mosso and Jacopo Ponticelli provided able research assistance. We are indebted to Katharina Pistor for extensive discussions. Gennaioli acknowledges financial support from the European Research Council and the Barcelona GSE. Ponzetto acknowledges financial support from the Spanish Ministry of Science and Innovation (grants ECO-2011-25624 and Juan de la Cierva JCI-2010-08414), the Spanish Ministry of Economy and Competitiveness through the Severo Ochoa Programme for Centres of Excellence in R&D (SEV-2011-0075), the Barcelona GSE Research Network and the Generalitat de Catalunya (2009 SGR 1157). E-mail: nicola.gennaioli@unibocconi.it, e.c.perotti@uva.nl, gponzetto@crei.cat.
1 Introduction

Market transactions can achieve efficient outcomes under the conditions of the Coase theorem (Coase 1960). These conditions include the reliable enforcement of contingent contracts (Grossman and Hart 1986). If instead courts are biased (Gennaioli 2013) or costly to use (Townsend 1970), contract flexibility needs to be sacrificed. Efficient arrangements may then become unattainable. Empirical evidence bears out the economic importance of courts. Legal rules and practices associated with common law positively predict the development of financial, labor and other markets (La Porta, Lopez-de-Silanes and Shleifer 2008).

These findings have revived the long-standing view that common law promotes efficient private contracting thanks to the efficiency-oriented nature of its judges and to the adaptability of precedents (Hayek 1960; Posner [1973] 2010). Evidence on court behavior, however, challenges this optimistic view. Adjudication is often swayed by judges’ preferences and biases (Frank 1930, 1949; Posner 2005), and judicial decisions have a degree of unpredictability even in areas of the law with mature precedents (Stone 1946, 1964; Niblett, Posner, and Shleifer 2010). Calabresi (1982) argues that the standardization of contracts undertaken in common-law systems in the last century was driven precisely by the desire to reduce the costs of unpredictable adjudication and chaotic precedents.

These observations raise two key questions. How does legal evolution affect contracting in a world where judges can be biased and precedents are imperfectly binding? What are the effects of different enforcement systems? And in particular, what is the role of contract standardization from both a static and a dynamic perspective? To answer these questions, we introduce imperfect judicial enforcement and legal evolution in an otherwise standard contracting setup. The buyer and the seller of a widget engage in a transaction in which quality-contingent pay is needed to induce productive effort. Verification of performance depends on how judges interpret the evidence presented by litigants in court. We make two assumptions concerning judicial verification.

First, following legal realism (Frank 1930; Posner 2005) and previous work (Gennaioli and Shleifer 2007; Gennaioli 2013), we assume that some judges are subject to biases. These biases may be idiosyncratic, reflecting for instance judicial attitudes towards specific litigants; or they may be systematic, such as when the transaction is international and local judges disfavor a buyer coming from a different jurisdiction. Either way, bias distorts judicial verification of quality. Contracting parties rationally expect that pro-seller judges will tend to over-estimate the quality of the widget while pro-buyer judges will tend to under-estimate it, and they write their contract accordingly.
Second, we assume that by contracting on legal precedents parties can limit judicial discretion, but only to some extent. Based on precedents, the contract may mandate, e.g., that an incentive payment should not be paid if the seller fails to deliver the widget on time. A pro-seller judge, though, may override this clause by arguing that the current transaction materially differs from precedent.\footnote{This is the canonical strategy of “spurious distinguishing,” whereby judges exploit spurious material aspects of the case, or purely abstract legal categories, to argue that the parties’ contract refers to a precedent different from the one parties originally intended (Fernandez and Ponzetto 2012). Many legal distinctions are at best dubiously grounded in objective efficiency considerations. E.g., manufacturers’ strict liability to the consumer was gradually introduced through a series of rulings that created exceptions for specific products such as soap, hair dye, dog food, and fish food. Cf. Kruper v. Procter & Gamble Co., 113 N.E.2d 605 (Ohio App. 1953); Graham v. Bottenfield’s, Inc., 176 Kan. 68, 269 P.2d 413 (1954); McAfee v. Cargill, Inc., 121 F. Supp. 5 (S.D. Cal. 1954); Midwest Game Co. v. M.F.A. Milling Co., 320 S.W.2d 547 (Mo. 1959).} This risk is higher if contract language is ambiguous and if the law attaches little binding power to precedents. In these cases, judges can either choose which exact precedent applies to the case at hand (Llewellyn 1930; Stone 1946, 1964), or alter the application of a given precedent to the current contract (Niblett, Posner, and Shleifer 2010).

Based on these assumptions, we study optimal contracts by adopting a mechanism-design approach. The optimal contract must provide incentives not only for the seller to produce high quality, but also for litigants to report their information truthfully and for judges to verify it at the enforcement stage. We obtain three main results.

First, the optimal contract is both path dependent and “innovative:” it conditions the seller’s payment on evidence based on precedents, but also on novel evidence that has not been used in court yet. Contracting on novel evidence and precedents is optimal because they both contain valuable information about quality. Judicial bias, however, reduces verifiability of quality by distorting the use of both sources of information. This distortion increases the cost of incentivizing the seller, destroying gains from trade. Stronger adherence to precedent restricts the discretion of biased judges, making quality more verifiable; but it cannot achieve the first best as long as novel evidence remains valuable.

Second, despite the presence of biased judges, efficiency increases over time through a virtuous cycle between the use of innovative contracts and legal evolution. Litigation of innovative contracts enriches the stock of precedents because judges are called to adjudicate upon new facts. As precedents become more informative, quality becomes more verifiable and parties are more willing to contract. Increased contacting speeds up legal evolution even more. Critically, though, we find that legal evolution does not eliminate the cost of judicial biases. On the contrary, as precedents become more informative it becomes ever more costly to have biased judges disregard them. Interventions removing distortions of
precedent become imperative when contracts and contract law mature.

This insight motivates our analysis of contract standardization, which yields our third result: this reform entails a static benefit but a dynamic cost. Consistent with real world practices, we model standardization as the creation of contracts that perfectly enforce established norms but do not incorporate novel evidence, because the standard-setter lacks the parties’ transaction-specific information. Standardization then has two static effects. It boosts the volume of trade among parties who would not use the innovative contract, but it crowds out the use of innovative contracts by some other parties. Owing to these effects, standardization immediately improves welfare by boosting the volume and efficiency of trade. On the other hand, it stifles legal evolution because judges are asked to evaluate the same features of every case and precedents no longer evolve. A key implication is that, although eventually it is always optimal to standardize, society should wait to do so after substantial legal evolution has taken place.

Our results help explain in a single framework both the evolution of contracts and the historical emergence of contract codification in common- but also civil-law systems. Choi, Gulati and Posner (2013) document that parties do not write contracts ex nihilo, as standard contract theory would predict, but rather introduce marginal innovations into slowly evolving established contract forms. In our model, relying on established contracts optimally reduces the enforcement frictions associated with novel contract terms. Then, as jurisprudence consolidates the enforcement of a certain novel contract clause, distortions in its enforcement decline, and the clause becomes a reference for all future transactions.

As shown by our model, this desirable process may actually increase the cost of judicial biases because judicial interference with highly developed contract practices is very costly. As we discuss in Section 6, this idea can shed light on the large-scale standardization that has occurred since the late nineteenth century across many different legal systems. This was a time when booming trade and industrialization created vast trading opportunities, but the different geographic and social backgrounds of trading partners created room for substantial enforcement risk. Consistent with our analysis, Calabresi (1982) views this process as a response to the need to make litigation cheaper and more predictable and he also warns that excessive standardization may stifle private initiative and innovation.

Our analysis contributes to the literature on legal evolution. Relative to early papers (Priest 1977; Rubin 1977), recent models of judge-made law focus on judicial behavior (Gennaioli and Shleifer 2007; Ponzetto and Fernandez 2008; Niblett 2013b; Anderlini, Felli, and Riboni 2014). Our view of legal evolution as a mechanism “completing the contract
space” is closest to Gennaioli and Shleifer’s (2007) model of distinguishing and to Hadfield’s (2011) portrayal of precedents as a form of judicial training. These papers, however, focus on torts and not on contracts, so they cannot account for slowly evolving contractual innovation. Anderlini, Felli and Riboni (2014) perform an analysis of contracting and precedents in the presence of time-inconsistent judges, but they consider neither contract innovation nor standardization.

Our approach to contract standardization is distinct from the legal literature on boilerplate and standard contracts. Ahdieh (2006) views standardization as a way to foster coordination. Kahan and Klausner (1997) view it as a way to exploit network effects and save on transaction costs. These approaches adopt the conventional industrial-organization perspective according to which standardization acts as a coordination device, particularly in sectors characterized by large network externalities (e.g., Varian, Farrell, and Shapiro 2004). In our model there is no benefit from coordination. Standardization acts as a constraint on judges who enforce contracts.

The paper is organized as follows. Sections 2 and 3 lay out the contracting setup and enforcement risk. Sections 4 and 5 study the optimal contract and legal evolution when partnerships are homogeneous. Section 6 studies standardization. Section 7 extends the model to the case in which partnerships are heterogeneous. Section 8 discusses real-world standardization episodes in light of our model. Section 9 concludes. Proofs are in the Appendix.

2 The Model

2.1 Setup

Time is discrete, with an infinite horizon. At the beginning of each period $t = 0, 1, \ldots$ a penniless entrepreneur (the seller) and a wealthy customer (the buyer) meet and choose whether to form a partnership involving the supply of a relationship-specific widget. During each period $t$, production occurs in two stages. First, the seller exerts effort $a \in [0, 1]$ at a non-pecuniary cost $C(a)$. Second, with probability $a$ the widget is realized to be “good,” taking value $v > 0$; with probability $1 - a$, the widget is “bad,” taking value zero. We keep time subscripts implicit until Section 3. The widget is an experience good, such as a medical or professional service, whose value is learned by the buyer only after consuming it. We impose the following restrictions on the seller’s cost function:

$$C(a) > 0, C'(a) > 0, C''(a) > 0, \text{ and } C'''(a) \geq 0 \text{ for all } a \in (0, 1),$$

(1)
with limit conditions $C(0) = 0$, $\lim_{a \to 0} C'(a) = 0$, and $\lim_{a \to 1} C'(a) > v$.

If at time $t$ the partnership is formed, the seller’s first-best effort level is:

$$a_{FB} = \arg \max_a \{av - C(a)\} = C^{-1}(v) \in (0, 1),$$

which corresponds to joint surplus

$$\Pi_{FB} = \max_a \{av - C(a)\} = vC^{-1}(v) - C\left(C^{-1}(v)\right) > 0.$$

If the partnership is not formed, the seller obtains $0$, while the buyer obtains utility $u_B \geq 0$. Forming the partnership is first-best efficient if and only if $\Pi_{FB} \geq u_B$.

### 2.2 Contracting

In period $t$, the seller and the buyer meet. If they form a partnership, the seller makes a take-it-or-leave-it contract offer to the buyer (so the seller has full bargaining power). Next, the seller exerts effort, which determines the likelihood of producing a valuable widget. The widget is produced, the buyer consumes it, and the contract is enforced.

Under full observability, the first best is implemented by requiring the buyer to pay the seller a price $p = a_{FB}v - u_B$ if he exerted effort $a_{FB}$, and zero otherwise. Unfortunately, effort is unobservable (and non-contractible), so this solution does not work.

If widget quality is observable and perfectly verifiable after consumption, the parties can specify a quality contingent price $p_q$, where $q \in \{0, v\}$. In the optimal contract, the buyer’s participation constraint is binding. Otherwise, the seller could raise $p_q$ for all $q$ and still ensure participation without affecting effort provision. As a result, the seller chooses $p_v \geq 0$ and $p_0 \geq 0$ to maximize joint surplus $av - C(a) - u_B$ subject to the buyer’s binding participation constraint $a(v - p_v) - (1 - a)p_0 = u_B$, and to the seller’s incentive-compatibility constraint $p_v - p_0 = C'(a)$. The problem can be rewritten as:

$$\max_{a \in [0, 1]} \{av - C(a) - u_B\}$$

subject to

$$av - u_B \geq \min_{p_v, p_0 \geq 0} \{ap_v + (1 - a)p_0\} \text{ s.t. } p_v - p_0 = C'(a).$$

In (5), the optimal price $p_q$ minimizes the cost of inducing any effort $a$. This minimum cost defines the set of effort levels that can be implemented given the buyer’s participation constraint. The seller chooses the surplus-maximizing effort $a$ from this set.
Proposition 1 When quality $q$ is contractible, the optimal contract sets a positive price only when quality is high ($p_0 = 0$ and $p_v = C'(a_{SB}) > 0$).

The first best is attainable if and only if the buyer’s outside option is nil ($u_B = 0$). The partnership is formed if and only if the buyer’s outside option is sufficiently low:

$$u_B \leq \max_{a \in [0, 1]} \{ a [v - C'(a)] \} .$$

(6)

Second-best effort and joint surplus decrease with the buyer’s outside option ($\partial a_{SB} / \partial u_B < 0$ and $\partial \Pi_{SB} / \partial u_B < 0$ for all $u_B \in (0, \max_{a \in [0, 1]} \{ a [v - C'(a)] \})$).

When the buyer’s outside option is positive, the first best cannot be achieved because of the seller’s wealth constraint. Ideally, the seller would like to pay $u_B$ to the buyer and “purchase the firm” from him, which would elicit first-best effort. However, this arrangement is infeasible because the seller is penniless. Second-best effort and joint surplus are below the first best and decrease with the buyer’s outside option. We assume that Equation (6) always holds, so the partnership is feasible when quality is fully contractible.

The optimal contract specifies zero payment to the seller in case of low quality, namely when $p_0 = 0$. This feature minimizes wasteful payments, thereby reducing the cost of incentive provision. A similar property will be at work when, owing to imperfect enforcement, parties can only contract on an imperfect signal of quality.

3 Litigation and Imperfect Verifiability

Quality is verified in court, depending on the evidence presented by parties and verified by judges. We now describe the structure of the evidence, as well as the preferences of judges.$^3$ We refer to the partnership occurring at time $t$ as “partnership $t$.”

3.1 General Evidence and Precedents

The quality of the seller’s job is reflected in a set of fundamental indicators, such as the functionality of the widget or its timely delivery, that are common to all partnerships and can be used in litigation. These indicators constitute a continuum $I$ of pieces of evidence, each of which is uniquely identified by an index $i \in [0, 1]$ and whose value $e_t(i) \in \{-1, 1\}$ depends on widget quality $q_t$. If quality at $t$ is high, all pieces of evidence take value 1:

$^3$In the appendix, we present a further microfoundation of our model of evidence collection. Here we provide a streamlined version.
formally \( e_t(i) = 1 \) for all \( i \). If quality is low, piece of evidence \( i \) takes value

\[
e_t(i) = \begin{cases} 
1 & \text{for } i < \xi_t \\
-1 & \text{for } i \geq \xi_t
\end{cases},
\]

where \( \xi_t \) is an i.i.d random variable with cumulative distribution function \( F_{\xi}(\cdot) \) and continuous density \( f_{\xi}(\cdot) > 0 \) on the interval \([0, 1]\). \( \xi_t \) captures the noisiness of evidence and thus the difficulty of proving the quality of the widget.

Negative material evidence \( e_t(i) = -1 \) is consistent with low quality only. Positive evidence \( e_t(i) = 1 \) signals high quality because it is more likely to occur when \( q_t = v \). Evidence carrying a higher index \( i \) is more informative. In the limit, \( e_t(1) \) almost surely takes values 1 if quality is high and \(-1\) if quality is low. In general, a piece of evidence \( e_t(i) \) is a sufficient statistic for \( q_t \) given a lower-indexed piece of evidence \( e_t(j) \) for all \( j \leq i \).

Thus, the index \( i \) measures the informativeness of a piece of evidence.

At any time \( t \), the set \( I \) is partitioned into a countable subset \( P_t \subset I \) of precedents and an uncountable subset \( I \setminus P_t \) of novel evidence. \( P_t \) contains all pieces of evidence \( i \) that have been used in past cases (before \( t \)) and cited in the judicial opinions justifying their outcome. The set \( I \setminus P_t \) contains all pieces of evidence that courts have not yet considered. Novel evidence and precedents differ with respect to their ease of collection and the predictability of enforcement.

Parties’ collection of novel evidence from \( I \setminus P_t \) is imperfect and undirected. During litigation, each litigant searches for one new piece of evidence. Search is imperfect because the litigant may come up empty-handed: a party is successful only with probability \( \ell \in (0, 1) \), and unsuccessful otherwise. Search is undirected because, if successful, the litigant samples a piece of evidence but it is not immediately apparent how informative such evidence is. Upon seeing novel evidence \( e_t(i_N) \), both litigants and the judge observe its realization in \( \{-1, 1\} \) but ignore its true informativeness \( i_N \in [0, 1] \). Owing to imperfect search, informative evidence is “hard”: litigants cannot falsify it, but they can hide it. If the litigant chooses to present novel evidence in court, the judge can likewise choose to hide it (e.g., by declaring it inadmissible or immaterial).

Precedents are more easily identified and enforced than novel evidence. Formally, the pieces of evidence \( e_t(i) \) for all \( i \in P_t \) are publicly observable in court. As a result, parties can contract on the realization of specific indices \( i \in P_t \). Let \( i_t \equiv \max i \in P_t \) denote the most informative precedent available at \( t \). If parties contract upon precedents, it is sufficient for them to contract on the realization of \( e_t(i_t) \) alone, which is a sufficient statistic for \( P_t \). The variable \( i_t \) summarizes the state of precedents at time \( t \). Precedent is harder to distort
than novel evidence, but judges have leeway to misinterpret contractual references to it. With probability $\alpha \in [0, 1]$ a judge can ignore the evidence corresponding to the precedent and interpret the relevant contract clause however he sees fit. With probability $1 - \alpha$, the judge must correctly verify the evidence and enforce the contract clause corresponding to precedent.

Our setup outlines a key distinction between precedents and novel evidence. The latter is subject to significant ambiguities because it does not have a track record of past use. Novel evidence is thus harder to interpret and easier to discard for both parties and judges. Precedents, by contrast, constitute the bedrock of judicial training. They incorporate information that is generated by higher courts and legal scholars, or that transpires from partnerships over time. Precedents thus allow contracting on more specific signals of performance. The fact that precedents are imperfectly binding is partly due to contractual ambiguity, which offers judges some interpretive leeway. But it also reflects, for a given quality of contract drafting, the stance toward stare decisis of the legal system. Precedents are less binding for contracts when adherence to stare decisis is weaker.

3.2 Transaction-Specific Evidence

In addition to the set $I$ of evidence that is informative about any partnership, we consider transaction-specific evidence that parties can contract upon. At each $t$ there is a partnership-specific set $U_t \equiv \{u^v_t, u^0_t\}$ of uninformative pieces of evidence: one is positive ($u^v_t = 1$) and the other is negative ($u^0_t = -1$), regardless of widget quality. Only parties recognize misleading evidence ex ante and can contract on $U_t$ if they wish to (just as they can contract on $P_t$). In particular, parties may want to rule out in their contract the use of evidence in $U_t$. Otherwise, judges would be unable to tell if a piece of evidence is informative about any partnership $t$ and thus sets a legal precedent for future cases becomes contractible at $t + 1$ (when $i_P \in P_{t+1}$). Two alternative mechanisms justify this assumption. On the one hand, at $t + 1$ judges may have learned the informativeness of signals in $P_{t+1}$. In line with legal realism, this may occur because the judge deciding a case cannot identify exactly and single-handedly its true ratio decidendi. Later courts, however, will determine the precise extent $i$ of the rule that a precedent had the power to establish (Cardozo 1921; Allen 1927; Radin 1933; Cohen 1935; Frank 1949; Montrose 1957; Llewellyn 1960; Dias 1985; Posner 1990; Garner 2009). A less demanding possibility is that $i$ becomes contractible at $t + 1$ because parties rather than judges learn the informativeness of precedents (owing to their personal experience, or communication with their peers) and they contractually reveal it to judges.

Formally, the informativeness $i_P$ of a novel piece of evidence $e_t (i_P)$ that decides a dispute for partnership $t$ and thus sets a legal precedent for future cases becomes contractible at $t + 1$ (when $i_P \in P_{t+1}$). Two alternative mechanisms justify this assumption. On the one hand, at $t + 1$ judges may have learned the informativeness of signals in $P_{t+1}$. In line with legal realism, this may occur because the judge deciding a case cannot identify exactly and single-handedly its true ratio decidendi. Later courts, however, will determine the precise extent $i$ of the rule that a precedent had the power to establish (Cardozo 1921; Allen 1927; Radin 1933; Cohen 1935; Frank 1949; Montrose 1957; Llewellyn 1960; Dias 1985; Posner 1990; Garner 2009). A less demanding possibility is that $i$ becomes contractible at $t + 1$ because parties rather than judges learn the informativeness of precedents (owing to their personal experience, or communication with their peers) and they contractually reveal it to judges.

This implies that, given the same quality of contract drafting, $\alpha$ is lower in common law regimes, which explicitly adhere to stare decisis. The weight accorded to precedent (and thus $\alpha$) is intermediate in Scandinavian legal system and those following the German tradition (Bankowski et al. 1997; Zweigert and Kötz 1998; Glendon, Gordon, and Carozza 1999). It is lowest (so $\alpha$ is highest) in a system of civil law following the French tradition, which formally denies the law-making role of precedent in theory, despite recognizing it more and more in practice (Lasser 1995; Zweigert and Kötz 1998).
mative or not (i.e., if it comes from \( U_t \) or \( I \setminus P_t \)). Novel evidence would then be turned into cheap talk, an outcome that parties want to avoid.

We think of uninformative evidence as factors that ex post look like valid performance proxies to any external observer, such as a judge, but that parties can identify ex ante as being irrelevant for performance in their particular partnership. Examples might include inappropriate measurements of quality (e.g., estimates of functionality relying on inappropriate methods) or idiosyncrasies in the seller’s production method (e.g., having an irregular work schedule). This transaction-specific evidence does not affect our main analysis. It only plays a role in our analysis of contract standardization in Section 6, for it implies that different parties may want to write contracts tailored to their needs, instead of being perfectly satisfied with a single one-size-fits-all standard contract. In particular, in light of their superior ex-ante knowledge of their transaction, parties can—and optimally will—rule out the use of such evidence in litigation (e.g., the contract may allow the seller to adjust his work schedule, or it may rule out certain misleading performance measures).\(^6\)

### 3.3 Judicial Preferences

Since the contract is litigated ex post, after production has taken place, the impact of contract enforcement is purely distributional. It is therefore natural for ex-post judicial decisions to be shaped by a judge’s idiosyncratic distributional preferences. There are three types of judges. A fraction \( \beta \) have a pro-buyer bias, and wish to minimize payment to the seller. A fraction \( \sigma \) have a pro-seller bias, and wish to maximize payment to the seller. The remaining \( 1 - \beta - \sigma \) judges are unbiased: they wish to enforce the contract faithfully. If judicial bias stems from personal idiosyncrasies, pro-seller and pro-buyer biases will roughly balance out in the population of judges and \( \beta \approx \sigma \). If \( \sigma >> \beta \), judicial bias is systematically pro-seller, if \( \beta >> \sigma \) judicial bias is systematically pro buyer. The difference between \( \beta \) and \( \sigma \) can capture the inequality of parties. When \(|\beta - \sigma|\) is large, one party expects a systematically more favorable treatment than the other in court. Our model allows us to consider idiosyncratic and systematic bias alike.

\(^6\)To sum up, in our model evidence has two dimensions: whether it is informative, and whether it is transaction-specific. We simplify the analysis by considering only evidence that is general and informative, or transaction-specific and uninformative. Thus, we neglect general uninformative evidence, and transaction-specific informative evidence. This restriction is reasonable. General uninformative evidence is irrelevant for the comparison between tailor-made and standard contracts. All private contracts rule it out identically, and the standard contract will be equally able to rule it out. Transaction-specific informative evidence simply increases the benefit for parties to write a tailor-made contract. Our assumptions maintain tractability while accounting for the fact that parties are aware of some idiosyncratic features of their transaction and may wish to contract on them in advance.
4 The Optimal Contract

A contract in our setting consists of a price schedule \( p (...) \geq 0 \) specifying a payment from the buyer to the seller contingent on the information presented by the litigants and verified by the judges. The contract cannot also specify a state-contingent trading rule because the widget is an experience good (such as a service). Thus, information is not generated or the good is not even produced until consumption by the buyer takes place. We rule out the possibility for the contract to specify punishments for the parties as a whole, such as non-pecuniary criminal penalties or incentive payments to judges. These punishments are illegal in the real world (and would not be robust to renegotiation).

Finally, the judge is the ultimate arbiter of the case. The evidence he chooses to verify is the sole determinant of contract enforcement. Payment cannot be made directly contingent on the information the parties have presented in court, bypassing the judge’s decision to verify it. Thus, litigants are powerless to work around judicial biases by impugning the judge’s ruling. This assumption captures a key feature of trial courts: judicial fact discretion (Frank 1949; Gennaioli and Shleifer 2008). The conflicting evidence produced by litigants before the court yields a menu from which the judge can select with substantial discretion in finding the facts of the case.\(^7\)

The contract is written at the beginning of the partnership, before any information on the value of the widget or the preferences of judges is revealed. The figure below summarizes the sequence of events within a generic period \( t \).

---

\(^7\)The appeal process determines no more than a marginal constraint. Appellate courts normally confine themselves to reviewing questions of law, and remain highly deferential to the findings of fact made by trial courts. In the U.S. federal court system, the standard of review for judge-found facts is that they will not be disturbed unless “clearly erroneous.” This is a very lenient standard, and findings of clear error by courts of appeal are very rare.
By the revelation principle, any contract \( p (...) \geq 0 \) can be represented by a two-stage direct revelation mechanism. First, both litigants are induced to report truthfully all their private information to the judge (stage 4). Parties are unable to communicate directly with the mechanism designer, bypassing the judge’s fact-finding discretion. Accordingly, the judge must be induced to report truthfully all his private information, including the litigants’ reports, to the mechanism designer (stage 5). The judge’s report represents judicial verification of the facts of the case. Contract enforcement depends directly on the judge’s report alone; it reflects the litigants’ reports to the judge only indirectly (stage 6).

Figure 1 illustrates the information possessed by different agents. After the buyer has consumed the widget, quality \( q_t \in \{0, v\} \) is observed by the litigants \( L \in \{B, S\} \) alone. Then, each litigant privately observes a novel piece of informative evidence \( e_t (i_t^L) = 0 \) denotes an unsuccessful search. Thus, in the first stage of the direct revelation mechanism (stage 4), each litigant reports to the judge quality \( q_L \) and the realization \( e_L \) of his private signal. The optimal contract obviously rules out the use of misleading evidence \( \{u_t^v, u_t^0\} \), so parties cannot report it.

The parties and the judge observe all remaining information: precedent-based evidence \( e_t (i_t) \in \{-1, 1\} \), the judge’s ability to disregard contractual references to precedent \( \omega_t \in \{0, 1\} \), and the judge’s preferences \( b_t \in \{b_B, u, b_S\} \), where \( b_L \) denotes a bias in favor of litigant \( L \) while \( u \) denotes an unbiased judge. The judge then “verifies” all these elements by reporting to the mechanism designer the litigants’ reports, the realization of precedent \( e_P \), and his own type \((b, \omega)\). As a result, the payment \( p (q_B, q_S; e_P, e_B, e_S; b, \omega) \) is enforced.

The optimal direct revelation mechanism is the contract that maximizes the seller’s expected payoff

\[
\max_{p(\cdot\cdot\cdot)} \left\{ a \mathbb{E} \left[ p (v, v; 1, e_B, e_S; b, \omega) | q_t = v \right] + (1 - a) \mathbb{E} \left[ p (0, 0; e_P, e_B, e_S; b, \omega) | q_t = 0 \right] - C (a) \right\}
\]

subject to the buyer’s participation constraint,

\[
a \{ v - \mathbb{E} \left[ p (v, v; 1, e_B, e_S; b, \omega) | q_t = v \right] \} - (1 - a) \mathbb{E} \left[ p (0, 0; e_P, e_B, e_S; b, \omega) | q_t = 0 \right] \geq u_B; \quad (9)
\]

the seller’s incentive compatibility constraint,

\[
C' (a) = \mathbb{E} \left[ p (v, v; 1, e_B, e_S; b, \omega) | q_t = v \right] - \mathbb{E} \left[ p (0, 0; e_P, e_B, e_S; b, \omega) | q_t = 0 \right]; \quad (10)
\]
and the seller’s wealth constraint,

\[ p(q_B, q_S; e_P, e_B, e_S; b, \omega) \geq 0. \] (11)

Relative to Proposition 1, this problem includes truth-telling constraints for the agents’ reports of their private information. Denote by \( \Omega^B_i \equiv \{ q_t = q, e_t(i_t) = e_P, e_t(i^B_t) = e_B, b_t = b, \omega_t = \omega \} \) the buyer’s information set when he reports \((q_B, e_B)\) to the judge, and by \( \Omega^S_i \equiv \{ q_t = q, e_t(i_t) = e_P, e_t(i^S_t) = e_S, b_t = b, \omega_t = \omega \} \) the corresponding set for the seller. Then, the buyer’s truth-telling constraints are

\[ \mathbb{E}[p(q, q; e_P, e_B, e_S; b, \omega) | \Omega^B_i] \leq \mathbb{E}[p(q'_B, q; e_P, e'_B, e_S; b, \omega) | \Omega^B_i] \] (12)

for any feasible report \( q'_B \in \{0, v\}, e'_B \in \{0, e_B\} \). The seller’s truth-telling constraints are

\[ \mathbb{E}[p(q, q; e_P, e_B, e_S; b, \omega) | \Omega^S_i] \geq \mathbb{E}[p(q, q'_S; e_P, e_B, e'_S; b, \omega) | \Omega^S_i] \] (13)

for any feasible report \( e'_S \in \{0, e_S\}, q'_S \in \{0, v\} \).

In the final stage, the truth-telling constraints for biased judges who must respect precedent are

\[ p(q_B, q_S; e_P, e_B, e_S; b_B, 0) \leq p(q'_B, q'_S; e_P, e'_B, e'_S; b'_B, 0) \leq p(q_B, q_S; e_P, e_B, e_S; b_S, 0) \] (14)

for any feasible ruling \( q'_B, q'_S \in \{0, v\}, e'_B \in \{0, e_B\}, e'_S \in \{0, e_S\} \) and \( b' \in \{b_B, 0, b_S\} \). The truth-telling constraints for pro-buyer judges who can disregard precedent are

\[ p(q_B, q_S; -1, e_B, e_S; b_B, 1) = p(q_B, q_S; 1, e_B, e_S; b_B, 1) = \min_{e_P \in \{-1, 1\}} p(q_B, q_S; e_P, e_B, e_S; b_B, 0); \] (15)

while those for pro-seller judges who can disregard precedent are

\[ p(q_B, q_S; -1, e_B, e_S; b_S, 1) = p(q_B, q_S; 1, e_B, e_S; b_S, 1) = \max_{e_P \in \{-1, 1\}} p(q_B, q_S; e_P, e_B, e_S; b_S, 0). \] (16)

Let us discuss the equations above, starting the with seller’s objective in (8). When widget quality is high \((q_t = v)\), an event that occurs with probability \(a\) equal to the seller’s effort, the litigants truthfully report it \((q_B = q_S = v)\) and the evidence based on precedent is positive \((e_P = 1)\) because no negative signals can be realized in this state.
Thus, the expected payment to the seller is $E[p(v; v; 1, e_B, e_S; b, \omega) | q_t = v]$, where the expectation is taken across realizations of the litigants’ evidence collection $(e_B, e_S \in \{0, 1\})$ and the judge’s type $(b, \omega)$. When widget quality is low $(q_t = 0)$, with complementary probability $1 - a$, the expected payment is $E[p(0, 0; e_P, e_B, e_S; b, \omega) | q_t = 0]$. In this case, the expectation is taken across realizations of all evidence, including evidence based upon precedent $(e_P \in \{-1, 1\}$ and $e_B, e_S \in \{-1, 0, 1\})$ as well as the judge’s type $(b, \omega)$.

The same expected payments affect the buyer’s participation constraint in equation (9) and the seller’s incentive compatibility constraint in equation (10), because effort and participation are chosen before any material information is revealed to the parties.

Consider now the truth-telling constraints. Constraint (12) means that the buyer cannot lower his expected payment either by misreporting quality (which is cheap talk) or by hiding his private piece of evidence. His expectation takes into account, through $\Omega^B_t$, knowledge of actual quality $(q_t = q)$, of evidence based on precedent $(e_t (i_t) = e_P)$, of his private search for additional evidence $(e_t (i_t^B) = e_B)$ and of the judge’s type $(b_t = b, \omega_t = \omega)$. The buyer does not know the outcome of the seller’s parallel search for private evidence, but he anticipates it rationally when making his report. Analogously, constraint (13) means that the seller cannot raise his expected payment by making an untruthful report conditional on his information $\Omega^S_t$.

The remaining truth-telling constraints concern judges, and do not involve expectations because judges move last, when the litigants have revealed to the judge all the information he had not directly observed. The first constraint in (14) means that pro-buyer judges cannot lower payment by untruthfully verifying the litigants’ reports $(q_B, q_S, e_B, e_S)$, or their preferences $b$. The second constraint in (14) means that, analogously, pro-seller judges cannot raise payment through such selective verification.

The remaining constraints deal with judges’ ability to disregard precedent. Constraint (15) means that a pro-buyer judge with the ability to ignore precedent can interpret the contract as if the precedent-based evidence had the realization that minimizes payment. Constraint (16) means that a pro-seller judges symmetrically uses this interpretive leeway to maximize payment. These last two constraints must hold with equality. Otherwise, biased judges without the ability to ignore precedent could improve their payoffs by misreporting their ability without actually needing to disregard precedents.

As in the analysis of Proposition 1, the contracting problem can be solved in two steps. In the first step, the seller minimizes the cost of implementing effort subject to incentive compatibility, non-negativity, and truth telling (equations (10) to (16)). In the second step, the seller chooses optimal effort. As we show in the Appendix, this is a linear programming
problem whose solution minimizes the ratio of expected payments when quality is low relative to when it is high:

$$\Lambda = \frac{\mathbb{E}[p(...)|q_t = 0]}{\mathbb{E}[p(...)|q_t = v]}.$$

(17)

As in Proposition 1, the optimal contract should minimize wasteful payments when quality is low and maximize incentive payments when quality is high. However, since now quality is not directly contractible and reports of quality are cheap talk, payments cannot be perfectly targeted to occur if and only if $q_t = v$. As a result, the second-best contract minimizes the cost of effort provision by loading payment onto verifiable signal realizations that are most indicative of high quality.

**Proposition 2** If the share of pro-buyer judges is sufficiently low that

$$\beta \leq 1 - \mathbb{E}\xi,$$

(18)

then the optimal contract for partnership $t$ stipulates that uninformative evidence $\{u^*_t, u^0_t\}$ is inadmissible, and that the buyer must pay the seller a price $p > 0$ if and only if the court verifies no evidence of low quality (either novel or based upon precedent) and it verifies novel evidence of high quality.

This optimal contract is the intuitive analogue of the full-verifiability contract of Proposition 1. The incentive payment $p$ is made only when evidence based on precedent is positive and additionally the seller presents a novel signal indicative of high quality, while the buyer fails to present evidence of low quality.$^8$

This contract is simultaneously path-dependent and open-ended. Parties take an established contract form and add to it an incremental provision based on novel evidence. By building upon contracts litigated in the past, parties both exploit the information embodied in precedents and enhance the predictability of enforcement because precedents have some binding power on judges. By making payment contingent on novel evidence, they increase the information upon which payment is conditioned, reducing the cost of providing incentives to the seller. Motivated by the latter characteristic, we dub this optimal contract the “innovative contract.”$^9$

$^8$Under imperfect verifiability, the optimal contract cannot rely on direct revelation of quality $(q_B, q_S)$. Such revelation is cheap talk, and the litigants’ interests are perfectly opposed because outcomes in which both litigants are punished are impossible. The contract relies instead, as far as possible, on evidence based on precedent, requiring that it should not prove low quality $(e_i(i_t) = 1)$.

$^9$As we discuss in greater detail below and prove in the Appendix, the optimal contract would retain these features even if condition 18 did not hold.
The downside of writing an innovative contract is that novel evidence can be distorted by partisan lawyers and biased judges. Parties can protect themselves from these distortions by fine tuning the price $p$, so that contracting on novel evidence is always optimal. The problem, however, is that when enforcement distortions are strong, the benefit of writing an innovative contract is small.

To see this, note that in our model, imperfect verifiability of quality arises endogenously from litigants’ conflicting interests, judges’ biases, and the workings of precedent. Under the optimal contract, payment of $p$ is enforced in two cases. First, when the judge is unbiased ($b_t = u$), evidence based on precedent is positive ($e_t (i_t^B) = 1$), the buyer does not present negative evidence ($e_t (i_t^B) \in \{0, 1\}$), and the seller presents positive evidence ($e_t (i_t^S) = 1$). Second, payment is also enforced if the judge is pro-seller ($b_t = b_S$), the seller presents positive evidence ($e_t (i_t^S) = 1$), and evidence based on precedent is positive ($e_t (i_t^S) = 1$) or can be disregarded ($\omega_t = 1$).

These cases allow both for the possibility that the bonus is paid even if quality is low (e.g., the judge is pro-seller and disregards negative evidence from the buyer), and for the possibility that the bonus is not paid despite high quality (the judge is pro buyer and disregards positive evidence from the seller). Because under Condition (18)—which we assume throughout—pro-buyer judges are few, the loss from the second event is small enough that parties optimally load payments onto novel signals of high quality.\(^{10}\)

In spite of optimal contracting, partisan litigants and judges reduce verifiability. The incompleteness of the innovative contract is fully characterized by the minimized likelihood ratio $\Lambda$ of low relative to high quality when $p$ is enforced (equation [17]). Owing to the binary nature of the optimal contract, this ratio coincides with the likelihood ratio of low relative to high quality when $p$ is enforced: $\Lambda = \Pr(q = 0 | p) / \Pr(q = v | p)$. If the information that courts can verify is less diagnostic ($\Lambda$ is higher), the innovative contract is more incomplete. As we prove in the Appendix, incompleteness $\Lambda$ is a sufficient statistic for the role of judicial biases ($\beta$ and $\sigma$) and precedent ($i_t$ and $\alpha$) on efficiency. When the contract is too incomplete ($\Lambda > \hat{\Lambda}$), it cannot provide incentives for the seller while leaving the buyer a sufficient payoff: the partnership cannot be formed.\(^{11}\) Lower incompleteness en-

---

\(^{10}\)As we show in the Appendix, when $\beta$ is large the optimal contract does not condition payment on novel evidence of high quality collected by the seller, but relies exclusively on precedents and on novel evidence of low quality presented by the buyer. By discarding some information, this contract reduces the precision of state verification when the judge is unbiased or pro-seller. However, it improves state verification by pro-buyer judges, who are forced to rule in favor of the seller when they cannot disregard a positive precedent and the buyer presents no negative evidence. Condition (18) is only sufficient for the contract of Proposition 2 to be optimal. The Appendix derives the weaker necessary and sufficient condition: $\alpha \sigma = 0$ or $\beta \leq (1 - \epsilon \xi) / (1 - \alpha \epsilon \xi)$.

\(^{11}\)The threshold for partnership formation is defined by $\max_{a \in [0, 1]} \left\{ a v - \left[ \hat{\Lambda} / \left( 1 - \hat{\Lambda} \right) + a \right] C' (a) \right\} = \underline{a}_B$. 

16
ables sharper incentives that allow partnership formation and then monotonically increase equilibrium effort and joint surplus ($\partial a / \partial \Lambda < 0$ and $\partial \Pi / \partial \Lambda < 0$). The result below then highlights the role of judicial biases ($\beta$ and $\sigma$) and precedent ($i_t$ and $\alpha$) in our model.

**Proposition 3** Under the optimal innovative contract:

1. **Verifiability of quality is higher when there are fewer biased judges** ($\partial \Lambda / \partial \beta \geq 0$ and $\partial \Lambda / \partial \sigma \geq 0$) and when precedents are more informative ($\partial \Lambda / \partial i_t \leq 0$) and more binding ($\partial \Lambda / \partial \alpha \geq 0$). **Full contractibility is achieved if and only if precedent is perfectly diagnostic of quality and there are no pro-seller judges with the ability to disregard it ($i_t = 1$ and $\alpha \sigma = 0$).**

2. **Greater respect for precedent reduces the cost of judicial biases** ($\partial^2 \Lambda / (\partial \alpha \partial \beta) \geq 0$ and $\partial^2 \Lambda / (\partial \alpha \partial \sigma) \geq 0$). **More informative precedents reduce the cost of judicial biases if and only if informativeness is sufficiently low ($\partial^2 \Lambda / (\partial \beta \partial i_t) \leq 0$ and $\partial^2 \Lambda / (\partial \sigma \partial i_t) \leq 0$ if and only if $i_t < 1 - \alpha / \nu$).**

Point 1 illustrates the first-order effect of the enforcement parameters. Judicial bias reduces verifiability by destroying information (embodied in precedents, novel evidence, or both). This cost arises whether bias is idiosyncratic or systematic.\(^{12}\) On the other hand, precedents increase verifiability by generating more precise information (the more so the higher $i_t$) that judges cannot easily discard (with greater difficulty the lower $\alpha$).

Point 2 analyzes the subtle interactions between precedent and judicial biases. On the one hand, the more judges are constrained by precedent (the lower $\alpha$), the less they can indulge their biases, which improves verifiability ($\partial^2 \Lambda / (\partial \alpha \partial \beta) \geq 0$ and $\partial^2 \Lambda / (\partial \alpha \partial \sigma) \geq 0$). This result captures a traditional rationale for stare decisis. Binding precedent prevents each judge from having undue influence on legal outcomes, so that successive judicial decisions build a gradually evolving body of rules whose value is greater than the sum of its parts (Burke [1790] 1999; Fernandez and Ponzetto 2010).\(^{13}\)

---

\(^{12}\)This does not mean that all biases are equally costly in our model. We can prove that pro-seller bias is more costly than pro-buyer bias. This asymmetry arises because pro-buyer judges introduce white noise in the adjudication process (holding for the buyer regardless of the true state), while pro-seller judges introduce a systematic distortion. When quality is low, they undesirably make payment more likely. When quality is high, they cannot desirably increase the probability of payment because they cannot fake novel informative evidence that has not been collected by the seller.

\(^{13}\)Greater respect for precedent also enhances the value of the information embodied in existing precedents ($\partial^2 \Lambda / (\partial \alpha \partial i_t) \geq 0$). In other words, stare decisis is naturally more valuable the higher the quality of precedent.
On the other hand, having more informative precedents may increase the cost of judicial bias. This result reflects the sum of two opposite effects. When precedent is more informative, it is costlier for parties that biased judges may disregard it. At the same time, precedents and novel evidence are substitutes. When precedent is more informative, it is less important that the litigants can generate additional information \( (\partial^2 \Lambda / (\partial i \partial i_t)) \geq 0 \) and less costly that biased judges can discard it.\(^{14}\) If precedents are sufficiently informative \( (i_t > 1 - \alpha / \iota) \), the first effect dominates. That is, if the law is highly developed and informative, the cost of biases that induce judges to disregard such valuable information may be very large. This effect has important implications for legal evolution.\(^{15}\)

Through their impact on incompleteness \( \Lambda \), the enforcement parameters \( \beta, \sigma \) and \( \alpha \) and the informativeness of precedent \( i_t \) determine partnership formation and efficiency.

**Corollary 1** Partnership \( t \) is formed if and only if precedent is sufficiently informative: \( i_t \geq \iota \). Partnership formation is more difficult when there are more biased judges \( (\partial \iota / \partial \beta \geq 0 \text{ and } \partial \iota / \partial \sigma \geq 0) \) and easier when precedents are more binding \( (\partial i / \partial \alpha \geq 0) \).

Conditional on a partnership being formed, the seller’s effort and thus joint surplus are lower when there are more biased judges \( (\partial a / \partial \beta \leq 0 \text{ and } \partial a / \partial \sigma \leq 0) \), and higher when precedents are more binding \( (\partial a / \partial \alpha \geq 0) \) and more informative \( (\partial a / \partial i_t \geq 0) \).

Figure 2 represents graphically Corollary 1 by depicting the regions of the parameter space where the partnership dissolves \( (NC) \) and where it forms \( (Inn) \). On the vertical axis, \( \theta \) denotes a combination of parameters that increases with each determinant of enforcement frictions \( (\beta, \sigma \text{ and } \alpha) \).

\(^{14}\)The collection of novel information naturally increases verifiability \( (\partial \Lambda / \partial \iota \leq 0) \). Greater ability to generate new information unambiguously increases the cost of judicial biases that prevent its verification \( (\partial^2 \Lambda / (\partial \beta \partial \iota) \geq 0 \text{ and } \partial^2 \Lambda / (\partial \alpha \partial \iota) \geq 0) \).

\(^{15}\)If novel evidence is scarce relative to disregard for precedent \( (\alpha \geq \iota) \), the first effect always dominates and any improvement in precedents makes judicial biases more costly.
When precedents are insufficiently informative relative to enforcement frictions \( (i_t < \dot{i}) \), the quality of verifiable information is so poor that it becomes prohibitively costly to harness effort, and the parties prefer not to contract (we are in region \( NC \)). If instead precedent is sufficiently informative, partnership \( t \) is formed.\(^\text{16} \)

The probability that partnership \( t \) is formed and the seller’s effort increase with the quality of verifiable evidence (i.e., decrease in \( \Lambda \)). The higher verifiability, the cheaper it becomes for the contract to reward high quality and provide incentives to the seller. Therefore, equilibrium effort and efficiency increase. In line with Proposition 3, then, contracting expands and efficiency rises when judicial biases are less severe and when precedents are more binding and informative.

## 5 Legal Evolution

The optimality of innovative contracts under distorted enforcement has important implications for legal evolution. As parties write and litigate these contracts, the stock of precedents can evolve. Whenever courts rely on a novel piece of evidence in their decision, such evidence is included in the stock of precedents. If the novel piece of evidence is sufficiently informative, subsequent private contracts are updated to incorporate it.

We now study this mutual feedback between the evolution of precedents and contracts. For simplicity, we assume that litigation is costless, so parties always go to court to enforce

\(^{16}\)Partnership formation is also more difficult when the buyer’s outside option is higher \( (\partial y / \partial u_B \geq 0) \). In the figure, such an increase in \( u_B \) induces a downward shift in the locus \( i_t = \dot{i} \).
their contract.\textsuperscript{17} Every time judges adjudicate, they must provide a detailed opinion that justifies their decision. Under the innovative contract from Proposition 2, a judge may write four different decisions.

1. The seller wins by presenting positive evidence \((e_t (i_S^t) = 1)\), while no negative evidence was verified. This decision establishes a new precedent \((P_{t+1} = P_t \cup \{i_S^t\})\).

2. The buyer wins because precedent is negative \((e_t (i_t) = -1)\). This decision is based on existing precedent and thus never establishes a new one \((P_{t+1} = P_t)\).

3. The buyer wins by presenting negative evidence \((e_t (i_B^t) = -1)\). This decision establishes a new precedent \((P_{t+1} = P_t \cup \{i_B^t\})\).

4. The buyer wins because the seller failed to present positive evidence. This decision is based on absence of evidence and never establishes a new precedent \((P_{t+1} = P_t)\).

Changes in the set of precedents do not necessarily improve its informativeness. Precedent improves when, as in case 3, the buyer wins by presenting novel negative evidence against a positive current precedent. In this case, the structure of informative evidence guarantees that \(i_{t+1} = i_B^t > i_t\). If instead, as in case 1, the seller wins by presenting novel positive evidence on top of a positive current precedent, informativeness improves if and only if the seller happens to draw a new piece of evidence that is more informative than precedents \((i_S^t > i_t)\).\textsuperscript{18}

In our model, then, changes in the stock of precedents do not automatically translate into uniform changes in the contracts written by private parties. Contracts change only when the new evidence incorporated into precedent is sufficiently informative. Such dynamics of contract evolution provide an enforcement-based explanation for the findings of

\textsuperscript{17}This is a conventional assumption in many studies of legal evolution. If litigation were costly, both litigants could prefer to settle out of court. The standard justification for why parties then go to court is that they hold different priors about the probability of winning the trial. We abstract from modeling this feature because none of our results would depend on the specific states leading or not leading to litigation. Reluctance to litigate would simply slow down legal evolution.

\textsuperscript{18}In some cases, a ruling in the buyer’s favor could be justified in several ways. In the limit, suppose that both precedent and the buyer produced negative evidence \((e_t (i_t) = e_t (i_B^t) = -1)\) while the seller failed to produce positive evidence \((e_t (i_S^t) \neq 1)\): then each of the three decisions in the buyer’s favor is possible. We assume that judges choose which decision to render on the basis of two principles. First, in accordance with stare decisis, if evidence based on precedent suffices to settle the case, it is summarily decided without considering novel evidence. Second, due to the need to justify their decision, judges always prefer citing novel evidence than grounding their ruling on the insufficiency of available evidence. As a consequence, judges consider the four decisions in the order given above. They proceed down the line only if they cannot (or do not want nor have to) stop at a lower-numbered decision. This assumption does not qualitatively affect our results, but merely influences the speed of legal evolution. Precedents would evolve more rapidly if judges preferred decision 3 to decision 2, or more slowly if they preferred decision 4 to decision 3.
Choi, Gulati and Posner (2013). They document that parties write contracts by slightly modifying the terms of contracts used in the past or by other parties. Most of these marginal changes are transient. Sometimes, however, these changes prove effective and new contract forms emerge which then become the reference for future contracting parties.

In light of these observations, we consider the joint evolution of precedents and contracts. This legal evolution can start from scratch only if it is profitable to form a partnership when there is no prior history of contract enforcement. Formally, this occurs when $\bar{i} = 0$ so we are in the bottom region of Figure 2 (where contracting takes place) even at $i_0 = 0$. This condition, which we henceforth assume, holds provided that judicial biases $\beta$ and $\sigma$ are sufficiently rare.\(^{19}\)

**Proposition 4** When $\bar{i} = 0$, the evolution of precedent is described by a time-homogeneous Markov chain. Given any body of precedents $i_t$, any weakly higher informativeness $j \geq i_t$ is accessible, but any strictly lower informativeness $j < i_t$ is inaccessible. The Markov chain is absorbing: its unique absorbing state is perfectly informative precedent $i = 1$, while all imperfectly informative states $i \in [0, 1)$ are transient.

The use of innovative contracts by parties causes a monotonic evolution of contract law towards greater informativeness, because in our model precedents do not depreciate. The quality of precedent is described by a monotone increasing and ratcheting process. If informativeness $i_t$ has been attained at time $t$, then less informative states $j < i_t$ are unattainable in the future. Conversely, any higher level of informativeness can be reached from the initial state $i_t$. In fact it can be reached directly through a single ruling. For any threshold $j \in [i_t, 1)$, the appendix derives the expression for the strictly positive probability with which the judicial decision for partnership $t$ establishes a new precedent whose informativeness is greater than $j$: $\Pr (i_{t+1} > j | i_t) > 0$. The informativeness of precedents is unchanged with complementary probability $1 - \lim_{j \to 0} \Pr (i_{t+1} > j | i_t) > 0$. The unique absorbing state of this process is perfectly informative precedent ($i_t = 1$), and the stationary distribution of the Markov chain is fully concentrated on the absorbing state.

The model depicts a joint evolution of contracts and the law. On the one hand, the use of innovative contracts promotes legal evolution, because judges are asked to evaluate new features all the time. On the other hand, legal evolution allows parties to update contracts in light of more informative precedents, so that legal evolution steadily raises verifiability.

\(^{19}\)As we prove in the Appendix, the condition can be written out explicitly as $\hat{\lambda} \geq \mathbb{E} \xi - [1 - \sigma / (1 - \beta)] \mathbb{E} [\xi (1 - \xi)]$, where the left-hand side is the maximum contract incompleteness consistent with partnership formation, while the right-hand side is the incompleteness of the optimal contract in the absence of precedents (i.e., for $i_0 = 0$).
and efficiency towards the second best. At the same time, though, legal evolution is no panacea for judicial bias: as hinted by Proposition 3 the cost of judges’ biases conversely tends to increase over time.

**Corollary 2** As long as the ability to draft contracts and adherence to precedent are imperfect ($\alpha > 0$), full verifiability is never attained. If $\alpha \geq i$, the cost of judicial bias increases monotonically over time ($\partial^2 \Lambda / (\partial \beta \partial i_t) \geq 0$ and $\partial^2 \Lambda / (\partial \sigma \partial i_t) \geq 0$). If instead $\alpha < i$, the cost of judicial bias initially declines over time but eventually increases over time after precedent becomes sufficiently developed (for $i_t \geq 1 - \alpha / i$).

Judicial bias becomes more damaging to verifiability and efficiency over time. This result follows from Propositions 3 and 4. To illustrate it, Figure 3 below reports joint surplus $\Pi$ as a function of the informativeness of precedent $i_t$ when there is no judicial bias (dashed line, $\beta = \sigma = 0$) and when some judicial bias is present (solid line, $\beta + \sigma > 0$). The case considered is $\alpha < i$.

![Figure 3: Efficiency Consequences of Precedent and Judicial Bias](image)

When judges are unbiased, social welfare monotonically increases as the law develops (i.e., as $i_t$ goes up), and the second-best contractible-quality outcome is attained when the law converges to its absorbing state ($i_t \rightarrow 1$). When some judges are biased, it is also the case that social welfare monotonically increases as the law develops over time. However, the cost of judicial biases (the vertical distance between the dashed and the solid line) initially falls but after some point it increases over time and becomes more severe at mature stages of legal evolution. Intuitively, when precedent is very informative, it is very
costly that biased judges disregard it. As a result, judicial bias causes welfare to be always below the second best, even as the law converges.

This result indicates that the ability of precedents to solve the problem of judicial bias is incomplete. Hence, there is scope for a regulatory intervention that aims at eliminating residual ambiguity in the binding role of precedent. This is the role of standardization, which codifies existing legal and trading practices to make their enforcement more predictable. The next section analyzes the role of contract standardization in our model.

6 Standardization

Standardization can be undertaken by the public legal system via commercial codification, e.g., by specifying default investor rights (La Porta at al. 1998), or by a private trade association (Bernstein 2001). In either case, reliable off-the-shelf contracts are created. The goal is to make the enforcement of frequently used contracts more predictable. This aim is typically achieved by systematizing precedents and by extensively training judges and legal professionals in the enforcement of standard contracts. Individual contracting parties cannot likewise train the courts that will enforce their contracts. As a result, privately drafted innovative contracts are unavoidably subject to enforcement risk.

To capture this idea, we model the standard contract as having $\alpha = 0$. The shortcoming of standardization is that public authorities are less informed about the specifics of each partnerships than the parties themselves. It is here that the presence of transaction-specific information bites. Obtaining such information for all partnerships is too costly for standard-setting bodies, so the standard contract cannot identify ex ante the misleading evidence $\{u_t', u_t^0\}$.

After standardization occurs at time $s$, an unambiguous contract form is created. We allow the standard prevailing at any time $t \geq s$ to systematize the most up-to-date state of precedent $i_t$. Thus, we do not stack the cards against standardization by considering obsolescent and increasingly suboptimal standards. In line with the real world, as the law evolves the standard contract is also updated.

---

$^{20}$In reality, even standard contracts are subject to some judicial subversion (Niblett 2013a), but for simplicity we rule out this possibility. The only thing we need is that standard contracts constrain judicial interpretation of precedent more than nonstandard contracts.
6.1 The Optimal Standard Contract

The standard setter writes a contract that solves the mechanism design problem of Section 4 with the two changes described above. First, the standard contract is unambiguous: \( \alpha = 0 \). Second, if any novel evidence is admitted, litigants can direct their search toward misleading evidence \( U_t \) because the latter is unknown to the standard setter and therefore cannot be ruled out.

The former change eliminates some truth-telling constraints for judges. Constraints (15) and (16) no longer bind since all judges are fully bound by precedent \( (\omega_t = 0) \) when enforcing the standard contract. The latter change adds some truth-telling constraints for the litigants. Constraints (12) and (13) become

\[
\mathbb{E} \left[ p(q, q; e_P, e_B, e_S; b, \omega) \mid \Omega_t^{B} \right] \leq \mathbb{E} \left[ p(q', q; e_P, e'_B, e_S; b, \omega) \mid \Omega_t^{B} \right] \tag{19}
\]

for any feasible report \( q'_B \in \{0, v\} \) and \( e'_B \in \{-1, 0, 1\} \), and

\[
\mathbb{E} \left[ p(q, q; e_P, e_B, e_S; b, \omega) \mid \Omega_t^{S} \right] \geq \mathbb{E} \left[ p(q, q'; e_P, e_B, e'_S; b, \omega) \mid \Omega_t^{S} \right] \tag{20}
\]

for any feasible report \( q'_S \in \{0, v\} \) and \( e'_S \in \{-1, 0, 1\} \). Litigants’ reports of novel evidence become cheap talk under a standard contract because \( U_t \) is not ruled out.

The optimal standard contract is the following.

**Lemma 1** At time \( t \), the optimal standard contract stipulates that the buyer should pay the seller a price \( p_{\text{Std}} > 0 \) if and only if the standard evidence based on precedents indicates high quality \( (e_t(i_t) = 1) \).

The optimal standard contract relies only on information based upon precedents, which can be standardized, while it disallows the consideration of any novel evidence. This rigidity is necessary in light of the standard-setter’s lack of knowledge about partnership-specific uninformative evidence \( U_t \). If novel evidence were allowed, it would then never generate more information, because parties would strategically present uninformative evidence (i.e., the seller would always produce \( u_t^0 = 1 \) and the buyer \( u_t^0 = -1 \)).

Under the standard contract described by Lemma 1, the realization that triggers payment has likelihood ratio

\[
\Lambda_{\text{Std}}(i_t) = 1 - F_\xi(i_t), \tag{21}
\]

which is independent of judicial biases and decreasing in the quality of the evidence that can be standardized on the basis of precedent at time \( t \) \( (\Lambda_{\text{Std}}'(i_t) = -f_\xi(i_t) < 0) \). Intuitively, the quality of the standard contract improves with the informativeness of precedent.
In the limit, if standardized precedents become perfectly informative it allows for full contractibility of the value of the widget \( \lim_{i_t \to 1} \Lambda_{\text{Std}} (i_t) = 0 \).

### 6.2 Contract Choice and Welfare

If a partnership is formed, parties choose between the standard contract of Lemma 1 and the innovative contract of Proposition 2. Both contracts exploit the same set of precedents \( P_t \), and parties opt for the standard contract if it features a lower likelihood ratio than the innovative one \( \Lambda_{\text{Std}} < \Lambda \).

**Proposition 5** The parties prefer the standard contract to the innovative contract if and only if standardized precedents are sufficiently informative: \( i_t > i \in [0, 1] \). The standard contract is more likely to be preferred when judicial bias is more prevalent \( (\partial i / \partial \beta \leq 0 \text{ and } \partial i / \partial \sigma \leq 0) \), and when innovative contracts are more ambiguous or adherence to precedent is low \( (\partial i / \partial \alpha \leq 0) \).

The benefit of the standard contract is to protect parties against judicial bias. Its cost is to preclude the use of novel informative evidence. Parties prefer the standard when judicial bias is too strong, when judges often disregard precedent in innovative contracts and when precedent is more informative.

The last property reflects differential improvements in the two contract forms. Both standard and innovative contracts improve as \( i_t \) becomes higher. However, the improvement is starker for the standard contract, because better precedents reduce the usefulness of including novel evidence in the enforcement of the contract. As a result, the standard contract is preferred when precedent is sufficiently informative.

For a standard to be used, it must be preferred not only to the optimal innovative contract, but also to dissolving the partnership. This also requires precedent to be informative enough, namely \( i_t \geq i_{\text{Std}} = F_{\xi}^{-1} \left( 1 - \hat{\Lambda} \right) \). Subject to this condition, giving parties the option of using the standard contract can improve welfare in two ways.

---

21 The optimal innovative contract does not change when the standard contract is introduced. The presence of innovative terms based on novel informative evidence and on partnership-specific misleading evidence leaves the contract open to ambiguities concerning contractual references to precedent. As a result, the precedent continues to be binding for the innovative contract only with probability \( \alpha \).

22 In the limit, standardization is useless if there are no pro-seller judges with the ability to disregard references to precedents in an innovative contract \( (\partial i / \partial \alpha = 0) \). This is the enforcement friction that a standard contract can eliminate. Instead, standardization cannot prevent either litigants or biased judges from hiding unfavorable novel evidence. On the contrary, it must discard all such new evidence, including some that could usefully be exploited by a partnership-specific innovative contract.

23 The standard is also more likely to be preferred when novel evidence is harder to collect \( (\partial i / \partial \xi \geq 0) \).
Proposition 6 The space consisting of enforcement frictions ($\beta, \sigma$ and $\alpha$) and the informativeness of precedent ($i_t$) can be partitioned into four regions of non-zero measure.

1. If enforcement frictions are sufficiently high and the informativeness of precedent sufficiently low, such that $i_t < \min\{\tilde{i}, i_{Std}\}$, then partnership $t$ cannot be formed either through an innovative or a standard contract.

2. If enforcement frictions are sufficiently low and the informativeness of precedent is intermediate, such that $\tilde{i} \leq i_t \leq \bar{i}$, then partnership $t$ is formed through an innovative contract, whether a standard contract is available or not.

3. If enforcement frictions are intermediate and the informativeness of precedents is sufficiently high, such that $i_t \geq \max\{\tilde{i}, \bar{i}\}$, then partnership $t$ is formed through an innovative contract if a standard contract is not available, but under standardization it chooses the standard contract instead.

4. If enforcement frictions are sufficiently high and the informativeness of precedent is intermediate, such that $i_{Std} \leq i_t < \bar{i}$, then partnership $t$ can be formed only through a standard contract.

The key result is that contract standardization, by protecting parties against judicial bias, entails two consequences. It expands the volume of trade but it crowds out the use of innovative contracts. To show how these two effects work, Figure 4 below illustrates the effect of standardization on the outcome of Figure 2.

![Figure 4: Effects of Standardization](image-url)
First, when enforcement frictions are too pervasive, some partnerships fail to form through an innovative contract. This is region $NC$ in Figure 2, above the upward-sloping locus $i_t = \bar{i}$. In this case, provided that precedent is informative enough ($i_t \geq \bar{i}_{\text{Std}}$), introducing a standard contract allows the formation of partnership that would not otherwise contract, expanding the volume of trade. This is case 4 in Proposition 6 and region $NC \rightarrow \text{Std}$ in Figure 4.\footnote{A higher reservation value makes standardization less likely to succeed: $\partial \bar{i}_{\text{Std}} / \partial u_B > 0$. As in Figure 2, it also makes the innovative contract less likely to be used, shifting the entire locus $i_t = \bar{i}$ down. Figure 4 is drawn for $\Lambda (u_B) \in (E^2, E^2)$. If $u_B$ were higher, the locus would shift so far down that the innovative contract could not be used when $i_t = 0$, no matter how small the enforcement frictions. If $u_B$ were lower, the locus would shift so far up that the innovative contract could always be used, even when $i_t = 0$ and enforcement frictions are maximal.}

Second, when judicial biases are sufficiently rare and precedents sufficiently binding, parties always contract with each other. This is region $\text{Inn}$ in Figure 2, below the locus $i_t = \bar{i}$. Here, the introduction of a standard crowds out the use of innovative contracts, provided that the quality of standardized evidence is high enough ($i_t > \bar{i} \geq \bar{i}_{\text{Std}}$). This is case 3 in Proposition 6 and region $\text{Inn} \rightarrow \text{Std}$ in Figure 4. Partnerships would write an innovative contract, but they prefer switching to a standard contract when it is introduced. The reason is that here the costs of judicial bias are not extreme but still significant.\footnote{Changes in the buyer’s reservation value have no effect on the locus $i_t = \bar{i}$, although an increase in $u_B$ shrinks the region by shifting down the locus $i_t = \bar{i}$.}

In both regions, standardization statically improves welfare. As we show next, however, standardization may not by dynamically welfare-improving when legal evolution is considered.

### 6.3 Legal Evolution with Standardization

To evaluate the dynamic efficiency consequences of introducing a standard contract we use an intertemporally additive time-consistent utilitarian welfare function

$$W_t = \sum_{s=0}^{\infty} \delta^s E_t \Pi_{t+s} \text{ for } \delta \in (0, 1). \quad (22)$$

Recall our assumption that the standard is immediately upgraded when new precedents emerge. If a standard was introduced at time $s < t$, the standard contract at $t$ can exploit the most informative precedent $i_t$ even if $i_t > i_s$. Then the following results obtain.

**Proposition 7** Suppose that $\bar{i} = 0$ and $\alpha \sigma > 0$. As long as the informativeness of precedent is low ($0 \leq i_t \leq \bar{i}$), partnership $t$ is formed by writing an innovative contract, whether a
standard contract is available or not. As precedent becomes informative enough \( \bar{i} < i_t \leq 1 \), if a standard contract is available the parties use it and legal evolution stops.

There is a threshold \( i^* \) such that it is welfare-increasing to introduce a standard contract when precedent is sufficiently informative \( (i_t > i^*) \), but it is welfare-reducing to introduce a standard contract before then \( (i_t < i^*) \). It is optimal to standardize when precedent is still imperfect \( (i^* < 1) \), but it is optimal not to standardize when precedent is so imperfect that the standard contract would initially remain unused \( (i^* > \bar{i}) \).

Dynamically, standardization causes legal evolution to stop before it has attained the absorbing state of perfectly informative precedent. Initially, precedents are uninformative, so that the use of novel evidence is critical for contracting. In this early stage the standard contract is not used, even if available. As the law develops, however, the benefit of using the standard contract progressively increases. At some point, parties switch to it. Litigation of novel pieces of evidence, and thus legal evolution, stop.

As a consequence, society should wait to standardize until the law is sufficiently developed, even though private parties would want to use the standard before that stage is reached. The intuition is that the use of innovative contracts is a public good. Writing such contracts entail a private cost due to judicial bias but also a future social benefit in terms of greater legal evolution. Atomistic and short-lived parties do not internalize this effect. As a result, standardization should not occur when the law is insufficiently developed.

At the same time, it is not optimal not to standardize at all. Imperfect enforcement imposes first-order costs even when precedents are perfectly informative \( (i_t \to 1) \), so long as there are judges who are both capable and desirous of ignoring them \( (a \sigma > 0) \). By contrast, the benefits of further marginal improvements in informativeness are themselves marginal as precedent approaches its absorbing state. As a result, it is never optimal to wait for legal evolution to reach perfection before standardizing \( (i^* < 1) \).

Figure 5 depicts one realized path of legal evolution when only innovative contracts are available, and the long-run precedents attained for two different timings of standardization. Under early standardization the standard is introduced since the start. Under optimal standardization the standard is introduced as soon as \( i_t \) rises above the optimal threshold \( i^* \). We denote by \( i_{ES} \) the long-run precedent attained under early standardization and by \( i_{OS} \) the long run precedent attained under optimal standardization.
When standardization occurs at $t = 0$, legal evolution stops too early. This premature adoption of a standard contract maximizes the static benefit of the parties switching to the standard as soon as threshold $\bar{i}$ is surpassed. However, it hurts future parties, who would benefit from more legal evolution. The welfare-maximizing level of legal evolution is attained by standardizing only after threshold $i^*$ is crossed.

7 Heterogeneous Partnerships

By focusing on a representative partnership, Propositions 4 and 7 cannot address the possibility for standardization to increase the volume of trade.\footnote{At the initial precedent $i_0 = 0$, the innovative contract dominates the standard contract. If at $i_0 = 0$ the representative partnership is not formed through an innovative contract, it cannot be formed through a standard contract either. If instead the representative partnership is formed through an innovative contract at $i_0 = 0$, it also is for any $i_t \geq 0$. Thus, the adoption of the standard can only crowd out open-ended contracts and not also increase trade among parties.} Furthermore, a model with homogeneous partnerships cannot account for the slow diffusion of contracts documented by Choi, Gulati and Posner (2013). To consider these additional possibilities, suppose that partnerships vary in their ability to uncover novel evidence. Parties that are more resourceful, or that rely on more skilled law firms, have a greater ability to produce novel evidence. Formally, each partnership $t$ has an independently drawn realization $\iota_t$ with cumulative distribution function $F_{\iota_t}(\cdot)$ on $[0,1]$. Consider first the case without standardization.
Lemma 2  Let the ability to uncover novel evidence \( \iota_t \) be i.i.d. across partnerships with a cumulative distribution function \( F(\cdot) \) having full support on \([0, 1]\). Some partnerships are formed by writing an innovative contract for every \( \iota_t \geq 0 \) if judicial biases \( \beta \) and \( \sigma \) are sufficiently rare that

\[
\frac{\sigma}{1 - \beta} < \frac{\Lambda - \mathbb{E}_\xi^2}{\mathbb{E}_\xi - \mathbb{E}_\xi^2}.
\]

(23)

Then legal evolution in the absence of standardization is described by a Markov chain with the same properties described in Proposition 4. There is a threshold \( \bar{i}_t|_{i=0} < 1 \), increasing in enforcement frictions (\( \partial \bar{i}/\partial \beta \geq 0 \), \( \partial \bar{i}/\partial \sigma \geq 0 \), and \( \partial \bar{i}/\partial \alpha \geq 0 \)), such that when \( \iota_t \geq \bar{i}_t|_{i=0} \) all partnerships are formed.

These dynamics are analogous to those described by Proposition 4. Initially, only parties with a sufficiently high ability to search for novel evidence \( \iota_t \) choose to contract. Their contracting and litigation promote legal evolution. As precedent improves, parties characterized by lower \( \iota_t \) find it beneficial to contract. This expansion in the volume of trade speeds up legal evolution. In the limit, precedent is fully informative (\( \iota_t = 1 \) remains the unique absorbing state) and all parties contract.

The volume of contracting (though not the joint surplus it generates) reaches the optimum before precedent becomes fully informative. As in Proposition 1, partnership \( t \) is formed if and only if precedent is sufficiently informative (\( \iota_t \geq \bar{i} \)). Intuitively, partnership formation is easier for parties that can more easily collect evidence (\( \partial \lambda/\partial \iota \leq 0 \)) because their innovative contracts are less incomplete (\( \partial \lambda/\partial \iota \leq 0 \)). A sufficiently high but still imperfect quality of precedent (\( \bar{i}_t|_{i=0} < 1 \)) suffices to allow a partnership to form even if the parties are completely unable to collect novel evidence (\( \iota_t = 0 \)). Then partnership formation is assured in spite of heterogeneity. The comparative statics of Proposition 1 imply that the optimal volume of contracting is attained earlier when judicial biases (\( \beta \) and \( \sigma \)) are rarer and judicial discretion (\( \alpha \)) more limited.

We consider next what changes if standardization is implemented.

Proposition 8  Suppose that Condition (23) holds and that \( \alpha \sigma > 0 \). As long as the informativeness of precedent is low (\( 0 \leq \iota_t < \bar{i}_{\text{Std}} \)), no partnership uses a standard contract even if it is available; but as soon as the informativeness of precedent becomes sufficiently high (\( \iota_t > \bar{i}_{\text{Std}} \)) some partnerships that would have written an innovative contract switch to using a standard contract if it is available. There is a threshold \( \bar{i}_t|_{i=1} \in (\bar{i}_{\text{Std}}, 1) \), decreasing in enforcement frictions (\( \partial \bar{i}/\partial \beta \leq 0 \), \( \partial \bar{i}/\partial \sigma \leq 0 \), and \( \partial \bar{i}/\partial \alpha \leq 0 \)), such that if the informativeness of precedent is above this threshold and a standard contract is available then all partnerships use it and the evolution of precedents stops.
Suppose that enforcement frictions $\beta$, $\sigma$ and $\alpha$ are sufficiently high that

$$\frac{\alpha\sigma}{1 - \beta} > \frac{\hat{\lambda} - \int_{\xi_{\text{Std}}}^{1} xdF_{\xi}(x)}{\mathbb{E}\xi - \int_{\xi_{\text{Std}}}^{1} xdF_{\xi}(x)}.$$  \hfill (24)

Then there is a non-empty range $[\tilde{i}_{\text{Std}}, \tilde{i}_{1=0}]$ such that for values of $i_t$ in this range the standard contract allows some partnerships to be formed that could not be formed through an innovative contract.

As in Proposition 7, at low levels of legal evolution ($i_t < i_{\text{Std}}$), the standard contract is not used and the law evolves exactly as it does without standardization. As precedents become sufficiently informative ($i_t \geq i_{\text{Std}}$), the standard contract becomes preferred to writing no contract at all. At this point, standardization exerts two effects.

First, it crowds out innovative contracts, as in Proposition 7. Given that partnerships are heterogeneous, crowding out occurs gradually. As some parties (those having high $i_t$) continue to write innovative contracts, precedents—and with them the standard—keep improving. This development fosters the use of the standard contract. When the crowding out is complete, legal evolution ceases. This stop occurs while precedent is still less than fully informative, and precisely when $i_t > \tilde{i}_{1=1}$.\footnote{Parties who can more easily collect novel evidence are more likely to prefer an innovative contract that makes use of such evidence ($\partial i_t / \partial i > 0$). Innovative contracts are completely crowded out when even parties that are sure of collecting novel evidence choose the standard contract ($i_t = 1$). The development of contracts and precedents ceases when standardized precedents are still imperfect ($\tilde{i}_{1=1} < 1$). The comparative statics of Proposition 6 imply that it cases sooner if enforcement frictions ($\beta$, $\sigma$, and $\alpha$) are more severe.}

The second effect of standardization is a static expansion in the volume of trade. When $i_t \geq i_{\text{Std}}$, everybody finds it profitable to transact through the standard contract. Thus, the volume of trade increases discontinuously if there are still some partnerships (those with low $i_t$) that could not be formed through an innovative contract. This is the case when $\hat{i}_{\text{Std}} < \tilde{i}_{1=0}$. This effect of standardization is present if and only if enforcement frictions ($\beta$, $\sigma$, and $\alpha$) are high enough (condition 24).\footnote{The condition is also more likely to be satisfied when the buyer’s outside option $u_B$ is sufficiently high. The Appendix proves that there is a non-empty range of intermediate values of $\beta$, $\sigma$, $\alpha$, and $u_B$ that satisfy simultaneously conditions (23) and (24).} In this case, early standardization manages to maximize the volume of trade sooner: everybody contracts for $i_t \geq i_{\text{Std}}$ whereas without a standard contract this occurs only for $i_t \geq \tilde{i}_{1=0} > i_{\text{Std}}$.

Our model thus predicts that some partnerships tend to write innovative contracts, while other partnerships act rather as late adopters of these contracts after they become standardized. Moreover, it predicts that certain contract terms undergo a diffusion process.
They are introduced by a sophisticated innovator and then, if they manage to be incorporated into informative precedents, they gradually diffuse to less and less sophisticated parties. These patterns are consistent with the evidence in Choi, Gulati and Posner (2013).

The positive effect of standardization on the volume of trade tempers our finding that early standardization has a negative welfare impact. An impatient society ($\delta \approx 0$) may wish to standardize early in order to reach maximal trade as soon as possible. On the other hand, if the social welfare function is patient enough ($\delta \approx 1$), premature standardization remains socially suboptimal (but still demanded by current parties). Early achievement of maximum trading merely anticipates an outcome that would also be ultimately reached absent standardization. But early standardization imposes a permanent cost because it interrupts legal evolution before the law has incorporated the optimal amount of information.

8 Real World Episodes of Contract Standardization

This section presents some historical evidence corroborating our key idea that standard contracts and commercial codes can be viewed as means to reduce legal uncertainty and thus allow parties to reap gains from trade.

The largest movement toward commercial codification in modern history is the so called “golden age of commercial codification” (Gutteridge 1935), which occurred in the nineteenth century in all leading economies, including common-law countries such as Britain and its colonies. Statutes systematized prior case law and converted existing precedents into a uniform act, as in the case of the British Bills of Exchange Act 1882 and Sales of Goods Act 1893 (Ilbert 1920). Following the British lead, the United States began enacting uniform commercial legislation (Uniform Negotiable Instruments Law of 1896, Uniform Sales Act of 1906), starting a process that would eventually culminate with Llewellyn’s Uniform Commercial Code.

Legal thinkers and historians viewed this trend toward codification of commercial law as a way to create a reliable basis for contracting. The nineteenth century was a period of booming industry and long-distance trade, but enforcement risk might have hampered the process. To make enforcement more reliable, reforms focused on harmonizing, standardizing sources and facilitating the understanding of the law by both judges and the public (Diamond 1968). In more unequal societies codification was seen as the fundamental tool to eliminate the privileges of traditionally powerful landowners, which encumbered the active use and transfer of assets necessary for trade and industry (Horwitz 1977). Class
stratification contributed to create enforcement distortions, not least because judges predominantly came from the upper classes. Long-distance trade was hindered by the legal inequality among geographically distant partners. We now review two specific episodes of standardization to see the main drivers and instruments of standardization.

8.1 The Indian Codification of Contract Law

The English admirers of the French *Code Civil*, including Bentham and Macaulay, believed that, by producing fairer and more reliable enforcement, standardization would encourage trade across the diverse peoples and nations of the British Empire. Under their influence, in the nineteenth century the British strictly codified criminal and contract law in India to overhaul a chaotic juridical situation. Under the original Law Charters of India, English, Muslim and Hindu residents were to be governed by their own laws in matters of contract. Soon there was broad dissatisfaction with this principle. Traditional laws differed across religions and castes, and had minimal tradition of supporting formal contracting, while English common law had a residual role. Contractual litigation was seen as producing arbitrary resolutions, and made contracting very difficult. This resembles enforcement of our innovative contracts, which is risky because legal ambiguity affects not just novel contract features but also precedents.

After a Penal Code based on a draft by Macaulay was enacted, its success gave impulse to efforts to codify contract law. The Indian Contract Act and the Evidence Act of 1872 imposed on Indian judges a strict statutory interpretation of contracts which took precedence on other sources of law, including common, Hindu and Muslim law as well as local traditions. It stipulated general principles to define and resolve contractual conflicts, formulated explicit rules on supplying evidence to courts, and provided templates in the form of “illustrations” to highlight how judicial decisions should be guided.

The authors of the India Law Commission admitted that “we have deemed it expedient to depart ... from English law in several particulars.” The Act simplified interpretation on specific issues relative to the more nuanced common-law practice, such as in the area of contractual damages for non-performance. In England, judges had discretion on determining whether contractual provisions represented damages or penalties, which were enforced differently depending on circumstances. This required more extensive evidence gathering and legal argument. The Indian Contract Act significantly simplified the enforcement of property transfers when a buyer in good faith acquired an asset from someone in possession who was not the legitimate owner.

Even if its adoption was not voluntary, the codification of Anglo-Indian law was warmly
received in India as a more rational system of law (Derrett 1968). Codes drawn from the Indian Contract Act were subsequently introduced in East Africa and other colonies.

Consistent with our model, contract standardization in India can be seen as an attempt to reduce legal uncertainty arising from conflicting laws and insufficient jurisprudence. Interestingly, the Indian Negotiable Instruments Act preceded the equivalent British Bills of Exchange Act 1882. One explanation is that the social stratification of India, its cultural heterogeneity, and the lower expertise of its judges all contributed to increase enforcement risk, making standardization more urgent there.

8.2 The Bills of Exchange Act 1882

The Bills of Exchange Act 1882, a milestone in the process of developing negotiability of financial contracts, “codifie[d] the greater portion of the common law relating to Bills of Exchange, Cheques, and Promissory Notes” (Diamond 1968). Before this code, English law relative to bills of exchange, promissory notes and cheques was to be found in 17 statutes dealing with specific issues, and about 2,600 cases scattered over some 300 volumes of reports. The code defined a template contract that could be chosen over general contracting under common law. This standardization remarkably simplified enforcement by reducing uncertainty, and it was critical for the diffusion of financial contracting (Diamond 1968).

The extensive commentary to the Act allows some insight in identifying its effect on common-law contracting rules. The authors went to great lengths to restate the supremacy of common law: “The rules of the common law, including the law merchant, save in so far as they are inconsistent with the express provisions of this Act, shall continue to apply.” Yet they also clearly indicated that “where a rule is laid out in express terms [in the Act] ... the general [common-law] rule ought not to be applied in ... limiting its effect.”

The sharpest innovation relative to common-law practice is mentioned in the commentary to the Act §29(2), and refers to the case when under common law “a signature to a bill obtained by force and fear is valueless even in the hand of an innocent third part.” In contrast, the Act established that any promissory note that conforms to the Act held by an acquirer in good faith is always valid irrespective of any irregularity in intermediate endorsements of the bill. This provision ensured entitlement by any holder, independently from the legitimacy of all previous transfers. Another innovation of the Act is that it sets the default rule that each bill of exchange is negotiable unless explicitly excluded by the text, while previously negotiability had to be explicitly included in the text. The spirit of the Bills of Exchange Act 1882 is thus also consistent with the notion that contract standardization ensured more reliable enforcement by reducing the uncertainties involved
in contract litigation.

9 Conclusions

We offer a novel perspective on private contracting and its interaction with legal evolution. In our theory, optimal contracts do not reflect merely the information of contracting parties, as in standard contract theory, but also judicial enforcement problems, and in particular the slow-moving stock of knowledge embodied in legal precedents. Enforcement frictions make contracts sticky, giving rise to a joint evolutionary process that progressively improves judicial verifiability on the one hand and contracting on the other.

Legal development, however, can alleviate enforcement problems only in part. Eventually, judicial agency becomes very costly because it can distort highly informative, mature precedents. One desirable remedy may be the creation of standard contract templates. This reform can be justified as a means to reduce enforcement risk and allow society to exploit new opportunities for trade, especially among diverse and distant parties who are subject to severe enforcement risk. However, a strict codification of specific contracts may contribute to a more rigid legal orientation, and suppresses contractual innovation (Beck and Levine 2005). In fact, our model suggests that standardization should optimally occur after private commercial practices have developed for a while.

We discussed some historical episodes of contract standardization, but our broad message holds some relevance for current efforts to improve contract enforcement in the face of endemic legal uncertainty. In line with current real-world trends, our model suggests that standardization should be beneficial in mature domains such as international trade. Here, conflict among national laws may create strong legal uncertainty, and existing laws and trade arrangements already provide a reliable basis for harmonization. The case of developing economies is more difficult. Here, softening legal uncertainty is critical, but the undeveloped state of the law makes standardization problematic. In this case, introducing very basic (if not rudimentary) contract templates that gradually lose force as new private practices develop may statically enhance trade without stifling commercial evolution.
References


A Mathematical Appendix (For Online Publication)

A.1. Proof of Proposition 1

The minimum cost to induce effort $a$ given the non-negativity constraint is $p_0 = 0$ and $p_v = C'(a)$. Then second best effort solves the surplus-maximization problem

$$\max_{a\in[0,1]} \{av - C(a)\}$$  \hspace{1cm} (A1)

subject to the participation constraint

$$\pi_B(a) \equiv a [v - C'(a)] \geq u_B.$$  \hspace{1cm} (A2)

The buyer’s share of joint surplus $\pi_B(a)$ is a concave function:

$$\pi''_B(a) = -2C''(a) - aC'''(a) < 0$$  \hspace{1cm} (A3)

because $C''(a) > 0$ and $C'''(a) \geq 0$ for all $a \in (0,1)$. It has limits $\pi_B(0) = \pi_B(a_{FB}) = 0$ and thus a unique maximum

$$a_B = \arg\max_{a\in[0,1]} \pi_B(a) \in (0, a_{FB}).$$  \hspace{1cm} (A4)

If $u_B > \pi_B(a_B)$ the partnership is infeasible. Otherwise, second-best effort is $a_{SB} \in [a_B, a_{FB})$ such that $\pi_B(a_{SB}) = u_B$ and $\pi'_B(a_{SB}) < 0$ for all $u_B \in (0, \max_{a\in[0,1]} \{a[v - C'(a)]\})$. By the implicit-function theorem

$$\frac{\partial a_{SB}}{\partial u_B} = \frac{1}{\pi'_B(a_{SB})} < 0.$$  \hspace{1cm} (A5)

Second-best surplus is $\Pi_{SB} = a_{SB}v - C(a_{SB})$ such that

$$\frac{\partial \Pi_{SB}}{\partial u_B} = [v - C'(a_{SB})] \frac{\partial a_{SB}}{\partial u_B} < 0 \text{ for all } a_{SB} < a_{FB} \Leftrightarrow u_B > 0.$$  \hspace{1cm} (A6)

A.2. Evidence Collection

Evidence about each partnership is drawn from a universe $W$, which admits a time-invariant partition into a set $I$ of informative signals and a set $U$ of uninformative signals. The realization $e_t(w)$ of each signal $w \in W$ is independent across partnerships $t$. The universe $W$ has the cardinality of the continuum, and so do $U$ and $I$. The informative set $I$ has measure $\iota \in (0,1)$, while the uninformative set $U$ has measure $1 - \iota$. Intuitively, we can visualize the universe $W$ as a filing cabinet, whose folders $w$ are filled in each period with information $e_t(w)$ about partnership $t$.

For all uninformative signals $w \in U$, the realization $e_t(w)$ is independent of the true quality $q_t$ of the widget. It occurs as soon as partnership $t$ is formed, before contracting, investment, and production take place. One signal $u^0_t \in U$ has a misleading negative
realization \( e_t(u_t^0) = -1 \). One signal \( u_t^0 \in U \) has a misleading positive realization \( e_t(u_t^i) = 1 \). These two signals are privately observed by both parties as soon as they are realized. Thus, the contract can—and optimally will—specify that evidence based on the signals \( u_t^0 \) and \( u_t^i \) is misleading and unacceptable in court. All uninformative signals other than \( u_t^0 \) and \( u_t^i \) are realized as missing for partnership \( t \): formally, 
\[
 e_t(w) = \emptyset \forall w \in U \setminus \{u_t^0, u_t^i\}.
\]
At the time of partnership formation, uninformative signals with a missing realization are indistinguishable from informative signals that are yet to be realized. Intuitively, when the parties first meet they observe privately that the filing cabinet \( W \) has two folders (\( u_t^0 \) and \( u_t^i \)) already filled with misleading information. All other folders are empty, and the parties do not know which ones will still be empty at the time of litigation (\( w \in U \)) and which ones will instead be filled with material evidence (\( w \in I \)).

Informative signals are realized after investment, production, and consumption have taken place. The realization \( e_t(w) \) of each signal \( w \in I \) depends on the true value \( q_t \) of the widget produced by partnership \( t \), and on the time-invariant informativeness \( i(w) \) of the signal. Informativeness is described by a one-to-one function \( i : I \to [0, 1] \) that maps the set of informative signals \( I \) onto the unit interval. If the widget produced in partnership \( t \) is of high quality \((q_t = w)\), then \( e_t(w) = 1 \) for all \( w \in I \). If instead the widget is of low quality \((q_t = 0)\), then the informative pieces of evidence \( w \in I \) take values
\[
 e_t(w) = \begin{cases}
 1 & \text{if } i(w) < \xi_t \\
 -1 & \text{if } i(w) \geq \xi_t
\end{cases},
\]
where \( \xi_t \) is an i.i.d random variable with cumulative distribution function \( F_{\xi_t}(. \) and continuous density \( f_{\xi_t}(. \) > 0 on the interval \([0, 1]\).

The mapping \( i(w) \) is generically unknown to all agents. Intuitively, there is no general index to the filing cabinet \( W \). Users approaching it do not know which folders they should look into (the informative set \( I \)) and which they should avoid (the uninformative set \( U \)), and a fortiori they do not know which folders contain more diagnostic information (i.e., have higher values of \( i(w) \)). When parties write and litigate contracts tailored to their specific needs, two non-generic exceptions to this lack of knowledge emerge.

First, each partnership \( t \) optimally rules out misleading evidence \( u_t^0 \) and \( u_t^i \). Thus, the set of contracts written up to and including \( t \) characterizes a set \( U_t = \bigcup_{s=1}^{t} \{u_s^0, u_s^i\} \subset U \) of signals that are publicly known to be uninformative. For all \( t \), the set \( U_t \) is countable and thus has measure zero, while \( U \setminus U_t \) is of full measure \( 1 - \iota \).

Second, signals \( \omega \in I \) whose realizations \( e_s(\omega) \) have been used cases prior to \( t \) (and cited in the judicial opinions justifying their outcome) constitute a body of precedents \( P_t \). For each signal \( \omega \in P_t \), not only is it common knowledge that it is informative \((P_t \subset I)\), but its precise informativeness \( i(\omega) \) is known and can be contracted upon. The state of precedents at time \( t \) is summarized by the informativeness
\[
 i_t = \max_{\omega \in P_t} i(\omega)
\]
that provides a sufficient statistic for \( q_t \) that characterizes the set \( P_t \) is countable for all \( t \) and thus has measure zero, while \( I \setminus P_t \) is of full measure \( \iota \).

In court, the realization \( e_t(\omega_t^P) \) that summarizes all evidence based on precedent is

\[\text{More generally, we could allow for a stochastic countable number of misleading realizations.}\]
observed by the judge as well as the parties. All signals $\omega \in P_1 \setminus \{\omega^B_t\}$ are publicly known to provide no information conditional on $e_t(\omega^B_t)$. All signals $\omega \in U_{t-1}$ are publicly known to be uninformative. Moreover, the parties can write a contract that rules out as evidence the misleading signals $u^0_t$ and $u^v_t$. If they do not, they have an opportunity for cheap talk because they can collect and present $e_t(u^0_t) = -1$ or $e_t(u^v_t) = 1$ to a judge who only knows that $u^0_t, u^v_t \in W \setminus (P_t \cup U_{t-1})$ but has no ability to discern they belong to $U$ instead of $I$.

If the contract optimally rules out $u^0_t$ and $u^v_t$, each litigant $L \in \{B, S\}$ can search for novel evidence by inspecting the realization of an unknown signal $w^L_t \in W \setminus (P_t \cup U_t)$. The parties have no information about these signals and thus can do no better than inspecting one at random. Intuitively, each litigant checks a random folder from all those that were not previously indexed ($P_t \cup U_{t-1}$) and were also not already full of misleading evidence at the moment of partnership formation ($u^0_t$ and $u^v_t$).

With probability $\iota$, litigant $L$ inspects a novel informative signal $w^L_t \in I \setminus P_t$ whose realization $e_t(w^L_t) \in \{-1, 1\}$ is informative of true quality $q_t$. With complementary probability $1 - \iota$, the litigant inspects instead an uninformative signal $w^L_t \in U \setminus (U_t \cup \{u^0_t, u^v_t\})$ that is realized as missing: $e_t(w^L_t) = \emptyset$. Signals with a missing realization cannot be produced as evidence in court because such a realization is unprovable. It is impossible to distinguish if a litigant who claims to have observed $e_t(w^L_t) = \emptyset$ has indeed observed it, or has not observed the realization of $w^L_t$ at all, or has observed a different realization but is hiding it. On the other hand, parties are unable to fake the realization of informative signals. Thus, novel informative evidence $e_t(w^L_t)$ for $w^L_t \in I \setminus P_t$ is “hard” in the sense that it can be hidden or presented truthfully, but not falsified.

A.3. Proof of Proposition 2

We begin by solving for the optimal mechanism when misleading evidence $U_t$ is ruled out, so the litigants’ truth-telling constraints are (12) and (13). If misleading evidence were allowed, these constraints would turn into (19) and (20) because the reports $(e_B, e_S)$ would become cheap talk instead of hard evidence. We shall show that the additional constraints impose further restrictions on the optimal mechanism, proving that ruling out misleading evidence $U_t$ is optimal, according to intuition.

The first-step problem of minimizing the cost of eliciting effort $a$ is

$$\min_{p(\cdots)} \mathbb{E} [p(0, 0; e_P, e_B, e_S; b, \omega) | q_t = 0] \quad (A8)$$

subject to three equality constraints—the incentive-compatibility constraint in equation (10) and the truth-telling constraints in equations (15) and (16)—and several inequality constraints—the non-negativity and truth-telling constraints in equations (11) to (14).

Only some of the inequality constraints are binding. A first set of binding constraints reflects the impossibility of making payment contingent on direct revelation of high quality ($q_t = v$) without also paying for low quality ($q_t = 0$) to induce truthful revelation. When payment is contingent on novel evidence ($e_t(i^B_t)$ and $e_t(i^S_t)$), the minimand likelihood ratio in equation (17) is increasing in the amount of negative evidence and decreasing in the amount of positive evidence. As a consequence, a second set of binding constraints
reflects the litigants’ and biased judges’ ability to hide positive or negative evidence.

A.3.1. Pro-Buyer Judges

Pro-buyer judges use cheap talk to minimize payment, so for any report \((q_B, q_S; e_B, e_S)\) made by the litigants they enforce the same price regardless of cheap talk \(q_B, q_S \in \{0, v\}\):

\[
p(q_B, q_S; e_P, e_B, e_S; b_B, \omega) = p(e_P, e_B, e_S; b_B, \omega).
\]  (A9)

Since the seller’s payoff and a pro-buyer judge’s are antithetical, revelation of the seller’s informative private signal \((e_t (i_t^S) \neq 0)\) through a pro-buyer judge requires a payment independent of \(e_S \in \{-1, 0, 1\}\):

\[
p(e_P, e_S; b_B, \omega) = p(e_P; b_B, \omega)
\]  (A10)

for all \(e_P \in \{-1, 1\}, e_B, e_S \in \{-1, 0, 1\},\) and \(\omega \in \{0, 1\}\). When the buyer presents a positive signal \(e_t (i_t^B) = 1\), pro-buyer judges’ ability to hide information implies the binding constraints

\[
p(e_P, 1; b_B, \omega) \leq p(e_P, 0; b_B, \omega) \text{ for } e_S \in \{0, 1\}.
\]  (A11)

for all \(e_P \in \{-1, 1\}\) and \(\omega \in \{0, 1\}\).

When the buyer presents a negative signal \(e_t (i_t^B) = -1\) it provides incontrovertible evidence of low quality. Thus the non-negativity constraints binds:

\[
p(e_P, -1; b_B, \omega) = 0
\]  (A12)

for all \(e_P \in \{-1, 1\}\) and \(\omega \in \{0, 1\}\). For non-negative realizations of the buyer’s private signals, the binding truth-telling constraints impose a single price

\[
p(e_P, 0; b_B, \omega) = p(e_P, 1; b_B, \omega) = p(e_P; b_B, \omega)
\]  (A13)

for all \(e_P \in \{-1, 1\}\) and \(\omega \in \{0, 1\}\).

Intuitively, the best verification that can be obtained from pro-buyer judges is to distinguish whether the buyer can prove low quality \((e_t (i_t^B) = -1)\). If he cannot, no further nuance is possible. The contract cannot rely on evidence of low quality presented by the seller against his own interest \((e_t (i_t^S) = -1)\), nor can it ask the pro-buyer judge to raise payment when the parties have produced positive signals that he can hide \((e_t (i_t^L) = 1)\).

Whenever \(e_t (i_t) = -1\) precedent suffices to establish incontrovertible evidence of low quality so the optimal payment is nil. By the truth-telling constraint (15),

\[
p(1; b_B, 1) = p(-1; b_B, 1) = p(-1; b_B, 0).
\]  (A14)

Thus, the optimal price schedule for pro-buyer judges consists of at most two prices \(\bar{p}_B \geq 0\) and \(p_B \geq 0\) such that

\[
\bar{p}_B \equiv p(1; b_B, 1) = p(-1; b_B, 1) = p(-1; b_B, 0) \leq p(1; b_B, 0) \equiv \bar{p}_B + p_B.
\]  (A15)
Then pro-buyer judges provide a reward for high quality

$$\mathbb{E}[p(v, v; 1, e_B, e_S; b_B, \omega) | q_t = v] = \tilde{p}_B + (1 - \alpha) p_B$$

(A16)

and a wasteful payment for low quality

$$\mathbb{E}[p(0, 0; e_P, e_B, e_S; b_B, \omega) | q_t = 0] =$$

$$\tilde{p}_B \Pr \{e_t (i_t^B) \neq -1 | q_t = 0\} + (1 - \alpha) p_B \Pr \{e_t (i_t) = 1, e_t (i_t^B) \neq -1 | q_t = 0\}$$

$$= \tilde{p}_B \int_0^1 (1 - \iota + \iota x) dF_\xi (x) + (1 - \alpha) p_B \int_{i_t}^1 (1 - \iota + \iota x) dF_\xi (x).$$

(A17)

If all judges have a pro-buyer bias ($\beta = 1$) and only one among $\tilde{p}_B$ and $p_B$ is positive, the minimand likelihood ratio is respectively

$$\Lambda (\tilde{p}_B) = \int_0^1 (1 - \iota + \iota x) dF_\xi (x) \geq \Lambda (p_B) = \int_{i_t}^1 (1 - \iota + \iota x) dF_\xi (x).$$

(A18)

Thus, the optimal contract has $\tilde{p}_B = 0 \leq p_B$, unless $\alpha = 1$.

If there are both pro-buyer and unbiased judges, the truth-telling constraint (14) imposes

$$p_B \leq \min_{q \in \{0, v\}, e_P \in \{0, 1\}, e_S \in \{-1, 0, 1\}} p (q, q; 1, e_B, e_S; u)$$

(A19)

### A.3.2. Unbiased Judges

Unbiased judges impose no truth-telling constraints of their own because their preferences consist in faithfully applying the contract. Thus, in particular, it is irrelevant if an unbiased judge is bound by precedent or not because he never wishes to disregard precedents: for all $\omega \in \{0, 1\}$,

$$p (q_B, q_S; e_P, e_B, e_S; u, \omega) = p (q_B, q_S; e_P, e_B, e_S; u).$$

(A20)

On the other hand, unbiased judges introduce additional truth-telling constraints for the litigants, who must honestly report quality to a judge who is willing to make payment depend on their cheap talk if the contract so stipulates.

The buyer must be induced to reveal truthfully $q_t = v$. Then $e_t (i_t) = 1$ with certainty, while $e_t (i_t^B) = 0$ with probability $1 - \iota$ and $e_t (i_t^B) = 1$ with probability $\iota$ independent of all other random variables. Hence, we can simplify his conditional expectation and write the constraint

$$\mathbb{E}[p(v, v; 1, e_B, e_S; u) - p(0, v; 1, e_B, e_S; u) | q_t = v] \leq 0 \text{ for } e_B \in \{0, 1\}.$$  

(A21)

For ease of notation, define the conditional probability

$$F_e (e_S | v) \equiv \Pr \{e_t (i_t^S) = e_S | q_t = v\}.$$  

(A22)
Then the buyer’s truth-telling constraint is
\[
\sum_{e_S \in \{0, 1\}} F_e(e_S|v) p(v, v; 1, e_B, e_S; u) \leq \sum_{e_S \in \{0, 1\}} F_e(e_S|v) p(0, v; 1, e_B, e_S; u)
\] (A23)
for \(e_B \in \{0, 1\}\).

The seller must be induced to reveal truthfully \(q_t = 0\) even if \(e_t(i_t) = 1\):
\[
\mathbb{E}[p(0, 0; 1, e_B, e_S; u) - p(0, v; 1, e_B, 0; u) | q_t = 0, e_t(i_t) = 1, e_t(i_t^S) = e_S] \geq 0.
\] (A24)

For ease of notation, define the conditional probability
\[
F_e(e_B|q, e_P, e_S) \equiv \Pr\{e_t(i_t^B) = e_B | q_t = q, e_t(i_t) = e_P, e_t(i_t^S) = e_S\}.
\] (A25)

Then the seller’s truth-telling constraint is
\[
\sum_{e_B \in \{-1, 0, 1\}} F_e(e_B|0, 1, e_S) p(0, 0; 1, e_B, e_S; u) \geq \sum_{e_B \in \{-1, 0, 1\}} F_e(e_B|0, 1, e_S) p(0, v; 1, e_B, e_S; u)
\geq \sum_{e_B \in \{0, 1\}} F_e(e_B|0, 1, e_S) p(0, v; 1, e_B, e_S; u) \text{ for } e_S \in \{-1, 0, 1\}.
\] (A26)

The second inequality follows by the non-negativity constraint. It reflects the intuitive optimality of punishing the seller when he falsely reports \(q_S = v\) and his lie is exposed by the buyer’s hard evidence \(e_t(i_t^B) = -1\).

The buyer’s and the seller’s constraints jointly imply that
\[
\sum_{e_B \in \{0, 1\}} F_e(e_B|0, 1, 1) \sum_{e_S \in \{0, 1\}} F_e(e_S|v) p(v, v; 1, e_B, e_S; u)
\leq \sum_{e_B \in \{0, 1\}} F_e(e_B|0, 1, 1) \sum_{e_S \in \{0, 1\}} F_e(e_S|v) p(0, v; 1, e_B, e_S; u)
\leq \sum_{e_S \in \{0, 1\}} F_e(e_S|v) \sum_{e_B \in \{0, 1\}} F_e(e_B|0, 1, e_S) p(0, v; 1, e_B, e_S; u)
\leq \sum_{e_S \in \{0, 1\}} F_e(e_S|v) \sum_{e_B \in \{0, 1\}} F_e(e_B|0, 1, e_S) p(0, 0; 1, e_B, e_S; u).
\] (A27)

The first and last inequality are linear combinations of equations (A23) and (A26), respectively. The inner inequality reduces to
\[
\frac{F_e(0|0, 1, 1)}{F_e(1|0, 1, 1)} p(0, v; 1, 0, 0; u) \leq \frac{F_e(0|0, 1, 0)}{F_e(1|0, 1, 0)} p(0, v; 1, 0, 0; u)
\] (A28)
which is true for all \(p(0, v; 1, 0, 0; u) \geq 0\) because the probability that the buyer’s search is unsuccessful is \(F_e(0|0, 1, 1) = F_e(0|0, 1, 0) = 1 - \nu\) independently of the seller’s signal,
while the probability that the buyer uncovers a positive signal is increasing in the seller’s signal:

\[
F_e(1|0, 1, 1) = \frac{\int q_x dF_e(x)}{\int q_x dF_e(x)} > F_e(1|0, 1, 0) = \frac{\int q_x dF_e(x)}{1 - F_e(x)}. \tag{A29}
\]

Intuitively, a positive signal given low quality induces inference of high \(\xi_i\) and thus a higher likelihood that another signal is also positive.

We conjecture that the only binding constraint for the litigants’ truthful reporting of quality \(q_e\) is

\[
\sum_{e_E \in \{0, 1\}} \sum_{e_S \in \{0, 1\}} \frac{F_e(e_E|0, 1, 1)}{F_e(1|0, 1, 1)} F_e(e_S|v) p(v, v; 1, e_E, e_S; u) \leq \sum_{e_E \in \{0, 1\}} \sum_{e_S \in \{0, 1\}} \frac{F_e(e_E|0, 1, e_S)}{F_e(1|0, 1, e_S)} F_e(e_S|v) p(0, 0; 1, e_E, e_S; u). \tag{A30}
\]

Then another binding constraint results from the need to induce the buyer to reveal truthfully a positive signal \((e_i(i^B) = 1)\) when quality is high \((q_e = v \Rightarrow e_i(i) = 1)\):

\[
\sum_{e_S \in \{0, 1\}} F_e(e_S|v) p(v, v; 1, 1, e_S; u) \leq \sum_{e_S \in \{0, 1\}} F_e(e_S|v) p(v, v; 1, 0, e_S; u). \tag{A31}
\]

Any combination of the four prices \(p(v, v; 1, e_E, e_S; u) \geq p_B\) for \(e_E, e_S \in \{0, 1\}\) such that

\[
\sum_{e_S \in \{0, 1\}} F_e(e_S|v) p(v, v; 1, e_E, e_S; u) = p_B + t p_U \text{ for } e_E \in \{0, 1\} \tag{A32}
\]

for some constant \(p_U \geq 0\) is optimal given the truth-telling constraints we have considered so far, though only those with \(p(v, v; 1, 0, 1; u) \geq p_v(v, v; 1, 1, 0; u)\) are actually feasible, because the seller must also be incentivized to disclose a positive private signal. Then unbiased judges provide a reward for high quality

\[
\mathbb{E}[p(v, v; 1, e_E, e_S; u) | q_e = v] = \sum_{e_E \in \{0, 1\}} \sum_{e_S \in \{0, 1\}} F_e(e_E|v) F_e(e_S|v) p(v, v; 1, e_E, e_S; u) = p_B + t p_U, \tag{A33}
\]

recalling that the success of the two litigants’ searches is independent.

The wasteful payment for low quality is minimized by minimizing payment whenever a negative signal is obtained. Thus, the non-negativity constraint is binding for \(e_i(i^B) = -1:\)

\[
p(0, 0; 1, -1, e_S; u) = 0 \text{ for all } e_S \in \{-1, 0, 1\}. \tag{A34}
\]

The truth-telling constraint (A19) is binding for \(e_i(i^S) = -1:\)

\[
p(0, 0; 1, e_B, -1; u) = p_B \text{ for } e_B \in \{0, 1\}. \tag{A35}
\]
Intuitively, the seller should be punished when quality is revealed to be low. When the buyer presents a negative signal \((e_t (i_t^B) = -1)\) punishment is constrained because the seller is judgment proof. When the seller collects a negative signal \((e_t (i_t^S) = -1)\) punishment is further limited by truth-telling constraints—as we are about to show, at the optimum \(p(0, 0; 1, e_B, 0; u) = p_B\) too.

For ease of notation, define the conditional probability

\[
F_q\left(e_P, e_B, e_S|q\right) \equiv \Pr\{e_t (i_t) = e_P, e_t (i_t^B) = e_B, e_t (i_t^S) = e_S|q_t = q\}. \tag{A36}
\]

Unbiased judges enforce a wasteful payment for low quality

\[
\mathbb{E}[p(0, 0; e_P, e_B, e_S; u)|q_t = 0] = 
\sum_{e_B \in \{0, 1\}} F_q(1, e_B, -1|0) p_B + \sum_{e_B \in \{0, 1\}} \sum_{e_S \in \{0, 1\}} F_q(1, e_B, e_S|0) p(0, 0; 1, e_B, e_S; u). \tag{A37}
\]

The four prices \(p(0, 0; 1, e_B, e_S; u)\) for \(e_B, e_S \in \{0, 1\}\) are optimally set to minimize it given the binding constraint for truthful reporting of \(q_t\):

\[
\sum_{e_B \in \{0, 1\}} \sum_{e_S \in \{0, 1\}} \frac{F_e(e_B|0, 1, e_S)}{F_e(1|0, 1, e_S)} F_e(e_S|v) p(0, 0; 1, e_B, e_S; u)
= \left[ 1 + \frac{F_e(0|0, 1, 1)}{F_e(1|0, 1, 1)} \right] \left( p_B + t p_u \right). \tag{A38}
\]

Thus, all prices should be minimized except those that minimize

\[
L(e_B, e_S) \equiv F_q(1, e_B, e_S|0) \frac{F_e(1|0, 1, e_S)}{F_e(e_B|0, 1, e_S) F_e(e_S|v)}, \tag{A39}
\]

such that

\[
L(0, 0) = L(1, 0) = t \int_{i_t}^{1} x dF_\xi(x) > L(0, 1) = L(1, 1) = t \int_{i_t}^{1} x^2 dF_\xi(x). \tag{A40}
\]

By the binding truth-telling constraint (A19), the optimum is

\[
p(0, 0; 1, e_B, 0; u) = p_B \text{ for } e_B \in \{0, 1\}, \tag{A41}
\]

with any pair \(p(0, 0; 1, e_B, 1; u) \geq p_B\) for \(e_B \in \{0, 1\}\) such that

\[
\sum_{e_B \in \{0, 1\}} F_e(e_B|0, 1, 1) p(0, 0; 1, e_B, 1; u) = [F_e(0|0, 1, 1) + F_e(1|0, 1, 1)] (p_B + p_u), \tag{A42}
\]

recalling that \(F_e(1|v) = t\). Any such pair is optimal given the truth-telling constraints we have considered so far, though only those with \(p(0, 0; 1, 1, 1; u) \leq p(0, 0; 1, 0, 1; u)\) are actually feasible, because the buyer must also be induced to reveal truthfully a positive
signal when quality is low.

Then unbiased judges enforce a wasteful payment for low quality

\[ \mathbb{E} [p(0, 0; e_P, e_B, e_S; u) | q_t = 0] = p_B \int_{i_t}^1 (1 - \nu + \nu x) dF_{\xi}(x) + \nu p_U \left( (1 - \nu) \int_{i_t}^1 xdF_{\xi}(x) + \nu \int_{i_t}^1 x^2 dF_{\xi}(x) \right). \]  

(A43)

Intuitively, pro-buyer judges can be made to pay \( p_B > 0 \) when the buyer fails to present evidence of low quality only if unbiased judges make the same payment in the same conditions. Moreover, unbiased judges can make an extra payment \( p_U \geq 0 \) when not only the buyer fails to present evidence of low quality, but the seller also manages to present evidence of high quality.

If there are no pro-seller judges (\( \sigma = 0 \)) and only one among \( p_B \) and \( p_U \) is positive, the minimand likelihood ratio is respectively

\[ \Lambda(p_B) = \int_{i_t}^1 (1 - \nu + \nu x) dF_{\xi}(x) \geq \Lambda(p_U) = (1 - \nu) \int_{i_t}^1 xdF_{\xi}(x) + \nu \int_{i_t}^1 x^2 dF_{\xi}(x). \]  

(A44)

Then the optimal contract for \( \sigma = 0 \) has \( p_B = 0 < p_U \) unless \( \beta = 1 \).

### A.3.3. Pro-Seller Judges

Pro-seller judges use cheap talk to maximize payment, so for any report \( (q_B, q_S; e_B, e_S) \) made by the litigants they enforce the same price regardless of cheap talk \( q_B, q_S \in \{0, v\} \):

\[ p(q_B, q_S; e_P, e_B, e_S; b_S, \omega) = p(e_P, e_B, e_S; b_S, \omega) \]  

(A45)

Since the buyer’s payoff and a pro-seller judge’s are antithetical, revelation of the buyer’s informative private signal \( e_t(i_t^B) \neq 0 \) through a pro-seller judge requires a payment independent of \( e_B \in \{-1, 0, 1\} \):

\[ p(e_P, e_B, e_S; b_S, \omega) = p(e_P, e_S; b_S, \omega) \]  

(A46)

for all \( e_P \in \{-1, 1\}, e_B, e_S \in \{-1, 0, 1\}, \) and \( \omega \in \{0, 1\} \). When the seller presents a negative signal \( e_t(i_t^S) = -1 \), pro-seller judges’ ability to hide information implies the binding constraints

\[ p(e_P, -1; b_S, \omega) \geq p(e_P, 0; b_S, \omega) \]  

(A47)

for all \( e_P \in \{-1, 1\}, e_B \in \{-1, 0, 1\}, \) and \( \omega \in \{0, 1\} \).

Whenever \( e_t(i_t) = -1 \) precedent suffices to establish incontrovertible evidence of low quality. Thus the non-negativity constraint is binding,

\[ p(-1, e_S; b_S, 0) = 0 \]  

(A48)
for all $e_B, e_S \in \{-1, 0, 1\}$. By the truth-telling constraint (16),

$$p(-1, e_S; b_S, 1) = p(1, e_S; b_S, 1) = p(1, e_S; b_S, 0). \quad (A49)$$

Thus, the optimal price schedule for pro-seller judges consists of at most two prices $\bar{p}_S \geq 0$ and $p_S \geq 0$ such that

$$\bar{p}_S \equiv p(-1, -1; b_S, 1) = p(-1, 0; b_S, 1) = p(1, -1; b_S, \omega) = p(1, 0; b_S, \omega) \leq p(-1, 1; b_S, 1) = p(1, 1; b_S, \omega) \equiv \bar{p}_S + p_S \quad (A50)$$

for all $\omega \in \{0, 1\}$.

Intuitively, the best verification that can be obtained from pro-seller judges is to distinguish whether the seller can present evidence of high quality ($e_t (i_t^S) = 1$). If he can, no further nuance is possible. The mechanism cannot rely on evidence of high quality presented by the buyer against his own interest ($e_t (i_t^B) = 1$), nor can it ask the pro-seller judge to lower payment when the parties have produced negative signals that he can hide ($e_t (i_t^B) = -1$).

Thus, pro-seller judges provide a reward for high quality

$$\mathbb{E}[p(v, v; 1, e_B, e_S; b_S)|q_t = v] = \bar{p}_S + p_S \Pr \{e_t (i_t^S) = 1|q_t = v\} = \bar{p}_S + \nu p_S \quad (A51)$$

and a wasteful payment for low quality

$$\mathbb{E}[p(0, 0; e_P, e_B, e_S; b_S)|q_t = 0] = \bar{p}_S \left[(1 - \alpha) \Pr \{e_t (i_t) = 1|q_t = 0\} + \alpha\right]$$

$$+ p_S \left[(1 - \alpha) \Pr \{e_t (i_t) = 1, e_t (i_t^S) = 1|q_t = 0\} + \alpha \Pr \{e_t (i_t^S) = 1|q_t = 0\}\right]$$

$$= [1 - (1 - \alpha) F_{\xi}(i_t)] \bar{p}_S + \nu \left[\int_{i_t}^{1} x dF_{\xi}(x) + \alpha \int_{0}^{i_t} x dF_{\xi}(x)\right] p_S. \quad (A52)$$

If all judges have a pro-seller bias ($\sigma = 1$) and only one among $\bar{p}_S$ and $p_S$ is positive, the minimand likelihood ratio is respectively

$$\Lambda(\bar{p}_S) = 1 - (1 - \alpha) F_{\xi}(i_t) \geq \Lambda(p_S) = \int_{i_t}^{1} x dF_{\xi}(x) + \alpha \int_{0}^{i_t} x dF_{\xi}(x). \quad (A53)$$

Then the optimal contract has $\bar{p}_S = 0 < p_S$.

If there are both pro-seller and unbiased judges, the truth-telling constraint (14) imposes

$$\bar{p}_S \geq \max_{q \in \{0, v\}, e_B \in \{-1, 0, 1\}, e_S \in \{-1, 0\}} p(q, q; 1, e_B, e_S; u) \quad (A54)$$

and

$$\bar{p}_S + p_S \geq \max_{q \in \{0, v\}, e_B \in \{-1, 0, 1\}} p(q, q; 1, e_B, 1; u). \quad (A55)$$
A.3.4. Optimal Contract

Since the optimal contract for pro-seller judges has $\tilde{p}_S = 0 < p_S$, the binding truth-telling constraint (14) uniquely pins down the optimal combination of the four prices $p(v, v; 1, e_B, e_S; u) \geq p_B$ for $e_B, e_S \in \{0, 1\}$:

$$p(v, v; 1, 0, 0; u) = p(v, v; 1, 1, 0; u) = p_B < p(v, v; 1, 0, 1; u) = p(v, v; 1, 1, 1; u) = p_B + p_U, \quad (A56)$$

which enables the minimization of

$$\bar{p}_S = p_B. \quad (A57)$$

The optimal contracts for the extreme cases in which judges are respectively all pro-seller or all unbiased are ranked by

$$\Lambda(p_S) = \int_{i_t}^{1} x dF_\xi(x) + \alpha \int_{0}^{i_t} x dF_\xi(x)$$

$$> \Lambda(p_U) = (1 - \alpha) \int_{i_t}^{1} x dF_\xi(x) + \alpha \int_{0}^{i_t} x^2 dF_\xi(x). \quad (A58)$$

Intuitively, unbiased judges provide the best verification, even if they cannot achieve perfect revelation of $q_t$ for any $i_t < 1$. Thus, it is optimal to minimize $p_S$ for any $p_U$, so the binding truth-telling constraint (14) also uniquely pins down the optimal pair $p(0, 0; 1, e_B, 1; u) \geq p_B$ for $e_B \in \{0, 1\}$:

$$p(0, 0; 1, e_B, 1; u) = p(0, 0; 1, e_B, 1; u) = p_B + p_U, \quad (A59)$$

which enables the minimization of

$$p_S = p_U. \quad (A60)$$

Then, for any $p_B = \bar{p}_S \geq 0$ and $p_U = p_S \geq 0$, the optimal contract provides a reward for high quality

$$\mathbb{E}[p(v, v; 1, e_B, e_S; b, \omega) | q_t = v] = (1 - \alpha \beta) p_B + (1 - \beta) \nu p_U \quad (A61)$$

and a wasteful payment for low quality

$$\mathbb{E}[p(0, 0; 0, e_P, e_B, e_S; b, \omega) | q_t = 0] =$$

$$\left[ (1 - \alpha \beta) \int_{i_t}^{1} (1 - \nu x) dF_\xi(x) + \sigma \int_{i_t}^{1} (1 - x) dF_\xi(x) + \alpha \sigma F_\xi(i_t) \right] p_B$$

$$+ \left[ (1 - \beta) \int_{i_t}^{1} x dF_\xi(x) - (1 - \beta - \sigma) \nu \int_{i_t}^{1} x (1 - x) dF_\xi(x) + \alpha \sigma \int_{0}^{i_t} x dF_\xi(x) \right] p_U.$$  

(A62)
If only one among \( p_B \) and \( p_U \) is positive, the minimand likelihood ratio is respectively

\[
\Lambda(p_B) = \int_{i_t}^{1} (1-t+ix) \, dF_\xi(x) + \frac{\sigma}{1-\alpha\beta} \left[ t \int_{i_t}^{1} (1-x) \, dF_\xi(x) + \alpha F_\xi(i_t) \right] \tag{A63}
\]

and

\[
\Lambda(p_U) = \int_{i_t}^{1} xf\,dF_\xi(x) - \frac{1-\beta-\sigma}{1-\beta} \int_{i_t}^{1} x(1-x) \, dF_\xi(x) + \frac{\alpha\sigma}{1-\beta} \int_{0}^{i_t} xf\,dF_\xi(x). \tag{A64}
\]

**A.3.5. Few Pro-Buyer Judges**

The optimal contract sets

\[
p_B = 0 < p_U \text{ if and only if } \Lambda(p_B) \geq \Lambda(p_U), \tag{A65}
\]

namely if and only if

\[
\left( 1-t + \frac{\sigma}{1-\alpha\beta}t \right) \int_{i_t}^{1} (1-x) \, dF_\xi(x) + \frac{1-\beta-\sigma}{1-\beta} \int_{i_t}^{1} x(1-x) \, dF_\xi(x)
\geq \alpha\sigma \left[ \frac{1}{1-\beta} \int_{0}^{i_t} xf\,dF_\xi(x) - \frac{1}{1-\alpha\beta} F_\xi(i_t) \right]. \tag{A66}
\]

The left-hand side of this expression is monotone decreasing in \( i_t \). The right-hand side is maximized at \( i_t = 1 \), since it is nil at \( i_t = 0 \) and has derivative

\[
\frac{\partial}{\partial i_t} \left[ \frac{1}{1-\beta} \int_{0}^{i_t} xf\,dF_\xi(x) - \frac{1}{1-\alpha\beta} F_\xi(i_t) \right] = \left[ \frac{1}{1-\beta} i_t - \frac{1}{1-\alpha\beta} \right] f_\xi(i_t), \tag{A67}
\]

implying a unique minimum at \( i_t = (1-\beta) / (1-\alpha\beta) \). Thus the condition is satisfied for all \( i_t \in [0,1] \) if and only if it is satisfied at \( i_t = 1 \):

\[
0 \geq \alpha\sigma \left( \frac{\mathbb{E}\xi}{1-\beta} - \frac{1}{1-\alpha\beta} \right). \tag{A68}
\]

Condition (18) is sufficient but not necessary for condition (A68):

\[
\beta \leq 1 - \mathbb{E}\xi \leq \frac{1 - \mathbb{E}\xi}{1 - \alpha\mathbb{E}\xi} \Rightarrow p_B = 0 < p_U. \tag{A69}
\]

When condition (A68) holds, the optimal mechanism stipulates that the price is nil \((p(\ldots) = 0) \) except in the following two cases in which the buyer must pay the seller a positive price \( p > 0 \).

1. The judge is unbiased, evidence based on precedent is positive, the buyer does not present novel negative evidence, and the seller presents novel positive evidence.
(p(q_B, q_S; 1, 0, 1; u, \omega) = p(q_B, q_S; 1, 1, 1; u, \omega) = p \text{ for all } q_B, q_S \in \{0, v\} \text{ and } \omega \in \{0, 1\}).

2. The judge is pro-seller, the seller presents novel positive evidence, and either

(a) evidence based on precedent is positive \((p(q_B, q_S; 1, e_B, 1, b_S, \omega) = p \text{ for all } q_B, q_S \in \{0, v\}, e_B \in \{-1, 0, 1\} \text{ and } \omega \in \{0, 1\});\) or

(b) the judge can disregard contractual references to precedent \((p(q_B, q_S; e_P, e_B, 1, b_S, 1) = p \text{ for all } q_B, q_S \in \{0, v\}, e_P \in \{-1, 1\} \text{ and } e_B \in \{-1, 0, 1\}).\)

Under this optimal mechanism, all the truth-telling constraints we conjectured to be non-binding are slack. Pro-buyer judges attain their bliss point because they never enforce payment. Thus, litigants are indifferent about their reports to pro-buyer judges. Pro-seller judges have no avenue to increase payment further: they would need to disregard payment. Thus, litigants are indifferent about their reports to pro-buyer judges. Pro-buyer judges attain their bliss point because they never enforce fake positive evidence that the seller failed to present because the optimal mechanism ignores their cheap talk.

When the judge is unbiased litigants are incentivized to report truthfully quality \(q_t\) because the optimal mechanism ignores their cheap talk \(q_B, q_S\). They are incentivized to report truthfully their private signals because they cannot improve their payoffs by hiding them. The buyer may lower payment to zero by presenting \(e_t(\hat{i}^B_t) = -1\) but can never raise it by presenting \(e_t(\hat{i}^B_t) = 1\). The seller may increase it to \(p > 0\) by presenting \(e_t(\hat{i}^S_t) = 1\) but can never lower it by presenting \(e_t(\hat{i}^S_t) = -1\).

If misleading evidence \(U_t\) were allowed, this mechanism would not induce collection and revelation of novel informative evidence. On the contrary, the buyer would always collect \(u^B_t\) and the seller would always collect \(u^S_t\). Hence, the choice to rule out misleading evidence \(U_t\) and thus to avoid the additional truth-telling constraints in equations (19) and (20) is optimal.

The correspondence between the optimal direct revelation mechanism and the intuitive contract in Proposition 2 is straightforward. Under the latter, the buyer hides positive evidence \(e_t(\hat{i}^B_t) = 1\) to minimize payment and the seller hides negative evidence \(e_t(\hat{i}^S_t) = -1\) to maximize it. An unbiased judge reports truthfully all evidence presented in court. Thus, he enforces payment if and only if evidence based on precedent is positive \((e_t(\hat{i}_t) = 1)\), the seller presented further positive evidence \((e_t(\hat{i}^B_t) = 1)\), and the buyer failed to present negative evidence \((e_t(\hat{i}^S_t) \neq -1)\). A pro-buyer judge can and does hide any positive evidence presented by the seller and thus succeeds in never enforcing payment. A pro-seller judge can and does hide any negative evidence presented by the buyer. He also disregards negative evidence based on precedent if he has the ability to do so. Thus, he enforces payment whenever the seller presents positive evidence \((e_t(\hat{i}^S_t) = 1)\), unless he is bound to respect contractual references to precedent and the corresponding evidence is negative \((\omega_t = 0 \text{ and } e_t(\hat{i}_t) = -1)\).
A.3.6. Many Pro-Buyer Judges

When condition (A68) fails, the optimal mechanism sets $p_U = 0 < p_B$ and is characterized as follows.

**Proposition A1** If there are pro-seller judges with the ability to disregard contractual references to precedent and the share of pro-buyer judges is so high that

$$\alpha \sigma > 0 \text{ and } \beta > \frac{1 - \mathbb{E}[\mathcal{E}]}{1 - \alpha \mathbb{E}[\mathcal{E}]} ,$$

(A70)

then the optimal contract for partnership $t$ stipulates that uninformative evidence $\{u_i^v, u_0^v\}$ is inadmissible, and that the buyer must pay the seller a price $p > 0$ if and only if the court verifies no evidence of low quality (either novel or based upon precedent).

When condition (A68) fails, the optimal mechanism stipulates that the price is nil ($p(...) = 0$) except in the following three cases in which the buyer must pay the seller a positive price $p > 0$.

1. The judge is pro-buyer and bound to respect contractual references to precedent, evidence based on precedent is positive, and the buyer does not present novel negative evidence $(p(q_B, q_S; 1, 0, e_S; b_B, 0) = p(q_B, q_S; 1, 1, e_S; b_B, 0) = p$ for all $q_B, q_S \in \{0, v\}$ and $e_S \in \{-1, 0, 1\})$.

2. The judge is unbiased, evidence based on precedent is positive, and the buyer does not present novel negative evidence $(p(q_B, q_S; 1, 0, e_S; u, \omega) = p(q_B, q_S; 1, 1, e_S; u, \omega) = p$ for all $q_B, q_S \in \{0, v\}$, $e_S \in \{-1, 0, 1\}$, and $\omega \in \{0, 1\})$.

3. The judge is pro-seller and either
   
   (a) evidence based on precedent is positive $(p(q_B, q_S; 1, e_B, e_S; b_S, \omega) = p$ for all $q_B, q_S \in \{0, v\}$, $e_B, e_S \in \{-1, 0, 1\}$ and $\omega \in \{0, 1\})$; or
   
   (b) the judge can disregard contractual references to precedent $(p(q_B, q_S; e_P, e_B, e_S; b_S, 1) = p$ for all $q_B, q_S \in \{0, v\}$, $e_P \in \{-1, 1\}$ and $e_B, e_S \in \{-1, 0, 1\})$.

Under this alternative mechanism, just as under the one described by Proposition 2, all the truth-telling constraints we conjectured to be non-binding are slack. The seller is always indifferent about his reports $q_S$ and $e_S$, which are disregarded. The buyer is similarly indifferent about his cheap talk about quality $q_B$.

Pro-buyer judges have no avenue to lower payment further: they would need to disregard precedent when lacking the ability to do so or to fake negative evidence that the buyer failed to present, both of which are impossible. The buyer is happy to report truthfully to a pro-buyer judge because their goals coincide.

Pro-seller judges attain their bliss point if they have the ability to disregard precedent, which enables them to enforce payment in all circumstances. If they lack this ability, they have no way to increase payment when evidence based on precedent is negative. The buyer is indifferent about his report to pro-seller judges, who will completely ignore it.
When the judge is unbiased the buyer is incentivized to report truthfully his private evidence because he cannot improve his payoffs by hiding them. He may lower payment to zero by presenting \( e_t (i_t^B) = -1 \) but can never raise it by presenting \( e_t (i_t^B) = 1 \).

If misleading evidence \( U_t \) were allowed, the mechanism described by Proposition A1 would not induce collection and revelation of novel informative evidence. On the contrary, the buyer would always collect \( u_t^0 \). Hence, the choice to rule out misleading evidence \( U_t \) and thus to avoid the additional truth-telling constraints in equations (19) and (20) is optimal.

The correspondence between the optimal direct revelation mechanism and the intuitive contract in Proposition A1 is straightforward. The seller can simply avoid to collect evidence, while the buyer could present all evidence truthfully or indifferently hide positive evidence. Pro-buyer judges can avoid enforcing payment by relying on hard negative evidence provided by the buyer \( (e_t (i_t^B) = -1) \), on negative evidence based upon precedent \( (e_t (i_t) = -1) \), or on their ability to disregard contractual references to precedent \( (\omega_t = 1) \). Pro-seller judges hide negative evidence presented by the seller, but are unable to enforce payment if and only if evidence based on precedent is negative \( (e_t (i_t) = -1) \) and they are bound to respect it \( (\omega_t = 0) \).

This alternative contract differs from the one described by Proposition 2 purely in that it disregards positive novel evidence (presented by the seller in equilibrium), so as to force pro-buyer judges to enforce payment under some circumstances, despite their ability to hide such evidence.

A.4. Proof of Proposition 3 and Corollary 2

Due to the binary nature of the optimal mechanism described by Proposition 2, we can define the probability that the incentive payment \( p \) is enforced given that \( q_t = v \),

\[
\eta_v (\iota, \beta) = (1 - \beta) \iota. \tag{A71}
\]

and the probability that it is enforced when \( q_t = 0 \),

\[
\eta_0 (i_t, \iota, \beta, \sigma, \alpha) = 
\iota \left[ (1 - \beta) \int_{i_t}^{1} x dF_x (x) - (1 - \beta - \sigma) \iota \int_{i_t}^{1} x (1 - x) dF_x (x) + \alpha \sigma \int_{0}^{i_t} x dF_x (x) \right]. \tag{A72}
\]

These probabilities characterize the minimized likelihood ratio

\[
A (i_t, \iota, \beta, \sigma, \alpha) \equiv \frac{\mathbb{E} [ p (0, 0; e_P, e_B, e_S; b, \omega) | q_t = 0]}{\mathbb{E} [ p (v, v; 1, e_B, e_S; b, \omega) | q_t = v]} = \frac{\eta_0 (i_t, \iota, \beta, \sigma, \alpha)}{\eta_v (\iota, \beta)} \tag{A73}
\]

and the solution of the first-stage cost-minimization problem.\(^{30}\)

Then the seller’s incentive-compatibility constraint implies that effort \( a \) is induced at

\[^{30}\text{Under the alternative mechanism described by Proposition A1 we would have instead } \eta_{L,F} (\beta, \alpha) = (1 - \alpha \beta) \text{ and } \eta_{L,F} (i_t, \iota, \beta, \sigma, \alpha) = (1 - \alpha \beta) \int_{i_t}^{1} (1 - \iota + \omega x) dF_x (x) + \sigma \int_{i_t}^{1} (1 - x) dF_x (x) + \alpha \sigma F_x (i_t).\]
minimum cost by an incentive payment

\[ p(a; \iota, \tau, \beta, \sigma, \alpha) = \frac{C'(a)}{\eta_v(\iota, \beta) - \eta_0(\iota, \tau, \beta, \sigma, \alpha)}. \]  \hfill (A74)

Substituting this solution, the optimal contract induces effort

\[ \hat{a} = \arg \max_{a \in [0,1]} \{ av - C'(a) \} \]  \hfill (A75)

subject to the buyer’s participation constraint

\[ \pi_B(a, \Lambda) \equiv av - \left( a + \frac{\Lambda}{1 - \Lambda} \right) C'(a) \geq u_B. \]  \hfill (A76)

The buyer’s share of joint surplus \( \pi_B(a) \) is a concave function:

\[ \frac{\partial^2 \pi_B}{\partial a^2} = -2C''(a) - \left( a + \frac{\Lambda}{1 - \Lambda} \right) C'''(a) < 0 \]  \hfill (A77)

because \( C''(a) > 0 \) and \( C'''(a) \geq 0 \) for all \( a \in (0,1) \). It has limit \( \pi_B(0, \Lambda) = 0 \) and a unique maximum

\[ a_B(\Lambda) \equiv \arg \max_{a \in [0,1]} \pi_B(a, \Lambda). \]  \hfill (A78)

For sufficiently high values of \( \Lambda(\iota, \tau, \beta, \sigma, \alpha) \), contract enforcement is so poor that \( \pi_B \) is maximized at \( a = 0 \):

\[ a_B(\Lambda) = 0 \text{ for all } \Lambda \geq \frac{v}{v + C''(0)}. \]  \hfill (A79)

because

\[ \frac{\partial \pi_B}{\partial a}(0, \Lambda) = v - \frac{\Lambda}{1 - \Lambda} C''(0). \]  \hfill (A80)

By the envelope theorem,

\[ \frac{\partial \pi_B}{\partial \Lambda}(a_B(\Lambda), \Lambda) = -\frac{C'(a_B(\Lambda))}{(1 - \Lambda)^2} < 0 \text{ for all } \Lambda < \frac{v}{v + C''(0)}. \]  \hfill (A81)

In the limit as \( \Lambda \to 0 \), quality becomes perfectly contractible and

\[ \lim_{\Lambda \to 0} \pi_B(a_B(\Lambda), \Lambda) = \max_{a \in [0,1]} \{ a[v - C'(a)] \}. \]  \hfill (A82)

as in Proposition 1. Condition (6) ensures that this is greater than \( u_B \). Therefore, there is a threshold

\[ \hat{\Lambda}(u_B) \in \left[ 0, \frac{v}{v + C''(0)} \right] \]  \hfill (A83)

such that partnership \( t \) is formed if and only if \( \Lambda(\iota, \tau, \beta, \sigma, \alpha) \leq \hat{\Lambda}(u_B) \). By the implicit function theorem, \( \hat{\Lambda}(u_B) \) is decreasing in the buyer’s outside option \( u_B \).
If the partnership can be formed, optimal effort is \( \hat{a} (\underline{u}_B, \Lambda) \) such that

\[
\pi_B (\hat{a}, \Lambda) = \underline{u}_B, \tag{A84}
\]

which implies

\[
a_B (\Lambda) \leq \hat{a} (\underline{u}_B, \Lambda) < a_{FB} \tag{A85}
\]

and

\[
\frac{\partial \pi_B}{\partial \hat{a}} (\hat{a}, \Lambda) < 0 \text{ for all } \Lambda < \overset{\Lambda}{\Lambda} (\underline{u}_B). \tag{A86}
\]

By the implicit-function theorem \( \hat{a} \) is decreasing in \( \underline{u}_B \) and \( \Lambda \). Welfare is given by joint surplus

\[
\Pi = \hat{a} v - C (\hat{a}), \tag{A87}
\]

which is monotone increasing in \( \hat{a} \) for all \( \hat{a} < a_{FB} \), namely whenever \( \Lambda \underline{u}_B > 0 \).

Quality is directly contractible if and only if \( (\hat{i}_t, \hat{i} \beta, \sigma, \alpha) = 0 \). By Proposition 1, the first best is then attainable if and only if, furthermore, \( \underline{u}_B = 0 \).

Under the optimal mechanism described by Proposition 2,

\[
\Lambda (\hat{i}_t, \hat{t}, \beta, \sigma, \alpha) =
\int_{i_t}^{1} x F_{\xi} (x) \frac{1 - \beta - \sigma}{1 - \beta} t \int_{i_t}^{1} x (1 - x) dF_{\xi} (x) + \frac{\alpha \sigma}{1 - \beta} \int_{0}^{i_t} x dF_{\xi} (x), \tag{A88}
\]

such that

\[
\Lambda (\hat{i}_t, \hat{t}, \beta, \sigma, \alpha) = 0 \iff \hat{i}_t = 1 \land \alpha \sigma = 0, \tag{A89}
\]

and more generally

\[
\frac{\partial \Lambda}{\partial \hat{i}_t} = - \left[ 1 - \frac{1 - \beta - \sigma}{1 - \beta} t (1 - \hat{i}_t) - \frac{\alpha \sigma}{1 - \beta} \right] i_t f_{\xi} (i_t) \leq 0, \tag{A90}
\]

\[
\frac{\partial \Lambda}{\partial \hat{t}} = - \frac{1 - \beta - \sigma}{1 - \beta} \int_{i_t}^{1} x (1 - x) dF_{\xi} (x) \leq 0, \tag{A91}
\]

\[
\frac{\partial \Lambda}{\partial \beta} = \frac{\sigma}{(1 - \beta)^2} \left[ t \int_{i_t}^{1} x (1 - x) dF_{\xi} (x) + \alpha \int_{0}^{i_t} x dF_{\xi} (x) \right] \geq 0, \tag{A92}
\]

\[
\frac{\partial \Lambda}{\partial \sigma} = \frac{1}{1 - \beta} \left[ t \int_{i_t}^{1} x (1 - x) dF_{\xi} (x) + \alpha \int_{0}^{i_t} x dF_{\xi} (x) \right] \geq 0, \tag{A93}
\]

and

\[
\frac{\partial \Lambda}{\partial \alpha} = \frac{\sigma}{1 - \beta} \int_{0}^{i_t} x dF_{\xi} (x) \geq 0. \tag{A94}
\]

Furthermore,

\[
\frac{\partial^2 \Lambda}{\partial \hat{i}_t \partial \hat{t}} = \frac{1 - \beta - \sigma}{1 - \beta} i_t (1 - i_t) f_{\xi} (i_t) \geq 0, \tag{A95}
\]
\[ \frac{\partial^2 \Lambda}{\partial \beta \partial \iota} = \frac{\sigma}{(1 - \beta)^2} \int_{i_t}^1 x (1 - x) \, dF_x(x) \geq 0, \quad (A96) \]

and

\[ \frac{\partial^2 \Lambda}{\partial \iota \partial \sigma} = \frac{1}{1 - \beta} \int_{i_t}^1 x (1 - x) \, dF_x(x) \geq 0; \quad (A97) \]

\[ \frac{\partial^2 \Lambda}{\partial \alpha \partial i_t} = \frac{\sigma}{1 - \beta} i_t f_x(i_t) \geq 0, \quad (A98) \]

\[ \frac{\partial^2 \Lambda}{\partial \alpha \partial \beta} = \frac{\sigma}{(1 - \beta)^2} \int_0^{i_t} x \, dF_x(x) \geq 0, \quad \quad \quad (A99) \]

and

\[ \frac{\partial^2 \Lambda}{\partial \alpha \partial \sigma} = \frac{1}{1 - \beta} \int_0^{i_t} x \, dF_x(x) \geq 0; \quad (A100) \]

and finally

\[ \frac{\partial^2 \Lambda}{\partial \beta \partial i_t} = \frac{\sigma}{(1 - \beta)^2} [\alpha - \iota (1 - i_t)] i_t f_x(i_t) \quad (A101) \]

and

\[ \frac{\partial^2 \Lambda}{\partial \sigma \partial i_t} = \frac{1}{1 - \beta} [\alpha - \iota (1 - i_t)] i_t f_x(i_t), \quad (A102) \]

such that

\[ \frac{\partial^2 \Lambda}{\partial \beta \partial i_t} \geq 0 \iff \frac{\partial^2 \Lambda}{\partial \sigma \partial i_t} \geq 0 \iff i_t \geq 1 - \frac{\alpha}{\iota} \quad (A103) \]

with joint equality.

**A.5. Proof of Corollary 1**

Since \( \Lambda \) is monotone decreasing in \( i_t \), ranging from

\[ \Lambda(0, \iota, \beta, \sigma, \alpha) = \mathbb{E} \xi - \frac{1 - \beta - \sigma}{1 - \beta} \iota (\mathbb{E} \xi - \mathbb{E} \xi^2) \quad (A104) \]

to

\[ \Lambda(1, \iota, \beta, \sigma, \alpha) = \frac{\alpha \sigma}{1 - \beta} \mathbb{E} \xi, \quad \quad \quad (A105) \]

for any

\[ \frac{\alpha \sigma}{1 - \beta} \mathbb{E} \xi \leq \hat{\Lambda}(u_B) \leq \mathbb{E} \xi - \frac{1 - \beta - \sigma}{1 - \beta} \iota (\mathbb{E} \xi - \mathbb{E} \xi^2) \quad (A106) \]

we can define a threshold \( \hat{i}(\iota, \beta, \sigma, \alpha, u_B) \in [0, 1] \) such that

\[ \Lambda(i_t, \iota, \beta, \sigma, \alpha) \leq \hat{\Lambda}(u_B) \iff i_t \geq \hat{i}(\iota, \beta, \sigma, \alpha, u_B). \quad (A107) \]

By the implicit-function theorem, \( \partial \hat{\Lambda} / \partial u_B < 0 \iff \partial \hat{i} / \partial u_B > 0 \), while for each parameter \( z \in (\iota, \beta, \sigma, \alpha) \) the derivative \( \partial \hat{i} / \partial z \) has the same sign as \( \partial \Lambda / \partial z \). We can extend the


\[ i_t^b = \begin{cases} 1 & \text{if } \Lambda(u_B) < \frac{\alpha \sigma}{1 - \beta} \mathbb{E}\xi, \\ 0 & \text{if } \Lambda(u_B) > \mathbb{E}\xi - \frac{1 - \beta - \sigma}{1 - \beta} \mathbb{E}(\xi - \xi^2), \end{cases} \]  

(A108)

and

\[ i_t^s = \begin{cases} 1 & \text{if } \Lambda(u_B) < \frac{\alpha \sigma}{1 - \beta} \mathbb{E}\xi, \\ 0 & \text{if } \Lambda(u_B) > \mathbb{E}\xi - \frac{1 - \beta - \sigma}{1 - \beta} \mathbb{E}(\xi - \xi^2). \end{cases} \]  

(A109)

and the derivatives are then nil.

**A.6. Proof of Proposition 4**

The judge may write four different decisions when the optimal innovative contract from Proposition 2 is litigated.

1. The seller wins the case because he presented positive evidence \((e_t(i_t^S) = 1)\), while no negative evidence was verified.

2. The buyer wins the case because evidence based on precedent is negative \((e_t(i_t) = -1)\).

3. The buyer wins the case because he presented negative evidence \((e_t(i_t^B) = -1)\).

4. The buyer wins the case because the seller failed to present positive evidence.

If evidence based on precedent suffices to settle the case, it is summarily decided without considering novel evidence. Moreover, judges prefer citing novel evidence than grounding their ruling on the insufficiency of available evidence. Given these assumptions, the conditions under which each decision is written are the following:

1. The seller presents positive evidence \((e_t(i_t^S) = 1)\) and one of two additional contingencies is realized.

   (a) The judge is unbiased \((b_t = u)\) and neither precedent nor the buyer produce negative evidence \((e_t(i_t) = 1)\) and \((e_t(i_t^B) \in \{0,1\})\).

   (b) The judge has a pro-seller bias \((b_t = b_S)\) and evidence based on precedent is positive \((e_t(i_t) = 1)\) or can be disregarded \((\omega_t = 1)\).

2. Evidence based on precedent is negative \((e_t(i_t) = -1)\), unless the seller presents positive evidence \((e_t(i_t^S) = 1)\) and the judge has a pro-seller bias and the ability to disregard precedent \((b_t = b_S\) and \(\omega_t = 1)\).

3. The buyer presents negative evidence \((e_t(i_t^B) = -1)\), evidence based on precedent is positive \((e_t(i_t) = 1)\), and the judge does not have a pro-seller bias \((b_t \in \{b_B, u\})\).

4. One of three residual cases is realized.

   (a) The judge is pro-buyer \((b_t = b_B)\) and neither precedent nor the buyer produce negative evidence \((e_t(i_t) = 1)\) and \((e_t(i_t^B) \in \{0,1\})\).
(b) The judge is unbiased ($b_t = u$), the seller fails to present positive evidence ($e_t (i_t^S) \in \{-1, 0\}$), and neither precedent nor the buyer produce negative evidence ($e_t (i_t) = 1$ and $e_t (i_t^B) \in \{0, 1\}$).

(c) The judge is pro-seller ($b_t = b_s$), the seller fails to present positive evidence ($e_t (i_t^S) \in \{-1, 0\}$), and evidence based on precedent is positive ($e_t (i_t) = 1$).

Precedent does not evolve ($P_{t+1} = P_t$) when decision 2 or 4 is made. When decision 1 is made precedent evolves ($P_{t+1} = P_t \cup \{i_t^S\}$) but its informativeness increases only if the seller’s novel evidence happens to be more informative than existing precedents ($i_{t+1} = i_t^S > i_t$), while it is unchanged otherwise ($i_{t+1} = i_t \geq i_t^S$). When decision 3 is made precedent evolves ($P_{t+1} = P_t \cup \{i_t^B\}$) and its informativeness certainly increases ($i_{t+1} = i_t^B > i_t$).

Suppose that given the current state of precedent $i_t$, partnership $t$ is formed with an innovative contract that induces optimal effort

$$a_t = \hat{a} (u_B, \Lambda (i_t, \beta, \sigma, \alpha)) > 0. \tag{A110}$$

Then the probability that the informativeness of precedent remains unchanged is

$$\Pr (i_{t+1} = i_t | i_t) = (1 - \beta - \sigma) i_t \left[ a_t + (1 - a_t) \int_{i_t}^1 (1 - t + \mu x) dF (x) \right]$$

$$+ \sigma \left( a_t i_t + (1 - a_t) \left\{ i_t [1 - F (i_t)] + \alpha \int_0^{i_t} x dF (x) \right\} \right)$$

$$+ (1 - a_t) F (i_t) - \alpha u \sigma (1 - a_t) \int_0^{i_t} x dF (x)$$

$$+ \beta \left[ a_t + (1 - a_t) \int_{i_t}^1 (1 - t + \mu x) dF (x) \right]$$

$$+ (1 - \beta - \sigma) \left\{ a_t (1 - i_t) + (1 - a_t) \int_{i_t}^1 [1 - t + \mu^2 x (1 - x)] dF (x) \right\}$$

$$+ \sigma \left[ a_t (1 - i_t) + (1 - a_t) \int_{i_t}^1 (1 - \mu x) dF (x) \right] \right), \tag{A111}$$

where the first two lines corresponds to each subcase of decision 1 with $i_t^S \leq i_t$, the third to decision 2, and the last three to each sub-case of decision 4. Simplifying,

$$\Pr (i_{t+1} = i_t | i_t) = 1 - a_t (1 - \beta) t (1 - i_t)$$

$$- (1 - a_t) t \int_{i_t}^1 \left\{ (1 - \sigma) (1 - x) + [\sigma + (1 - \beta - \sigma) (1 - t + \mu x)] (x - i_t) \right\} dF (x). \tag{A112}$$

This rewriting is intuitive because it highlights the cases in which the informativeness of precedent improves ($i_{t+1} > i_t$). If quality is high (with probability $a_t$), a valuable new precedent is created if the seller’s search is successful (with probability $t$), his evidence happens to be more informative than the best existing precedent (with probability $1 -$
and the judge is willing to verify it because he doesn’t have a pro-buyer bias (with probability $1 - \beta$). If quality is low (with probability $1 - a_t$), a valuable new precedent can be created only if evidence based on precedent is positive ($\xi_t > i_t$). Then, one possibility is that the buyer finds negative evidence (with probability $\tau (1 - \xi_t)$), and the judge is willing to verify it because he doesn’t have a pro-seller bias (with probability $1 - \sigma$). The opposite possibility is that the seller finds evidence that is positive and yet more informative than precedents ($i_t < i_t^S < \xi_t$, with probability $\tau (\xi_t - i_t)$). A pro-seller judge always reports it to rule in the seller’s favor (with probability $\sigma$). An unbiased judge (who decides the case with probability $1 - \beta - \sigma$) does the same if and only if the buyer does not simultaneously report negative evidence (with probability $1 - \tau + \tau \xi_t$).32

The informativeness of precedent improves when decision 3 is made, and also when decision 1 is made and the seller’s novel evidence happens to be more informative than existing precedents ($i_{t+1} = i_t^S > i_t$). For every value $j \in [i_t, 1]$, the probability that the new precedent is more informative equals

$$
\Pr (i_{t+1} > j | i_t) = (1 - \beta - \sigma) \tau \left[ a_t (1 - j) + (1 - a_t) \int_j^1 (1 - \tau + \tau x) (x - j) dF_\xi (x) \right] \\
+ \sigma \tau \left[ a_t (1 - j) + (1 - a_t) \int_j^1 (x - j) dF_\xi (x) \right] \\
+ (1 - \sigma) \tau (1 - a_t) \left[ \int_{i_t}^j (1 - j) dF_\xi (x) + \int_j^1 (1 - x) dF_\xi (x) \right], \quad (A113)
$$

where the first two lines correspond to each subcase of decision 1 with $i_t^S > j$, and the last one to decision 3. Simplifying,

$$
\Pr (i_{t+1} > j | i_t) = a_t (1 - \beta) \tau (1 - j) \\
+ (1 - a_t) \tau \int_j^1 [\sigma + (1 - \beta - \sigma) (1 - \tau + \tau x)] (x - j) dF_\xi (x) \\
+ (1 - a_t) \tau (1 - \sigma) \left\{ (1 - j) [F_\xi (j) - F_\xi (i_t)] + \int_j^1 (1 - x) dF_\xi (x) \right\}. \quad (A114)
$$

In this intuitive rewriting, the first line describes the probability that precedent improves above informativeness $j$ when quality is high (with probability $a_t$). The seller’s search must be successful (with probability $\tau$), his evidence must happen to be more informative than $j$ (with probability $1 - j$), and the judge must be willing to verify it because he doesn’t have a pro-buyer bias (with probability $1 - \beta$). The second line represents the same decision in the seller’s favor when quality is actually low (with probability $1 - a_t$). Then evidence

---

31If instead a pro-seller judge disregards a negative precedent to rule in the seller’s favor, he creates a new precedent by reporting positive evidence presented by the seller, but this evidence is certainly less informative than the disregarded precedent: $e_t (i_t^B) = -1 < e_t (i_t^S) = 1 \Rightarrow i_t^S < \xi_t \leq i_t^B$.

32If an unbiased judge reports the buyer’s negative evidence, he may also report the seller’s positive evidence, but the latter is not only irrelevant for the outcome of the case but also necessarily less informative: $e_t (i_t^B) = -1 < e_t (i_t^S) = 1 \Rightarrow i_t^S < \xi_t \leq i_t^B$. 

---
based on precedent must be positive ($\xi_t > i_t$). The seller’s search must be successful (with probability $\lambda$) and it must yield evidence that is positive and yet more informative than $j$ ($j < i_t^2 < \xi_t$, with probability $\xi_t - j$). Moreover, either the judge must have a pro-seller bias (with probability $\sigma$), or else he must be unbiased (with probability $1 - \beta - \sigma$) and have observed no negative evidence produced by buyer. The latter condition obtains when the buyer’s search fails or when it uncovers positive evidence (with probability $1 - t + i\xi_t$).

Given any starting point $i_0 \geq i(t, \beta, \sigma, \alpha, u_B)$ consistent with partnership formation, the informativeness of precedent $i_t$ evolves as a time-homogeneous Markov chain with transition kernel

$$P(i, dj) = p(i, j) \, dj + r(i) \, 1_i(dj),$$

where $1_i$ denotes the indicator function $1_i(dj) = 1$ if $i \in dj$ and 0 otherwise;

$$r(i) = 1 - (1 - \beta) t (1 - i) \hat{\alpha}(u_B, \Lambda(i, t, \beta, \sigma, \alpha)) - t [1 - \hat{\alpha}(u_B, \Lambda(i, t, \beta, \sigma, \alpha))] \cdot \int_i^1 \{ (1 - \sigma)(1 - x) + [\sigma + (1 - \beta - \sigma)(1 - t + tx)](x - i) \} \, dF_\xi(x)$$

(A116)

describes the discrete probability of a transition from $i_t = i$ to $i_{t+1} = i$; and finally

$$p(i, j) = 0$$

for all $j \in [0, i]$.

(A117)

and

$$p(i, j) = (1 - \beta) i \hat{\alpha}(u_B, \Lambda(i, t, \beta, \sigma, \alpha)) + t [1 - \hat{\alpha}(u_B, \Lambda(i, t, \beta, \sigma, \alpha))] \cdot \left\{ \int_j^1 [\sigma + (1 - \beta - \sigma)(1 - t + tx)] \, dF_\xi(x) + (1 - \sigma) [F_\xi(j) - F_\xi(i)] \right\}$$

(A118)

for all $j \in (i, 1]$ jointly describe the continuous probability density of a transition from $i_t = i$ to $i_{t+1} = j$, which is positive if and only if $j > i$.

It follows that state $j$ is accessible from state $i$ if and only if $j \geq 1$. The state $i = 1$ is absorbing because it is impossible to leave: $r(1) = 1$ and $p(1, j) = 0$ for all $j \in [0, 1]$. The absorbing state is immediately accessible from any other state, so the Markov chain is absorbing.

The Markov chain can start from any $i_0 \geq 0$ if $\hat{\Lambda}(t, \beta, \sigma, \alpha, u_B) = 0$. The definition of $\hat{\Lambda}$ immediately implies that this condition coincides with

$$\hat{\Lambda} \geq \mathbb{E} \xi - \frac{1 - \beta - \sigma}{1 - \beta} t (\mathbb{E} \xi^2 - \mathbb{E} \xi^2).$$

(A119)

**A.7. Proof of Lemma 1**

Since the standard contract cannot rule out uninformative evidence $U_t$, the litigants’ reports of private signals $e_B$ and $e_S$ become mere cheap talk. Formally, the litigants’ truth-telling constraints (12) and (13) are replaced by the more restrictive (19) and (20).

Moreover, the judge’s type $\omega_t$ becomes irrelevant because all judges are fully bound
by references to precedent made by standard contracts. Thus we can drop \( \omega_t \) from the remainder of the proof, and consider judicial preferences \( b_t \in \{b_B, u, b_S\} \) only.

Litigants facing an unbiased judge can report multidimensional cheap talk \( c_L = (q_L, e_L) \). For a given value of \( e_t(i_t) = e_P \in \{-1, 1\} \) the litigants’ cheap-talk zero-sum game has a unique value \( p(e_P) \), so there is no loss of generality in making payment independent of cheap talk.

When the judge is based, he enforces the same price regardless of cheap talk \( q_B, q_S \in \{0, v\} \):

\[
p(q_B, q_S; e_P, e_B, e_S; b) = p(e_P; e_B, e_S; b) \quad \text{for } b \in \{b_B, b_S\}
\]

(\text{A120})

for all \( e_P \in \{-1, 1\} \) and \( e_B, e_S \in \{-1, 0, 1\} \). A biased judge cannot similarly manipulate evidence \( e_B, e_S \), which is cheap talk for the parties but not for judges. However, since a biased judge’s payoffs are antithetical to one litigant’s, revelation of this litigant’s evidence through a judge with the opposite bias requires a payment independent of the evidence revealed. The payment must also be independent of the evidence presented by the litigant whom the biased judge favors. Otherwise, the favored litigant would always report, and the biased judge verify, the single realization that induces the most favorable outcome.

Thus, contract enforcement can only be conditional on the evidence based on precedent \( e_t(i_t) \) and on the judge’s type:

\[
p(q_B, q_S; e_P, e_B, e_S; b) = p(e_P; b)
\]

(\text{A121})

for all \( q_B, q_S \in \{0, v\} \), \( e_P \in \{-1, 1\} \), \( e_B, e_S \in \{-1, 0, 1\} \), and \( b \in \{b_B, u, b_S\} \). The judge’s type is independent of quality \( q_t \), so without loss of generality the optimal mechanism sets a single price \( p(e_P) \) irrespective of judicial preferences.\(^{33}\)

Whenever \( e_t(i_t) = -1 \) precedent suffices to establish incontrovertible evidence of low quality. Thus the non-negativity constraint is binding,

\[
p(q_B, q_S; -1, e_B, e_S; b) = 0
\]

(\text{A122})

for all \( q_B, q_S \in \{0, v\} \), \( e_B, e_S \in \{-1, 0, 1\} \), and \( b \in \{b_B, u, b_S\} \).

The optimal standard contract requires payment of a positive price if and only if the evidence based on standardized precedents is positive:

\[
p(q_B, q_S; 1, e_B, e_S; b) = p_{\text{Std}} > 0
\]

(\text{A123})

for all \( q_B, q_S \in \{0, v\} \), \( e_B, e_S \in \{-1, 0, 1\} \), and \( b \in \{b_B, u, b_S\} \).

\(^{33}\)Since all agents are risk neutral, the mechanism could identically involve stochastic prices, randomizing on the basis of the judges’ preferences or equivalently of sunspots.
A.8. Proof of Proposition 5

Parties prefer the standard contract to the innovative contract if and only if

\[ \Lambda_{\text{Std}}(i_t) = 1 - F_\xi(i_t) < \Lambda(i_t, \nu, \beta, \sigma, \alpha) \]

\[ = \int_{i_t}^{1} x dF_\xi(x) - \frac{1 - \beta - \sigma}{1 - \beta} \int_{i_t}^{1} (1 - x) dF_\xi(x) + \frac{\alpha \sigma}{1 - \beta} \int_{0}^{i_t} x dF_\xi(x), \quad (A124) \]

namely if and only if

\[ \Delta(i_t, \nu, \beta, \sigma, \alpha) \equiv \alpha \sigma \int_{0}^{i_t} x dF_\xi(x) - (1 - \beta) \int_{i_t}^{1} (1 - x) dF_\xi(x) \]

\[ - (1 - \beta - \sigma) i_t \int_{i_t}^{1} x (1 - x) dF_\xi(x) > 0. \quad (A125) \]

The function \( \Delta \) is increasing in \( i_t \),

\[ \frac{\partial \Delta}{\partial i_t} = \alpha \sigma i_t f_\xi(i_t) + (1 - \beta) (1 - i_t) f_\xi(i_t) + (1 - \beta - \sigma) i_t (1 - x) f_\xi(i_t) > 0 \quad (A126) \]

and has limit behavior

\[ \lim_{i_t \to 0} \Delta = -(1 - \beta) \int_{0}^{1} (1 - x) dF_\xi(x) - (1 - \beta - \sigma) i_t \int_{0}^{1} x (1 - x) dF_\xi(x) \leq 0 \quad (A127) \]

and

\[ \lim_{i_t \to 1} \Delta = \alpha \sigma \int_{0}^{1} x dF_\xi(x) \geq 0. \quad (A128) \]

Thus, the condition can be rewritten

\[ i_t > \bar{i}(\nu, \beta, \sigma, \alpha) \quad (A129) \]

for a threshold \( \bar{i} \in [0, 1] \) such that

\[ \alpha \sigma \int_{0}^{\bar{i}} x dF_\xi(x) = (1 - \beta) \int_{\bar{i}}^{1} (1 - x) dF_\xi(x) + (1 - \beta - \sigma) \int_{\bar{i}}^{1} x (1 - x) dF_\xi(x). \quad (A130) \]

By the implicit-function theorem,

\[ \frac{\partial \bar{i}}{\partial \alpha} = -\frac{1}{\partial \Delta / \partial i_t} \sigma \int_{0}^{i_t} x dF_\xi(x) \leq 0, \quad (A131) \]

\[ \frac{\partial \bar{i}}{\partial \beta} = -\frac{1}{\partial \Delta / \partial i_t} \int_{i_t}^{1} (1 - x) (1 + \nu x) dF_\xi(x) \leq 0, \quad (A132) \]

\[ \frac{\partial \bar{i}}{\partial \sigma} = -\frac{1}{\partial \Delta / \partial i_t} \left[ \alpha \int_{0}^{i_t} x dF_\xi(x) + i \int_{i_t}^{1} x (1 - x) dF_\xi(x) \right] \leq 0 \quad (A133) \]
and
\[
\frac{\partial \tilde{t}}{\partial t} = \frac{1}{\partial \Delta/\partial t} (1 - \beta - \sigma) \int_{t}^{1} x (1 - x) dF_\xi(x) \geq 0. \tag{A134}
\]

### A.9. Proof of Proposition 6

Recall from Proposition 1 that partnership \( t \) is formed by writing an innovative contract if \( t \) is above a threshold \( \tilde{t}(\iota, \beta, \sigma, \alpha, u_B) \) defined by
\[
\Lambda (\tilde{t}(\iota, \beta, \sigma, \alpha, u_B), \iota, \beta, \sigma, \alpha) = \hat{\Lambda}(u_B), \tag{A135}
\]
which is increasing in the buyer’s reservation value \( (\partial \tilde{t}/\partial u_B \geq 0) \) and all enforcement frictions \( (\partial \tilde{t}/\partial \alpha \geq 0, \partial \tilde{t}/\partial \beta \geq 0, \partial \tilde{t}/\partial \sigma \geq 0, \text{ and } \partial \tilde{t}/\partial \iota \leq 0) \).

Recall from Proposition 5 that partnership \( t \) prefers the standard contract to an innovative contract if \( t \) is above a threshold \( \tilde{\iota}(\iota, \beta, \sigma, \alpha) \) defined by
\[
\Lambda_{\text{Std}} (\tilde{\iota}(\iota, \beta, \sigma, \alpha)) = \Lambda (\tilde{\iota}(\iota, \beta, \sigma, \alpha), \iota, \beta, \sigma, \alpha), \tag{A136}
\]
which is independent of the buyer’s reservation value and decreasing in all enforcement frictions \( (\partial \tilde{\iota}/\partial \alpha \leq 0, \partial \tilde{\iota}/\partial \beta \leq 0, \partial \tilde{\iota}/\partial \sigma \leq 0, \text{ and } \partial \tilde{\iota}/\partial \iota \geq 0) \).

Partnership \( t \) prefers the standard contract to no contract if \( t \) is above the threshold
\[
i_{\text{Std}}(u_B) = F_\xi^{-1} \left( 1 - \hat{\Lambda}(u_B) \right) \text{ such that } \Lambda_{\text{Std}} (i_{\text{Std}}(u_B)) = \hat{\Lambda}(u_B), \tag{A137}
\]
which is increasing in the buyer’s reservation value \( (\partial \Lambda_{\text{Std}}/\partial u_B < 0 \Rightarrow \partial i_{\text{Std}}(u_B)/\partial u_B > 0) \) and independent of enforcement frictions.

Since \( 0 > \partial \Lambda/\partial \iota_t > \partial \Lambda_{\text{Std}}/\partial \iota_t \), these definitions imply that
\[
i(\iota, \beta, \sigma, \alpha, u_B) > i_{\text{Std}}(u_B) \iff \\
\Lambda_{\text{Std}} (i_{\text{Std}}(u_B)) = \hat{\Lambda}(u_B) = \Lambda (i(\iota, \beta, \sigma, \alpha, u_B), \iota, \beta, \sigma, \alpha) < \Lambda (i_{\text{Std}}(u_B), \iota, \beta, \sigma, \alpha) \iff i_{\text{Std}}(u_B) > \tilde{\iota}(\iota, \beta, \sigma, \alpha). \tag{A138}
\]

Therefore, the five-dimensional parameter space consisting of the buyer’s reservation value \( (u_B) \) and all enforcement frictions \( (\beta, \sigma, \alpha, \text{ and } 1 - \iota) \) can be partitioned into two regions separated by the four-dimensional plane
\[
\hat{\Lambda}(u_B) = \Lambda \left( F_\xi^{-1} \left( 1 - \hat{\Lambda}(u_B) \right) \right), \iota, \beta, \sigma, \alpha) \iff \\
i(\iota, \beta, \sigma, \alpha, u_B) = i_{\text{Std}}(u_B) = \tilde{\iota}(\iota, \beta, \sigma, \alpha). \tag{A139}
\]
Depending on the informativeness of precedent \( (i_t) \), each of these two subspaces can be further partitioned into three regions.
1. If the buyer’s reservation value and enforcement frictions are sufficiently low, then

\[ \hat{\Lambda}(\hat{u}_B) \geq \Lambda \left( F_{\xi}^{-1} \left( 1 - \hat{\Lambda}(\hat{u}_B) \right) \right), \lambda, \beta, \sigma, \alpha \] \iff

\[ \hat{i}(\lambda, \beta, \sigma, \alpha, \hat{u}_B) \leq \hat{i}_{\text{Std}}(\hat{u}_B) \leq \hat{i}(\lambda, \beta, \sigma, \alpha). \] (A140)

As the informativeness of precedent varies:

(a) For \( i_t < \hat{i}(\lambda, \beta, \sigma, \alpha, \hat{u}_B) \) partnership \( t \) cannot be formed through either type of contract.

(b) For \( \hat{i}_{\text{Std}}(\hat{u}_B) \leq i_t \leq \hat{i}(\lambda, \beta, \sigma, \alpha) \) partnership \( t \) is always formed through an innovative contract.

(c) For \( i_t > \hat{i}(\lambda, \beta, \sigma, \alpha) \) partnership \( t \) is formed through an innovative contract, but under standardization it uses the standard contract.

2. If the buyer’s reservation value and enforcement frictions are too high, then

\[ \Lambda \left( F_{\xi}^{-1} \left( 1 - \hat{\Lambda}(\hat{u}_B) \right) \right), \lambda, \beta, \sigma, \alpha > \hat{\Lambda}(\hat{u}_B) \iff

\[ \hat{i}(\lambda, \beta, \sigma, \alpha, \hat{u}_B) < \hat{i}_{\text{Std}}(\hat{u}_B) < \hat{i}(\lambda, \beta, \sigma, \alpha). \] (A141)

As the informativeness of precedent varies:

(a) For \( i_t < \hat{i}_{\text{Std}}(\hat{u}_B) \) partnership \( t \) cannot be formed through either type of contract.

(b) For \( \hat{i}_{\text{Std}}(\hat{u}_B) \leq i_t < \hat{i}(\lambda, \beta, \sigma, \alpha, \hat{u}_B) \) partnership \( t \) cannot be formed through an innovative contract, but it can be formed through a standard contract.

(c) For \( i_t \geq \hat{i}(\lambda, \beta, \sigma, \alpha, \hat{u}_B) \) partnership \( t \) is formed through an innovative contract, but under standardization it uses the standard contract.

The economic situation of partnership \( t \) can be summarized by four cases.

1. Partnership \( t \) cannot be formed either through either type of contract if

\[ \hat{\Lambda}(\hat{u}_B) < \min \{ \Lambda(i_t, \lambda, \beta, \sigma, \alpha), \Lambda_{\text{Std}}(i_t) \} \]

\[ \iff i_t < \min \{ \hat{i}(\lambda, \beta, \sigma, \alpha, \hat{u}_B), \hat{i}_{\text{Std}}(\hat{u}_B) \}. \] (A142)

This region has non-zero measure, because

\[ \hat{\Lambda}(\hat{u}_B) \leq \frac{v}{v + C''(0)} < 1 \implies \hat{i}_{\text{Std}}(\hat{u}_B) > 0 \text{ for all } \hat{u}_B \geq 0, \] (A143)

while

\[ \hat{i}(\lambda, \beta, \sigma, \alpha, \hat{u}_B) > 0 \]

\[ \iff \hat{\Lambda}(\hat{u}_B) < \Lambda(0, \lambda, \beta, \sigma, \alpha) = \mathbb{E}\xi - \frac{1 - \beta - \sigma}{1 - \beta} \mathbb{E}(\mathbb{E}\xi - \mathbb{E}\xi^2) \] (A144)
is satisfied by a non-empty range of reservation values consistent with condition (6):

\[
\max_{a \in [0,1]} \left\{ av - \left( a + \frac{\Lambda (0, t, \beta, \sigma, \alpha)}{1 - \Lambda (0, t, \beta, \sigma, \alpha)} \right) C'(a) \right\} < u_B \leq \max_{a \in [0,1]} \{ a [v - C'(a)] \}, \quad (A145)
\]

which is non-empty even in the absence of enforcement frictions \((\Lambda (0, 1, 0, 0, 0) = \mathbb{E} \xi^2 \geq 0)\) and expands as enforcement frictions increase \((\text{up to } \Lambda (0, 0, 0, 1, 1) = \mathbb{E} \xi)\).

2. Partnership \(t\) is formed through an innovative contract irrespective of standardization if

\[
\Lambda (i_t, t, \beta, \sigma, \alpha) \leq \min \left\{ \hat{\Lambda} (u_B), \Lambda_{\text{std}} (i_t) \right\} \leftrightarrow \check{i} (t, \beta, \sigma, \alpha, u_B) \leq i_t \leq \check{i} (t, \beta, \sigma, \alpha), \quad (A146)
\]

which requires sufficiently low reservation value \((u_B)\) and enforcement frictions \((\beta, \sigma, \alpha, \text{and } 1 - t)\).

In the limit as enforcement frictions disappear

\[
\Lambda (i_t, 1, 0, 0, 0) = \int_{i_t}^{1} x^2 dF_\xi (x) \leq \Lambda_{\text{std}} (i_t) = 1 - F_\xi (i_t), \quad (A147)
\]

with strict inequality if \(i_t < 1\). If furthermore the buyer’s reservation value is nil

\[
\Lambda (i_t, 1, 0, 0, 0) = \int_{i_t}^{1} x^2 dF_\xi (x) \leq \hat{\Lambda} (0) = \frac{v}{v + C''(0)} \quad (A148)
\]

for a non-empty range of values \(i_t < 1\). By continuity, the conditions are satisfied for \(\beta \approx 0, \sigma \approx 0, \alpha \approx 0, t \approx 1\) and \(u_B \approx 0\) in a region with non-zero measure.

3. Partnership \(t\) is formed through an innovative contract but under standardization it uses the standard contract if

\[
\Lambda_{\text{std}} (i_t) < \Lambda (i_t, t, \beta, \sigma, \alpha) \leq \hat{\Lambda} (u_B) \leftrightarrow i_t \geq \max \{ \check{i} (t, \beta, \sigma, \alpha, u_B), \check{i} (t, \beta, \sigma, \alpha) \}. \quad (A149)
\]

For all \(\alpha \sigma > 0\), in the limit as precedent becomes perfectly informative

\[
\Lambda_{\text{std}} (1) = 0 < \Lambda (1, t, \beta, \sigma, \alpha) = \frac{\alpha \sigma}{1 - \beta} \mathbb{E} \xi. \quad (A150)
\]
Moreover

\[ \dot{i}(\iota, \beta, \sigma, \alpha, u_B) < 1 \]

\( \Leftrightarrow \Lambda(1, \iota, \beta, \sigma, \alpha) = \frac{\alpha \sigma}{1 - \beta} \int_0^{i_t} xdF_\xi(x) < A(u_B) \leq \frac{v}{v + C''(0)} \) (A151)

for a non-empty range of values \( \alpha \sigma > 0 \) and \( u_B \simeq 0 \). By continuity, the conditions are satisfied for \( u_B \simeq 0 \) and \( i_t \simeq 1 \) in a region with non-zero measure.

4. Partnership \( t \) cannot be formed through an innovative contract but it can be formed through a standard contract if

\[ A_{Std}(i_t) \leq \hat{\Lambda}(u_B) < A(i_t, \iota, \beta, \sigma, \alpha) \Leftrightarrow \hat{\iota}_{Std}(u_B) \leq \dot{i}_t < \hat{i}(\iota, \beta, \sigma, \alpha, u_B), \] (A152)

which requires sufficiently high reservation value \( (u_B) \) and enforcement frictions \( (\beta, \sigma, \alpha, \text{and} \ 1 - \iota) \).

In the limit as enforcement frictions reach their maximum,

\[ \Lambda(i_t, 0, 0, 1, 1) = E \xi \text{ for all } i_t, \] (A153)

so the condition reduces to

\[ \max_{a \in [0,1]} \left\{ av - \left( a + \frac{E \xi}{1 - E \xi} \right) C'(a) \right\} < \]

\[ u_B \leq \max_{a \in [0,1]} \left\{ av - \left( a + \frac{1 - F_\xi(i_t)}{F_\xi(i_t)} \right) C'(a) \right\}, \] (A154)

which is satisfied by a non-empty range of reservation values provided that \( i_t > F_\xi^{-1}(1 - E \xi) \). By continuity, the conditions are satisfied for \( i_t \simeq 0, \beta \simeq 0, \sigma \simeq 1 \) and \( \alpha \simeq 1 \) in a region with non-zero measure.

**A.10. Proof of Proposition 7**

The first part of the proposition follows immediately from Propositions 5 and 4. The standard contract is preferred if and only if \( i_t \geq \tilde{i}(\iota, \beta, \sigma, \alpha) > 0 \). The evolution of precedent is ever improving: \( i_{t+1} \geq i_t \). Initially, for \( i_0 = 0 \) and any subsequent \( i_t \leq \tilde{i}(\iota, \beta, \sigma, \alpha) \), partnership \( t \) is formed with an innovative contract irrespective of the availability of a standard contract. As soon as the Markov chain reaches for the first time a value \( i_t \geq \tilde{i}(\iota, \beta, \sigma, \alpha) \), parties switch to the standard contract. Since no new evidence is verified under the optimal standard contract, the evolution of precedent stops.

The joint surplus of partnership \( t \) is

\[ \Pi_t = \Pi(u_B, \Lambda_t) \equiv va(u_B, \Lambda_t) - C(a(u_B, \Lambda_t)) \] (A155)
where
\[
a (u_B, \Lambda_t) = \arg \max_{a \in [0,1]} \{au - C(a)\} \text{ s.t. } au - \left(a + \frac{\Lambda_t}{1 - \Lambda_t}\right) C'(a) \geq u_B \tag{A156}
\]
and
\[
\Lambda_t = \max \{\Lambda_{Std}(i_t), \Lambda(i_t, \iota, \beta, \sigma, \alpha)\}. \tag{A157}
\]

By Proposition 1, \(\Pi\) is a continuous and monotone strictly decreasing function of \(\Lambda_t \in [0, \hat{\Lambda}(u_B)]\).

For \(i_t = 1\), the likelihood ratio of the two contracts is
\[
\Lambda_{Std}(1) = 0 < \Lambda(1, \iota, \beta, \sigma, \alpha) = \frac{\alpha \sigma}{1 - \beta} \hat{E} \xi \text{ for all } \alpha \sigma > 0 \tag{A158}
\]

Then, given \(i_t = 1\) social welfare without a standard is
\[
W_{Inn}(u_B, 1, \iota, \beta, \sigma, \alpha) = \frac{1}{1 - \delta} \Pi \left(\frac{\alpha \sigma}{1 - \beta} \hat{E} \xi\right) \tag{A159}
\]

while under standardization it is
\[
W_{Std}(u_B, 1) = \frac{1}{1 - \delta} \Pi(u_B, 0) > W_{Inn}(u_B, 1, \iota, \beta, \sigma, \alpha). \tag{A160}
\]

By continuity, for \(i_t = 1 - \varepsilon\) the likelihood ratios are
\[
\Lambda_{Std}(1 - \varepsilon) = o(\varepsilon) \text{ and } \Lambda(1 - \varepsilon, \iota, \beta, \sigma, \alpha) = \frac{\alpha \sigma}{1 - \beta} \hat{E} \xi + o(\varepsilon) \tag{A161}
\]

and joint profits are respectively
\[
\Pi(u_B, \Lambda_{Std}(1 - \varepsilon)) = \Pi(u_B, 0) - o(\varepsilon) \tag{A162}
\]

and
\[
\Pi(u_B, \Lambda(1 - \varepsilon, \iota, \beta, \sigma, \alpha)) = \Pi \left(u_B, \frac{\alpha \sigma}{1 - \beta} \hat{E} \xi\right) - o(\varepsilon). \tag{A163}
\]

Social welfare under standardization is
\[
W_{Std}(u_B, 1 - \varepsilon) = \frac{1}{1 - \delta} \Pi(u_B, \Lambda_{Std}(1 - \varepsilon)) = \frac{1}{1 - \delta} \Pi(u_B, 0) - o(\varepsilon), \tag{A164}
\]

68
while social welfare without a standard is

\[ W_{Inn}(u_B, 1 - \varepsilon, \iota, \beta, \sigma, \alpha) = \]

\[ \Pi(u_B, \Lambda(1 - \varepsilon, \iota, \beta, \sigma, \alpha)) + \sum_{s=1}^{\infty} \delta^s \mathbb{E}[\Pi(u_B, \Lambda(i_{t+s}, \iota, \beta, \sigma, \alpha)) | i_t = 1 - \varepsilon] \]

\[ < \Pi(u_B, \Lambda(1 - \varepsilon, \iota, \beta, \sigma, \alpha)) + \frac{\delta}{1 - \delta} \Pi(u_B, \Lambda(1, \iota, \beta, \sigma, \alpha)) \]

\[ = \frac{1}{1 - \delta} \Pi\left(\frac{\alpha \sigma}{1 - \beta} \mathbb{E} \xi\right) - o(\varepsilon), \quad (A165) \]

considering that the best case is a jump to the absorbing state \( i_{t+1} = 1 \). Thus,

\[ W_{Std}(u_B, 1 - \varepsilon) = \frac{1}{1 - \delta} \Pi(u_B, 0) - o(\varepsilon) \]

\[ > \frac{1}{1 - \delta} \Pi\left(\frac{\alpha \sigma}{1 - \beta} \mathbb{E} \xi\right) - o(\varepsilon) > W_{Inn}(u_B, 1 - \varepsilon, \iota, \beta, \sigma, \alpha) \quad (A166) \]

for strictly positive values of \( \varepsilon \). As a consequence, there is a non-empty left neighborhood of 1, which we can call \((i^*, 1)\) for \( i^* < 1 \), such that standardization is welfare-increasing for \( i_t \in (i^*, 1) \).

For \( i_t = \bar{i}(\iota, \beta, \sigma, \alpha) \), by definition the likelihood ratio of the two contracts is identical:

\[ \Lambda_{Std}(\bar{i}(\iota, \beta, \sigma, \alpha)) = \Lambda(\bar{i}(\iota, \beta, \sigma, \alpha), \iota, \beta, \sigma, \alpha) \equiv \bar{\Lambda}. \quad (A167) \]

Thus joint surplus for partnership \( t \) is identical under the two contracts. Then social welfare without a standard is

\[ W_{Inn}(u_B, \bar{i}(\iota, \beta, \sigma, \alpha), \iota, \beta, \sigma, \alpha) = \]

\[ \Pi(u_B, \bar{\Lambda}) + \sum_{s=1}^{\infty} \delta^s \mathbb{E}[\Pi(u_B, \Lambda(i_{t+s}, \iota, \beta, \sigma, \alpha)) | i_t = \bar{i}(\iota, \beta, \sigma, \alpha)] > \frac{1}{1 - \delta} \Pi(u_B, \bar{\Lambda}) \quad (A168) \]

since there is strictly positive probability that the transient state \( i_t = \bar{i}(\iota, \beta, \sigma, \alpha) \) will be abandoned.

If parties adopted the standard for \( i_t = \bar{i}(\iota, \beta, \sigma, \alpha) \), when they are indifferent between using it or writing an open ended contract, social welfare under standardization would be

\[ W_{Std}(u_B, \bar{i}(\iota, \beta, \sigma, \alpha)) = \frac{1}{1 - \delta} \Pi(u_B, \bar{\Lambda}) < W_{Inn}(u_B, \bar{i}(\iota, \beta, \sigma, \alpha), \iota, \beta, \sigma, \alpha). \quad (A169) \]

By continuity, if \( i_t = \bar{i}(\iota, \beta, \sigma, \alpha) + \varepsilon \) the likelihood ratios are

\[ \Lambda_{Std}(\bar{i} + \varepsilon) = \bar{\Lambda} - o(\varepsilon) \quad (A170) \]
and
\[ \Lambda (i + \varepsilon, 1, \beta, \sigma, \alpha) = \bar{\Lambda} - o(\varepsilon), \]  
(A171)
such that
\[ W_{\text{inn}}(u_B, i + \varepsilon, 1, \beta, \sigma, \alpha) > \frac{1}{1-\delta} \Pi(u_B, \bar{\Lambda}) + o(\varepsilon) \]  
(A172)
and
\[ W_{\text{std}}(u_B, i + \varepsilon) = \frac{1}{1-\delta} \Pi(u_B, \bar{\Lambda}) + o(\varepsilon). \]  
(A173)

Thus for sufficiently small \( \varepsilon > 0 \) the adoption of the standard contract is welfare reducing; but
\[ \Lambda_{\text{std}}(i + \varepsilon) < \Lambda (i + \varepsilon, 1, \beta, \sigma, \alpha) \]  
(A174)
so the standard is adopted by parties if it is available. Thus, there is a non-empty right neighborhood of \((i, \beta, \sigma, \alpha), (i, \beta, \sigma, \alpha), i^*\), such that for \( i \in (i, i^*) \) standardization is welfare-reducing.

If a standard is introduced while \( i \in [0, \bar{i}] \), it is adopted as soon as \( i_{t+s} > i(i, \beta, \sigma, \alpha) \). By Proposition 4, there is strictly positive probability that the first such jump leads to a state \( i_{t+s} \in (i, i^*) \) such that standardization is welfare-reducing. If instead the jump leads to a state in \( i_{t+s} \in (i^*, 1] \) the standard can be introduced before partnership \( t+s \) is formed. Thus, standardization when \( i \in [0, \bar{i}] \) is also welfare-reducing.

### A.11. Proof of Lemma 2

There are partnerships that are willing to form through an innovative contract if and only if
\[ i_t \geq i(1, \beta, \sigma, u_B) \iff \Lambda (i_t, 1, \beta, \sigma, \alpha) \leq \hat{\Lambda}(u_B), \]  
(A175)
where
\[ i(1, \beta, \sigma, u_B) = 0 \iff \Lambda (0, 1, \beta, \sigma, \alpha) \equiv \frac{\sigma}{1-\beta} E\xi + \left(1 - \frac{\sigma}{1-\beta}\right) E\xi^2 \leq \hat{\Lambda}(u_B). \]  
(A176)

All partnerships are willing to form through an innovative contract if and only if
\[ i_t \geq i(0, \beta, \sigma, u_B) \iff \Lambda (i_t, 0, \beta, \sigma, \alpha) \leq \hat{\Lambda}(u_B), \]  
(A177)
where
\[ i(0, \beta, \sigma, u_B) > 0 \iff \hat{\Lambda}(u_B) < \Lambda (0, 0, \beta, \sigma, \alpha) \equiv E\xi \]  
(A178)
and
\[ i(0, \beta, \sigma, u_B) < 1 \iff \hat{\Lambda}(u_B) > \Lambda (1, 0, \beta, \sigma, \alpha) \equiv \frac{\alpha \sigma}{1-\beta} E\xi, \]  
(A179)
which is weaker than condition (A176) and necessary for an innovative contract ever to be written.

Suppose that \( \beta + \sigma < 1 \) and that condition (A176) holds. Then for all \( i \in [0, i(0, \beta, \sigma, u_B) < 1 \iff \hat{\Lambda}(u_B) > \Lambda (1, 0, \beta, \sigma, \alpha) \equiv \frac{\alpha \sigma}{1-\beta} E\xi, \]  
(A179)
which is weaker than condition (A176) and necessary for an innovative contract ever to be written.

Suppose that \( \beta + \sigma < 1 \) and that condition (A176) holds. Then for all \( i \in [0, i(0, \beta, \sigma, u_B) < 1 \iff \hat{\Lambda}(u_B) > \Lambda (1, 0, \beta, \sigma, \alpha) \equiv \frac{\alpha \sigma}{1-\beta} E\xi, \]  
(A179)
which is weaker than condition (A176) and necessary for an innovative contract ever to be written.

Suppose that \( \beta + \sigma < 1 \) and that condition (A176) holds. Then for all \( i \in [0, i(0, \beta, \sigma, u_B) < 1 \iff \hat{\Lambda}(u_B) > \Lambda (1, 0, \beta, \sigma, \alpha) \equiv \frac{\alpha \sigma}{1-\beta} E\xi, \]  
(A179)
which is weaker than condition (A176) and necessary for an innovative contract ever to be written.
the partnership is not formed. If precedent is described by Proposition 1 it follows that

\[ \alpha, u_B \] there is a threshold

\[ \xi (u_B, i_t, \beta, \sigma, \alpha) = \frac{(1 - \beta) \left( \int_{i_t}^{1} xdF_\xi (x) - \hat{A} (u_B) \right) + \alpha \sigma \int_{0}^{i_t} xdF_\xi (x)}{(1 - \beta - \sigma) \int_{i_t}^{1} x(1 - x) dF_\xi (x)} \in [0, 1] \quad \text{(A180)} \]

such that partnership \( t \) is formed if and only if \( \iota_t \in [\xi, 1] \). From the comparative statics derived in Proposition 1 it follows that \( \partial\xi / \partial u_B > 0, \partial\xi / \partial i_t < 0, \partial\xi / \partial \beta \geq 0, \partial\xi / \partial \sigma > 0, \) and \( \partial\xi / \partial \alpha \geq 0 \). We extend the definition to \( \xi (u_B, i_t, \beta, \sigma, \alpha) = 0 \) for all \( \iota_t \geq i_t (\beta, \sigma, \alpha, u_B) \).

In period \( t \), if the parties draw an ability to collect novel evidence \( \iota_t < \xi (u_B, i_t, \beta, \sigma, \alpha) \) the partnership is not formed. If \( \iota_t \geq \xi (u_B, i_t, \beta, \sigma, \alpha) \) the partnership is formed and the seller exerts effort

\[ a_t = \hat{a} (u_B, \Lambda (i_t, \iota_t, \beta, \sigma, \alpha)) > 0. \quad \text{(A181)} \]

Considering that \( \iota_t \) is a random draw from the distribution \( F_i (.) \), the evolution of precedent is described by

\[
\Pr (i_{t+1} > j | i_t) = (1 - \beta) \iota (1 - j) \int_{\xi (u_B, i_t, \beta, \sigma, \alpha)}^{1} \hat{a} (u_B, \Lambda (i_t, h, \beta, \sigma, \alpha)) dF_i (h) \\
+ \iota \left\{ \int_{j}^{i_t} [\sigma + (1 - \beta - \sigma)(1 - \iota + tx)] (x - j) dF_\xi (x) + (1 - \sigma) \int_{i_t}^{1} (1 - \max \{j, x\}) dF_\xi (x) \right\} \\
\cdot \int_{\xi (u_B, i_t, \beta, \sigma, \alpha)}^{1} [1 - \hat{a} (u_B, \Lambda (i_t, h, \beta, \sigma, \alpha))] dF_i (h) \text{ for all } j \in [i_t, 1]. \quad \text{(A182)}
\]

Thus, it is represented by an absorbing Markov chain with the same qualitative properties described by Proposition 4.

### A.12. Proof of Proposition 8

There are partnerships that prefer a standard contract to an innovative contract if and only if

\[ i_t > \bar{i} (0, \beta, \sigma, \alpha) > 0 \Leftrightarrow \Lambda_{Std} (i_t) < \Lambda (i_t, 0, \beta, \sigma, \alpha), \quad \text{(A183)} \]

where \( \Lambda_{Std} (0) \equiv 1 > \Lambda (0, 0, \beta, \sigma, \alpha) \equiv \mathbb{E} \xi \) implies that \( \bar{i} (0, \beta, \sigma, \alpha) > 0 \).

All partnerships prefer a standard contract to an innovative contract if and only if

\[ \bar{i} (1, \beta, \sigma, \alpha) < i_t \leq 1 \Leftrightarrow \Lambda_{Std} (i_t) < \Lambda (i_t, 1, \beta, \sigma, \alpha), \quad \text{(A184)} \]

where \( \Lambda_{Std} (1) \equiv 0 < \Lambda (1, 1, \beta, \sigma, \alpha) \equiv [\alpha \sigma / (1 - \beta)] \mathbb{E} \xi \) implies \( \bar{i} (1, \beta, \sigma, \alpha) < 1 \) for all \( \alpha, \sigma \). If \( \beta + \sigma < 1 \), for all \( i_t \in [\bar{i} (0, \beta, \sigma, \alpha), \bar{i} (1, \beta, \sigma, \alpha)] \) there is a threshold

\[ \bar{i} (i_t, \beta, \sigma, \alpha) \equiv \frac{\alpha \sigma \int_{0}^{i_t} xdF_\xi (x) - (1 - \beta) \int_{i_t}^{1} (1 - x) dF_\xi (x)}{(1 - \beta - \sigma) \int_{i_t}^{1} x(1 - x) dF_\xi (x)} \in [0, 1] \quad \text{(A185)} \]
such that partnership $t$ prefers the standard contract to an innovative contract if and only if $i_t \in [0, \bar{i}(i_t, \beta, \sigma, \alpha)]$. Its comparative statics are $\partial \bar{i} / \partial i_t > 0$, $\partial \bar{i} / \partial \beta > 0$, $\partial \bar{i} / \partial \sigma > 0$, and $\partial \bar{i} / \partial \alpha \geq 0$. We extend the definition to $\bar{i}(i_t, \beta, \sigma, \alpha) = 0$ for all $i_t \leq \bar{i}(0, \beta, \sigma, \alpha)$ and $\bar{i}(i_t, \beta, \sigma, \alpha) = 1$ for all $i_t \geq \bar{i}(1, \beta, \sigma, \alpha)$.

Recall that each and every partnership is willing to form under a standard contract if and only if

$$i_t \geq \bar{i}_{std}(u_B) \equiv F_{\xi}^{-1} \left(1 - \hat{\Lambda}(u_B)\right) \Leftrightarrow \Lambda_{std}(i_t) \equiv 1 - F_{\xi}(i_t) \leq \hat{\Lambda}(u_B). \quad (A186)$$

The thresholds for forming a partnership through an innovative contract, for forming a partnership through a standard contract, and for switching from an innovative to a standard contract are linked by

$$i_t \geq \bar{i}_{std}(u_B) \Leftrightarrow \xi(u_B, i_t, \beta, \sigma, \alpha) \leq \bar{i}(i_t, \beta, \sigma, \alpha). \quad (A187)$$

Outside of this region ($i_t < \bar{i}_{std}(u_B)$) a standard contract is not used even if it is available. In this region, the availability of a standard contract has two effects.

1. **It crowds out some innovative contracts for all**

$$i_t > \bar{i}_{std}(u_B) \Leftrightarrow \Lambda_{std}(i_t) < \hat{\Lambda}(u_B) \Leftrightarrow \xi(u_B, i_t, \beta, \sigma, \alpha) < \bar{i}(i_t, \beta, \sigma, \alpha). \quad (A188)$$

It crowds out all innovative contracts and stops the evolution of precedent if and only if

$$i_t \geq \bar{i}(1, \beta, \sigma, \alpha) \Leftrightarrow \bar{i}_{std}(i_t) \leq \Lambda(i_t, 1, \beta, \sigma, \alpha) \Leftrightarrow \bar{i}(i_t, \beta, \sigma, \alpha) = 1 \quad (A189)$$

There is a non-empty range of values $i_t \in [\bar{i}_{std}(u_B), \bar{i}(1, \beta, \sigma, \alpha)]$ for which the standard and innovative contracts coexist if and only if the buyer’s reservation value $u_B$ and adjudication frictions $\beta$, $\sigma$, and $\alpha$ are low enough that

$$\hat{\Lambda}(u_B) > \Lambda \left(F_{\xi}^{-1} \left(1 - \hat{\Lambda}(u_B)\right), 1, \beta, \sigma, \alpha\right) \Leftrightarrow \xi(1, \beta, \sigma, \alpha, u_B) < \bar{i}_{std}(u_B) < \bar{i}(1, \beta, \sigma, \alpha), \quad (A190)$$

which is implied by condition (A176) because $\Lambda(i_t, 1, \beta, \sigma, \alpha)$ is decreasing in $i_t$.

2. **It expands the static volume of contracting if and only if**

$$i_t < \bar{i}(0, \beta, \sigma, \alpha, u_B) \Leftrightarrow \Lambda(i_t, 0, \beta, \sigma, \alpha) > \hat{\Lambda}(u_B) \Leftrightarrow \xi(u_B, i_t, \beta, \sigma, \alpha) > 0. \quad (A191)$$

This occurs for a non-empty range of values $i_t \in [\bar{i}_{std}(u_B), \bar{i}(0, \beta, \sigma, \alpha, u_B)]$ if and only if the buyer’s reservation value $u_B$ and adjudication frictions $\beta$, $\sigma$, and $\alpha$ are
high enough that
\[
\Lambda \left( F_{\xi}^{-1} \left( 1 - \hat{\Lambda}(\underline{u}_B) \right), 0, \beta, \sigma, \alpha \right) \equiv \\
\int_{F_{\xi}^{-1}(1-\hat{\Lambda}(\underline{u}_B))}^{1} x dF_{\xi}(x) + \frac{\alpha \sigma}{1-\beta} \int_{0}^{F_{\xi}^{-1}(1-\hat{\Lambda}(\underline{u}_B))} x dF_{\xi}(x) > \hat{\Lambda}(\underline{u}_B)
\]
\[
\Leftrightarrow \hat{i}(0, \beta, \sigma, \alpha) < \hat{i}_{\text{Std}}(\underline{u}_B) < \hat{i}(0, \beta, \sigma, \alpha, \underline{u}_B). \quad (A192)
\]

Conditions (A176) and (A192) can be rewritten respectively
\[
\frac{\sigma}{1-\beta} \leq \frac{\hat{\Lambda}(\underline{u}_B) - \mathbb{E}\xi^2}{\mathbb{E}\xi - \mathbb{E}\xi^2} \quad (A193)
\]
and
\[
\frac{\alpha \sigma}{1-\beta} > \frac{\hat{\Lambda}(\underline{u}_B) - \int_{F_{\xi}^{-1}(1-\hat{\Lambda}(\underline{u}_B))}^{1} x dF_{\xi}(x)}{\mathbb{E}\xi - \int_{F_{\xi}^{-1}(1-\hat{\Lambda}(\underline{u}_B))}^{1} x dF_{\xi}(x)}.
\]

Using the definition of \( i_{\text{Std}}(\underline{u}_B) \equiv F_{\xi}^{-1} \left( 1 - \hat{\Lambda}(\underline{u}_B) \right) \Leftrightarrow \hat{\Lambda}(\underline{u}_B) \equiv 1 - F_{\xi}(i_{\text{Std}}(\underline{u}_B)) \), we can write these jointly as
\[
\frac{1 - F_{\xi}(i_{\text{Std}}(\underline{u}_B)) - \int_{i_{\text{Std}}(\underline{u}_B)}^{1} x dF_{\xi}(x)}{\mathbb{E}\xi - \int_{i_{\text{Std}}(\underline{u}_B)}^{1} x dF_{\xi}(x)} \leq \frac{\alpha \sigma}{1-\beta} \leq \frac{\sigma}{1-\beta} \leq \frac{1 - F_{\xi}(i_{\text{Std}}(\underline{u}_B)) - \mathbb{E}\xi^2}{\mathbb{E}\xi - \mathbb{E}\xi^2}. \quad (A195)
\]
These conditions define a non-empty range for \( \sigma/(1-\beta) \), given a large enough \( \alpha \), if and only if
\[
\left[ \int_{i_{\text{Std}}(\underline{u}_B)}^{1} x dF_{\xi}(x) - \mathbb{E}\xi^2 \right] \left[ 1 - F_{\xi}(i_{\text{Std}}(\underline{u}_B)) - \mathbb{E}\xi \right] < 0. \quad (A196)
\]
E.g., if \( \xi \sim U[0,1] \) then conditions (A176) and (A192) simplify to
\[
\left[ \frac{1 - i_{\text{Std}}(\underline{u}_B)}{i_{\text{Std}}(\underline{u}_B)} \right]^2 \leq \frac{\alpha \sigma}{1-\beta} \leq \frac{\sigma}{1-\beta} \leq 2 \left[ 2 - 3i_{\text{Std}}(\underline{u}_B) \right], \quad (A197)
\]
and the range is non-empty if and only if
\[
\frac{1}{2} < i_{\text{Std}}(\underline{u}_B) < \frac{\sqrt{3}}{3}. \quad (A198)
\]