Optimal Litigation Strategies with Observable Case Preparation*

Philippe Choné† and Laurent Linnemer‡

March 18, 2010

Abstract

This paper investigates the strategic effects of case preparation in litigation. Specifically, it shows how the pretrial efforts incurred by one party may alter its adversary's incentives to settle. We build a sequential game with one-sided asymmetric information where the informed party first decides to invest in case preparation, and the uninformed party then makes a settlement offer. Overinvestment, or bluff, always prevails in equilibrium: with positive probability, plaintiffs with weak cases take a chance on investing, and regret it in case of trial. Furthermore, due to the endogenous investment decision, the probability of trial may (locally) *decrease* with case strength. Overinvestment generates inefficient preparation costs, but may trigger more settlements, thereby reducing trial costs.

Keywords: Case preparation, Settlement, Trial, signaling.
***JEL* Classification: K410**

1 Introduction

The vast majority of tort disputes never reach a trial verdict. Litigants, indeed, have mutual incentives to save on trial costs by settling out of court. Moreover, a settlement shortens the dispute and might help to keep it confidential.[1] For example, out of the 98,786 tort cases that were terminated in U.S. district courts during fiscal years 2002 and 2003, 1,647 or 2% were decided by a bench or jury trial.[2] Data about settlement are most of the time not available but it is commonly believed that cases that go to trial involve larger damages.[3]

*We are grateful to seminar participants at the 2007 CESifo Area Conference on Applied Microeconomics, the 22nd Meeting of the European Economic Association in Budapest, the 2007 ASSET Meeting in Padova, as well as to Philippe Février and Thibaud Vergé for insightful comments.

†CREST (LEI), 15 bd Gabriel Péri 92245 Malakoff. Email : chone@ensae.fr

‡CREST (LEI), 15 bd Gabriel Péri 92245 Malakoff. Email : laurent.linnemer@ensae.fr (contact author)

[1] See Daughety and Reinganum (1999) for the issue of confidentiality.

[2] Source: Bureau of Justice Statistics Bulletin, August 2005, NCJ 208713.

[3] See Black, Silver, Hyman, and Sage (2005) and Chandra, Shantanu, and Seabury (2005). Kaplan, Sadka, and Silva-Mendez (2008) use a data set from labor tribunals in Mexico that provides information about settled cases as well as tried cases. They find that about 70% of lawsuits are settled, 15% dropped

The amount at stake in a settlement dispute can be very important: in March 2006 the Canadian firm Research In Motion who manufactures the Blackberry email device agreed to pay a \$612.5m settlement amount to end a patent dispute with NTP Inc. a little known Virginia firm.[4]

In this article, we examine how the incentives to settle are modified when litigants can enhance the strength of their case by investing in case preparation during the pretrial phase. We assume that pretrial efforts incurred by the parties can change the probability that the defendant will be found liable at trial and/or the damage awarded to the plaintiff should liability be established. The seminal contributions in the field, Bebchuk (1984) and Reinganum and Wilde (1986), assume that the expected award is fixed, but known to one party only. The former paper considers a screening game: the uninformed party (the defendant, say) makes a settlement offer, which is rejected by plaintiffs with strong cases. The latter paper studies a signaling game: the informed party makes an offer which positively depends on the strength of his case, and the defendant refuses to pay a larger settlement amount with a higher probability.[5]

With few exceptions, the subsequent literature has treated the expected award in court as exogenous. Litigants, however, do invest in case preparation with the purpose of improving their position at trial and, consequently, at the negotiation table. During the pretrial phase, the parties take various actions: getting additional evidence, taking thorough initial interviews and depositions, obtaining statements from witnesses, issuing interrogatories, selecting expert witnesses, etc. In practice, the precise form of pretrial preparation depends on the legal procedure in force.

To show how the investment in case preparation of one party can affect its adversary's incentives to settle, we build a sequential game, where the informed party first decides to invest, or not, in case preparation, and the uninformed litigant, after observing this decision, makes a take-it-or-leave-it settlement offer.[6] We assume that case preparation efforts entail a sunk cost, which is incurred during the pretrial phase, and that they are effective in raising or reducing the expected award (depending on the party who invests). Conditionally on the investment decision, litigants play a screening game with a continuum of types *à la* Bebchuk, leading to settlement or trial.[7] The endogenous investment decision, however, introduces a signaling dimension. The informed party can potentially use the investment to manipulate the other side's beliefs and alter her incentives to settle.

The observability assumption is critical as it is the basis of the signaling mechanism. Admittedly, a party may not observe the exact amount of resources devoted by her adversary to prepare his case. At the very least, however, the counsel chosen by a litigant to

---

and 15% go to trial. They find, however, that plaintiffs that go to trial receive significantly lower final payments. They explain this difference by a selection effect as workers who exaggerate their claims settle less often, and may be punished in terms of final-payment amounts.

[4]The settlement, which was not easy to reach, ended four years of legal dispute in the U.S. between the two companies. Maybe the largest amounts that make newspapers front pages correspond to drug related civil action trials but they do not necessarily lead to the largest amounts per plaintiff.

[5]See **?** and Daughety and Reinganum (2005) for comprehensive surveys.

[6]For a model with alternative offers, see Spier (1992).

[7]The informational asymmetry is one-sided. For models where both parties have private information, see Schweizer (1989) and Daughety and Reinganum (1994).

assist him during the pretrial phase is known to the other party as counsels have many opportunities to interact during this phase. The counsel choice is a good indicator of case preparation expenses. Lawyer's fees vary substantially from one lawyer to another according to experience and reputation. For example, the Laffey Matrix[8] allows an experienced federal court litigator to charge twice as much as a junior associate. Hiring a prominent law firm rather than an ordinary attorney is a major strategic decision, and this choice is public information before the settlement offers are made.[9]

To present our findings, we suppose, for convenience, that the informed party is an injured plaintiff, and the uninformed party a potentially negligent defendant. Case preparation raises the value of the claim, but entails a sunk cost. We assume that, under symmetric information, only plaintiffs with strong cases do invest. For low expected damage types, the costs of case preparation exceed its return. In other words, the case preparation technology is tailored for plaintiffs with large damages.

Under asymmetric information, low-damage plaintiffs mimic plaintiffs with more serious cases in the hope of a larger settlement offer. Such an incentive is well understood by the defendant and, when total trial costs are not too large, a complex equilibrium pattern stands out. Plaintiffs with strong cases, who invest in case preparation under symmetric information, maintain this choice under asymmetric information. Plaintiffs with weak cases, who do not invest under symmetric information, however, are made indifferent between investing or not, and randomize between both options. When the defendant observes that the plaintiff has invested, she herself randomizes between a high and a low settlement offer. When she observes no investment, she makes a deterministic low offer. Plaintiffs with strong cases reject all equilibrium settlement offers and proceed to trial. Plaintiffs with weak cases can be further distinguished with respect to their settlement strategy. Plaintiffs with very weak cases accept all equilibrium offers (whether they have invested or not), and earn an informational rent. Intermediate types settle if and only if they have invested and the defendant offers a large amount. That is, these types settle more often out of court if they invest than if they do not.

Overinvestment in case preparation is generic, and its extent is constant across equilibria. Investment by weak plaintiffs is tantamount to *bluff*: a weak plaintiff who invests knows that he will regret it, should he receive the low offer and go to court. Plaintiffs with intermediate types go to court with positive probability, and indeed regret to have invested when a trial takes place. Strong and very weak plaintiffs, on the contrary, never regret their decision.

Furthermore, our model predicts that the probability of trial can decrease with the strength of the case. This is in sharp contrast with both Bebchuk and Reinganum and Wilde

[8]A list of hourly rates (adapted each year to take into account inflation) for attorneys of varying experience levels prepared by the Civil Division of the United States Attorney's Office for the District of Columbia. This list is intended to be used in cases in which a fee-shifting statute permits the prevailing party to recover reasonable attorney's fees.

[9]Garoupa and Gomez-Pomar (2008) offer a number of explanations of why corporate clients acting as plaintiffs prefer to hire large law firms and resort to hourly fees arrangements. Our results may provide an extra rationale for such a policy: the reason why plaintiffs are ready to incur expensive attorney fees is that they use these sunk costs to signal the strength of their claims.

models, which predict that the probability of trial increases with the expected damages. Indeed, the more demanding the plaintiff, the less likely settlement occurs, otherwise all types of plaintiff would demand more. In our model, this logic fails because of a *selection effect*: the larger their expected damage, the larger their probability of investment and, in turn, the larger their probability of settlement. This strategic effect reduces trial costs, and may outweigh the socially inefficient increase in preparation sunk costs.

Finally, we examine the robustness of our results in a game where the investment in case preparation is continuous. As in the binary model, the probability of trial is driven by two forces that play in opposite directions: the selection effect counters the usual effect, tending to make the plaintiff with the strongest case settle *more often* than weaker types. We highlight the interplay of these opposite forces, but do not know whether the selection effect may dominate, as is the case in the binary model. We also show that the probability of bluff (an agent choosing the complete information effort of the highest type) is positive and constant across equilibria as in the binary game. We stress, however, a major difference between the two environments: All plaintiffs choosing the same effort level (full pooling) can never be an equilibrium when investment is continuous, while such an outcome occurs under a relatively mild condition in the binary game. Our analysis also shows that separation of types cannot occur in equilibrium and suggests that extensive and complicated randomization from both parties is necessary to sustain an equilibrium of the continuous game.

An important pretrial topic is discovery. Following Shavell (1989), the interaction between mandatory or voluntary information disclosure and settlement strategies has been thoroughly explored. Depending on the model, discovery completely[10] or partially [11] reveals the type of the plaintiff. In any case, as is widely recognized, discovery is costly, and litigants have incentives to settle their dispute before discovery costs are sunk. In this respect, discovery and trial share common features; in particular, both are costly proceedings that unveil information. In this paper, we abstract away from the strategic effects of discovery that are outlined below. Our framework, however, accommodates the presence of a discovery phase, provided that expected payoffs depend on pre-existing heterogeneity and case preparation efforts only.

Sobel (1989), Cooter and Rubinfeld (1994), Hay (1995), Schrag (1999), and Schwartz and Wickelgren (2008) have investigated settlement prior to discovery. Some of these papers consider litigation efforts, thereby endogenizing case strength. None of them, however, allows for *pre-discovery* case preparation, which is our focus here.

Hay (1995) considers pre-discovery settlement in a model with two types, but assumes that case preparation occurs when all exogenous heterogeneity concerning case strength has vanished. At the final stage of his game, plaintiffs differ only through their discovery effort, which is not observed by the defendant: there is hidden action, but symmetric information. If efforts were observed, intense case preparation would induce generous settlement offers. Since effort is unobserved, the plaintiff is tempted to shirk and, in equilibrium, both players

[10]See for instance Shavell (1989), Sobel (1989), Hay (1995), Farmer and Pecorino (2005), Schwartz and Wickelgren (2008). See also Daughety and Reinganum (2008) for a tradeoff between signaling and disclosure.
[11]Mnookin and Wilson (1998).

resort to mixed strategies. In contrast to Hay's framework, investment occurs in the present paper prior to discovery, and is observed by the defendant.

Schrag (1999) investigates the effects of regulating discovery efforts in a model with two types. If they fail to settle early, the litigants engage in simultaneous discovery efforts, the payoffs of which are modeled through reduced-form profit functions. Case strength is endogenous, but efforts are made after the strategic settlement phase under consideration. The settlement decisions alter the beliefs of the uninformed party and, in turn, her discovery efforts, which is anticipated at the time of the settlement bargaining. Schrag finds that the prospect of future discovery can undermine the incentives to settle early.

Schwartz and Wickelgren (2008) study the incentives of plaintiffs with negative expected value cases to file a suit. The problem of NEV cases has first been tackled in the seminal contributions of Nalebuff (1987) and Bebchuk (1988). Schwartz and Wickelgren examine pre-discovery settlement bargaining in a model *à la* Bebchuk. As in Schrag (1999), the beliefs of the uninformed party are revised following the early settlement bargaining. In contrast, in the present article, the Bayesian revision occurs after case preparation, which is the source of the bluffing strategy. In other words, we concentrate on case preparation investment as well as settlement offers made before any activities that uncover and spread information.

A methodological contribution of this article is the characterization, in a signaling game with a continuum of types, of semi-pooling equilibria where *both* sides play in mixed strategies. The informed party, playing first, makes a binary decision (preparing or not), and the uninformed party replies with a continuous strategic variable (the settlement offer). Despite of the multiplicity of equilibria, we are able to show that important economic features (in particular, the extent of overinvestment) are constant across equilibria. We also demonstrate that the equilibria involve non-degenerated mixed strategies of both players. As already said, the main modeling difference with Hay (1995) is the presence of private information. Another difference is the timing of the game: sequential in the present paper, simultaneous in Hay. In contrast to the signaling game of Reinganum and Wilde (1986), both players, in our framework, resort to mixed strategies. Specifically, in their paper, the uninformed defendant randomizes between accepting or rejecting the settlement offer made by the informed plaintiff, while, in our model, no randomization takes place once a settlement offer is made as it is the informed party who accepts or rejects the offer. Here, the defendant randomizes between a generous and a conservative offer when the plaintiff opts for case preparation which induces the intermediate plaintiff types to accept or reject the offer. As to the plaintiff, he randomizes between investing, or not, in case preparation, using a probability that is not necessarily monotonic in case strength. Yet our model predicts a simple average pattern: above a critical threshold for case strength, all types invest and proceed to trial. Below the threshold, the average probability of investment depends on the fundamentals of the game (sunk and trial costs and effectiveness of case preparation) in a simple manner.

The paper is organized as follows. Section 2 presents the model. Section 3 details the strategies of the parties and presents some preliminary results. Section 4 characterizes the unique equilibrium when trial costs are relatively large, while section 5 deals with the relatively low trial cost case. Section 6 presents comparative statics and qualitative properties of the equilibria. Section 7 investigates the game with continuous investment. Section 8 suggests alternative interpretations of the model and avenues for future research. Most proofs are in a technical appendix.

2 The model

We consider a litigation framework with one-sided informational asymmetry. The expected award, also referred to as "case strength", is the plaintiff's private information. The defendant only knows the distribution of case strength. The plaintiff can invest in preparation to enhance his case. We posit a multiplicative effect: the investment multiplies the expected award by a constant greater than one.[12] Both litigants are risk neutral.

2.1 The litigation game

The extensive form of the game is illustrated in Figure 1. Nature determines the plaintiff's type, noted x, according to a distribution F with positive density f on $[a, b]$. The plaintiff decides to invest in case preparation, which we note $e = H$, or to exert the basic level of effort, $e = L$. The investment involves a sunk (pretrial) cost $c > 0$. The defendant, after observing the investment decision, makes a take-it-or-leave-it settlement offer. The plaintiff either accepts or refuses the offer. In the latter case, the case proceeds to either discovery or trial. The plaintiff's expected gain is μx if he has invested, x otherwise. The parameter $\mu > 1$ is common knowledge. In other words, the return of case preparation is a higher expected award in the subsequent litigation.

In addition to the sunk cost, the initial investment may alter the plaintiff's *trial* costs. The choice of a reputable attorney in the initial phase may imply larger trial cost as it might be costly to switch to a less expensive lawyer who would have to start from scratch. The plaintiff's litigation costs are noted $t_H^P \geq t_L^P \geq 0$. For simplicity, we assume here that the plaintiff's case preparation does not impact the defendant's litigation cost, that we note t_D.[13] Total trial costs in each situation are denoted $T_L = t_L^P + t_D$ and $T_H = t_H^P + t_D$. Negative expected claims are assumed away: the expected award in court is greater than the trial's costs for both technologies, even for the weakest case: $a > t_L^P$ and $\mu a > t_H^P + c$.

Under symmetric information, litigants never go to court. The defendant observes x, the case strength, as well as the plaintiff's investment decision. Therefore, she holds all the necessary information to make personalized offers which are accepted. Formally, she offers $x - t_L^P$ after $e = L$, and $\mu x - t_H^P$ after $e = H$. The plaintiff accepts such an amount (but would refuse any smaller one) because this is exactly the expected reward he would have in case of a trial. Anticipating these settlement amounts, the plaintiff invests if and only if

[12] Such a multiplicative structure captures the specific case where private litigation expenditures influence the plaintiff's probability of winning but the damages conditional on winning are fixed.

[13] In an earlier draft Choné and Linnemer (2008), we allowed for the possibility that the plaintiff's investment makes the life of the opposite party harder, forcing her to incur higher litigation costs. The qualitative results are the same in this more general framework.

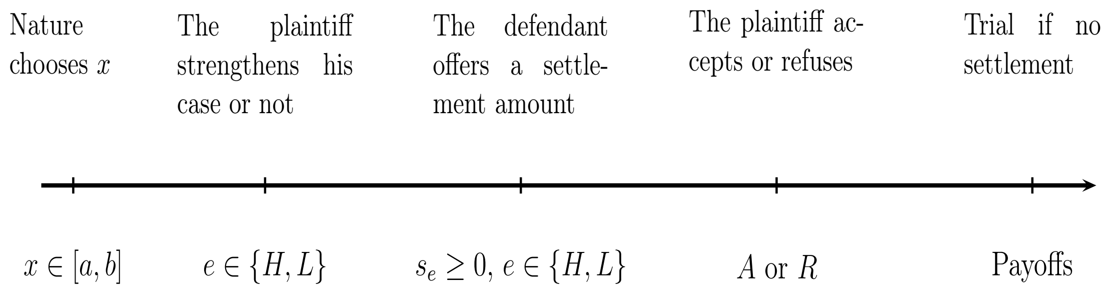

Figure 1: Timeline of the game

$\mu x - t_H^P - c \geq x - t_L^P$. Let $\widetilde{x}$ be the type of the plaintiff who is indifferent, under perfect information, between both technologies:

$$\widetilde{x} = \frac{t_H^P - t_L^P + c}{\mu - 1}.$$

We refer to the plaintiff $\widetilde{x}$ as the marginal type. Investment is optimal for plaintiffs with strong cases ($x > \widetilde{x}$), while it is not for weak cases ($x < \widetilde{x}$). Throughout, we assume that no technology is superior to the other for all types of plaintiff. Formally:

Assumption 1. *The marginal type is interior: $a < \widetilde{x} < b$.*

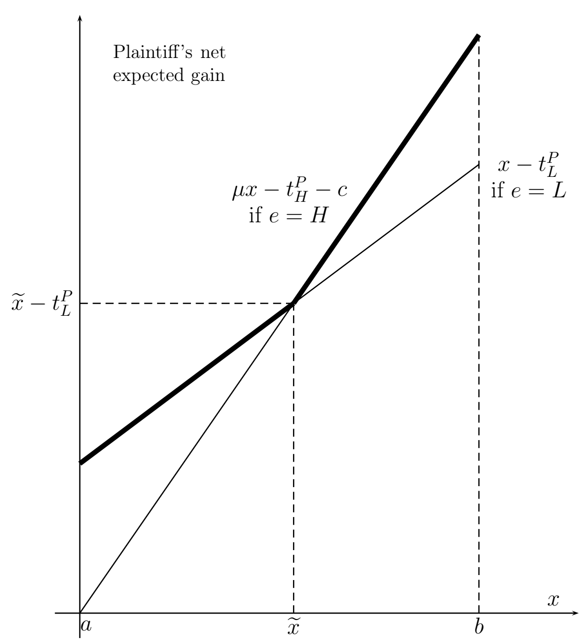

Figure 2: The optimal technology choices and the marginal type $\widetilde{x}$

In Figure 2, the bold line represents the plaintiff's symmetric information gain as a function of his type x. Under asymmetric information, this line corresponds to the minimum gain the plaintiff can secure by going to court. This gain is the plaintiff's reservation utility, which is type-dependent. We also assume that a plaintiff who has invested in case preparation does not want to switch back to the basic technology.

Assumption 2. *Once the sunk cost c has been incurred, a plaintiff has no incentives to give up the return of case preparation. Formally: $a - t_L^P < \mu a - t_H^P$.*

Combined with Assumption 1, Assumption 2 implies, for the weakest case: $\mu a - t_H^P - c < a - t_L^P < \mu a - t_H^P$, which entails a positive lower bound for the sunk cost: $c > (\mu a - t_H^P) - (a - t_L^P) > 0$. Notice that assumption 2 is satisfied when $t_L^P = t_H^P$.

2.2 The one-technology benchmarks

Throughout, we note $\{H\}$ and $\{L\}$ the situations where only one technology is available, and $\{HL\}$ the situation where the plaintiff can choose his preferred technology. Following Bebchuk, we first examine the benchmark cases $\{H\}$ and $\{L\}$.

A settlement offer partitions the population of plaintiffs into two groups. In case $\{L\}$, the plaintiff of type x accepts a settlement offer s if and only if $x \leq s + t_L^P$. The corresponding threshold in case $\{H\}$ is $(s + t_H^P)/\mu$. It is convenient to parameterize settlement offers with the type of the indifferent plaintiff, rather than with the settlement amount itself. The offer leaving plaintiff x indifferent yields the following utility to plaintiff y:

$$v_{\{L\}}(y; x) = \max(y - t_L^P, x - t_L^P) \quad \text{and} \quad v_{\{H\}}(y; x) = \max(\mu y - t_H^P - c, \mu x - t_H^P - c), \quad (1)$$

and the following profit to the defendant: $\pi_{\{L\}}(x) = -(x - T_L) F(x) - \int_x^b y f(y) \mathrm{d}y - t_D$ and $\pi_{\{H\}}(x) = \mu \left[-\left(x - \frac{T_H}{\mu} \right) F(x) - \int_x^b y f(y) \mathrm{d}y \right] - t_D$. The latter formulae express the defendant's tradeoff between rent extraction and trial cost savings. Throughout the paper, we maintain the following assumption.

Assumption 3. *The distribution of case strength is strictly log-concave.*

Assumption 3 amounts to $\tau = F/f$ being increasing on $[a, b]$ and guarantees that the profit functions $\pi_{\{L\}}$ and $\pi_{\{H\}}$ attain their maximum at a unique value.[14] We denote the optimal offers by x_L^* and x_H^*, and assume they are interior ($x_H^*, x_L^* \in (a, b)$). We have:

$$\tau(x_H^*) = T_H/\mu \quad \text{and} \quad \tau(x_L^*) = T_L.$$

If case strength is uniformly distributed on $[a, b]$, then $\tau(x) = x - a$, $x_L^* = a + T_L$, $x_H^* = a + T_H/\mu$. By assumption, $T_L \leq T_H$, but T_H/μ could be either larger or smaller than T_L and there is *a priori* no restriction on the ordering of x_H^* and x_L^*.

[14] In the three situations $\{H\}$, $\{L\}$ and $\{HL\}$, existence results only require τ to be nondecreasing, but uniqueness results depend on τ increasing. In particular, the strict monotonicity guarantees the uniqueness of x_H^* and x_L^*.

In such a litigation environment, total welfare equals the opposite of litigation costs, and is not affected by transfers from one party to another.[15] The expected litigation costs in equilibrium when only one technology is available are given by

$$\mathcal{C}^*_{\{H\}} = c + (1 - F(x^*_H))T_H \quad \text{and} \quad \mathcal{C}^*_{\{L\}} = (1 - F(x^*_L))T_L. \tag{2}$$

The comparison of the expected trial costs in the situations $\{H\}$ and $\{L\}$ involves a direct cost effect and a strategic effect. Formally,

$$\mathcal{C}^*_{\{H\}} - \mathcal{C}^*_{\{L\}} = c + (1 - F(x^*_L))(T_H - T_L) - (F(x^*_H) - F(x^*_L))T_H.$$

Since $c \geq 0$ and $T_H \geq T_L$, the sum of the first two terms is positive, and tends to make $\mathcal{C}^*_{\{H\}}$ higher than $\mathcal{C}^*_{\{L\}}$ (direct cost effect). The last term reflects the change in the incentives to settle. If $T_H/\mu \leq T_L$, trial occurs less often in $\{L\}$ than in $\{H\}$, so both effects play in the same direction. This happens, in particular, when $T_L = T_H$. On the other hand, if T_H/μ is larger than T_L, the strategic effect tends to make $\mathcal{C}^*_{\{H\}}$ lower than $\mathcal{C}^*_{\{L\}}$. When x^*_H tends to b, the strategic effect may dominate the direct cost effect.

3 Notations and Preliminary results

We now examine the incentives to invest and to settle in the situation $\{HL\}$ where both technologies are available. As will shortly become clear, we must consider mixed strategies of the defendant. Parameterizing settlement offers with the type of the indifferent plaintiff as explained above, the most general defendant's strategy is represented by a pair (P_H, P_L) of probability measures on the interval $[a, b]$. Facing $e = H$ (resp. $e = L$), the defendant randomizes across offers $\mu x - t^P_H$ (resp. $x - t^P_L$), where x is drawn in $[a, b]$ according to the distribution P_H (resp. P_L).

The defendant's strategy and the plaintiff's payoffs: Facing a defendant's strategy (P_H, P_L), a plaintiff of type y gets the following expected payoffs:

$$v_H(y) = \int_a^b v_{\{H\}}(y; x) dP_H(x) \quad \text{and} \quad v_L(y) = \int_a^b v_{\{L\}}(y; x) dP_L(x), \tag{3}$$

where the base utility functions $v_{\{H\}}(.; x)$ and $v_{\{L\}}(.; x)$ are defined in (1). In Appendix, Lemma A.1 shows that the functions v_L and v_H are nondecreasing and convex and that their derivatives are linked to the probability of trial (conditional on effort) by: $P_H(x \leq y) = v'_H(y)/\mu$ and $P_L(x \leq y) = v'_L(y)$. Furthermore, Lemma A.1 shows that we can interchangeably use the probability measures P_H and P_L or the expected payoff functions $v_H(.)$ and $v_L(.)$ of the plaintiff to represent the defendant's strategy. This result plays a critical role in the following analysis, where the geometric properties of v_H and v_L are extensively used.

[15] The focus of the paper is on the settlement issue. Therefore we do not take into account the administrative costs of a trial nor the deterrence effects that trial and settlement cost might have.

The plaintiff's strategy and the defendant's payoffs: A plaintiff's behavioral strategy is represented by a map $\sigma : [a,b] \to [0,1]$, which specifies the probability $\sigma(x)$ that the plaintiff of type x invests in case preparation. After observing the plaintiff's decision $e \in \{H, L\}$, the defendant revises her beliefs about the distribution of case strength. For a given plaintiff's strategy σ, we note f_e and F_e, the density and c.d.f. of the defendant's posterior distributions. Assuming that both technologies are used in equilibrium, the revised densities are given by the Bayes' rule

$$f_H(x) = \frac{\sigma(x)f(x)}{\int_a^b \sigma(t)f(t)dt} \quad \text{and} \quad f_L(x) = \frac{[1-\sigma(x)]\,f(x)}{\int_a^b [1-\sigma(t)]\,f(t)dt}. \tag{4}$$

Consequently, conditional on e, an offer x yields the following expected revenues

$$\pi_H(x) = \mu\left[-\left(x - \frac{T_H}{\mu}\right)F_H(x) - \int_x^b t f_H(t)dt\right] - t_D \tag{5}$$

$$\pi_L(x) = -(x - T_L)\,F_L(x) - \int_x^b t f_L(t)dt - t_D \tag{6}$$

to the defendant after $e = H$ or $e = L$. The difference between the above expressions of π_e and the expression of $\pi_{\{e\}}$ used in the one-technology worlds of Section 2.2 is the underlying distributions of heterogeneity: the profits π_e refer to the posterior distributions, while the prior distributions of case strength are used in $\pi_{\{e\}}$.

If the defendant randomizes across offers according to the probability distribution P_e, her payoff is

$$\Pi_e = \int_a^b \pi_e(x)\mathrm{d}P_e(x).$$

Lemma B.1 in the technical appendix shows how the defendant's profits can be expressed in terms of the utility she leaves to the plaintiff.

A perfect Bayesian equilibrium of the game is a function $\sigma^*(.)$ and two probability measures P_H^* and P_L^* on $[a,b]$ such that (i) given P_H^* and P_L^*, $\sigma^*(x)$ maximizes the expected payoff of type x, for all x in $[a,b]$; (ii) given σ^*, P_e^* maximizes the defendant expected payoff after she has observed e, for $e = H, L$; (iii) beliefs are updated according to Bayes' rule (4).

Throughout, we assume that the the plaintiff's strategy σ has a left and a right limit at any point x, which are noted $\sigma(x^-)$ and $\sigma(x^+)$ respectively.[16] It follows that the posterior densities f_H and f_L have left and right limits at any point x and that the defendant's payoff function π_H and π_L have a left and a right derivative at any point x; the right derivatives are given by $\pi_H'(x^+)/\mu = -F_H(x) + \frac{T_H}{\mu} f_H(x^+)$ and $\pi_L'(x^+) = -F_L(x) + T_L f_L(x^+)$ (the left derivatives are given by analog formulae). It is useful to observe that $\pi_H'(x)$ and $\pi_L'(x)$ are respectively equal, up to a positive multiplicative constant, to $\sigma(x)f(x) - \int_a^x \sigma\mathrm{d}F$ and $(1-\sigma(x))f(x) - \int_a^x (1-\sigma)\mathrm{d}F$.

[16] More precisely, we assume that σ is a "function of bounded variation", i.e. that it can be written as the difference between two nondecreasing functions. Such a function everywhere admits a right and a left limit, which coincide except possibly at countably many points. See Ziemer (1989).

General properties of equilibria:

The *support* of a mixed strategy is the set of pure strategies to which a positive probability is assigned. For an offer to be in the support, a necessary condition is that it maximizes the defendant expected payoff. Whenever the maximum of her payoff is attained at many points, she can randomize across several of the corresponding offers. Formally:

$$\text{supp } P_e^* \subset \text{ argmax } \pi_e,$$

where π_e, given by (5) or (6), uses the updated beliefs. If, for $e = H$ or L, the defendant's payoff function attains its maximum at a unique point, then she makes the corresponding offer with probability 1: the support of distribution P_e is thus a singleton, or, equivalently, P_e is a mass point. (As seen above, this happens if all x opt for the same technology.)

We now show that the investment decision of plaintiffs with strong cases is never distorted.

Lemma 1. *In equilibrium, plaintiffs with strong cases invest in case preparation.*

Proof. The result, obviously, holds when the overall probability that $e = H$, $\int_a^b \sigma \mathrm{d}F$, is 1. We concentrate, therefore, on the case where the overall probability of observing the basic technology, $\int_a^b (1 - \sigma)\mathrm{d}F$, is positive, and prove that, under this assumption, the defendant makes no offer greater than $\widetilde{x} - t_L^P$ after observing $e = L$, that is, $v_L(\widetilde{x}) = \widetilde{x} - t_L^P$.

To this aim, we use a standard unraveling argument. Suppose that $v_L(\widetilde{x}) > \widetilde{x} - t_L^P = \mu\widetilde{x} - t_H^P - c$. Since $v_L(b) = b - t_L^P < \mu b - t_H^P - c$, the curve v_L must cross the segment $\mu x - t_H^P - c$ on $(\widetilde{x}, b]$. They only cross once since the segment has slope $\mu > 1$ and v_L has slope no greater than 1. Let x_0 be the unique intersection point. For $x > x_0$, we have: $v_H > v_L$, so the plaintiff chooses $e = H$ with probability $\sigma = 1$. The defendant therefore knows that all the plaintiffs who choose $e = L$ necessarily have: $x \leq x_0$. Therefore she could reduce the utility $v_L(.)$ by the constant amount $z = v_L(x_0) - [x_0 - t_L^P] > 0$. Such a change would increase her payoff by $z \int_a^{x_0} (1 - \sigma)\mathrm{d}F$, which is positive by assumption. We conclude that we must have: $v_L(\widetilde{x}) = \widetilde{x} - t_L^P$, which is equivalent to saying that the support of P_L is a subset of $[a, \widetilde{x}]$ or that the defendant, after observing $e = L$, does not make any offer greater than $\widetilde{x} - t_L^P$ with positive probability.

It follows that, for $x > \widetilde{x}$, we have: $v_L(x) = x - t_L^P < \mu x - t_H^P - c$, so plaintiff x chooses to invest.

□

4 Equilibria when trial costs are large

This section is devoted to the case $\widetilde{x} < x_H^*$ or, equivalently, $\tau(\widetilde{x}) < T_H/\mu$. This assumption expresses that, in the benchmark situation $\{H\}$, the marginal plaintiff $\widetilde{x}$ settles in equilibrium. The next proposition shows that, under this assumption, the only possible equilibrium configuration when both technologies are available is the same as in $\{H\}$.

Proposition 1. *Assume $\widetilde{x} < x_H^*$. Then, in equilibrium, any plaintiff invests in case preparation: $\sigma^* = 1$ on $[a, b]$. After observing $e = H$, the defendant offers $\mu x_H^* - t_H^P$ to settle the case.*

Proof. First, we prove the existence of out-of-equilibrium beliefs that are consistent with this configuration. Suppose that, after she observes $e = L$, the defendant believes that the deviation comes from plaintiff a, and, accordingly, offers only $a - t_L^P$ to settle. It follows that a plaintiff of type x receives utility $x - t_L^P$ if he does not invest, while he gets $v_{\{H\}}(x; x_H^*) > x - t_L^P$ if he invests (see Figure 3). In turn, any plaintiff invests and, as seen in Section 2.2, it is indeed optimal for the defendant to offer $\mu x_H^* - t_H^P$, leaving plaintiff x_H^* indifferent between accepting or rejecting the offer.

Second, we check that the above configuration is the only possible one in equilibrium. The proof proceeds by contradiction. Assuming that the investment probability $\Pr(e = H)$ is smaller than 1, we use the convexity of the functions v_H and v_L to show that there must exist $x_1 \in (a, \widetilde{x}]$ such that, after observing $e = H$, the defendant makes the offer $\mu x_1 - t_H^P$ with positive probability, and we show that this is not possible given the assumption $\widetilde{x} < x_H^*$. The detailed proof is presented in Appendix C.

$\square$

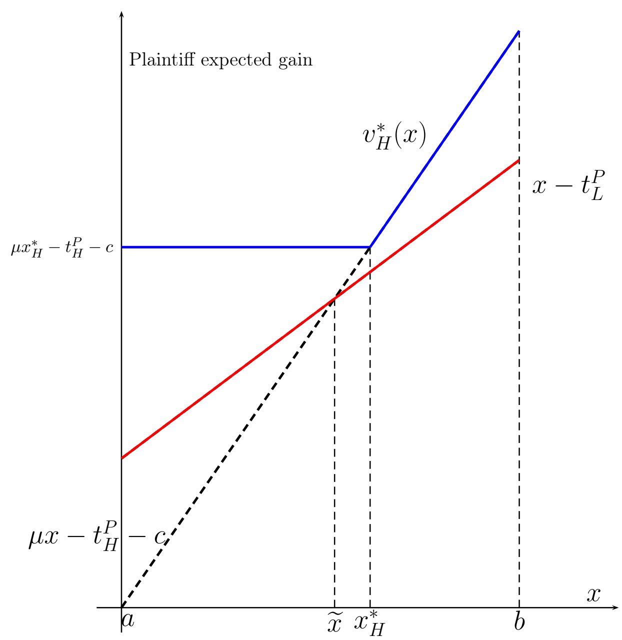

Figure 3: Equilibrium gains when $x_H^* > \widetilde{x}$

According to Proposition 1, the plaintiff's strategy, as well as the defendant's strategy after she has observed $e = H$, are unique in equilibrium. Any distribution P_L such that $v_L \leq v_H^*$ sustains the equilibrium. But the equilibrium configuration is unique, and is the same as in $\{H\}$. The uniqueness result is the heart of Proposition 1.

The intuition behind this result is fairly simple. When T_H/μ is large enough, the trial is relatively costly to at least one party and the plaintiff and/or the defendant are eager to settle. Therefore a relatively high settlement offer is made after $e = H$ which attracts all types of plaintiff.

The absence of the basic technology in equilibrium can harm some types of plaintiff. Indeed, assume that x_L^* is larger than x_H^* and such that $x_L^* - t_L^P$ is larger than $\mu x_H^* - t_H^P - c$ (more precisely, we are concerned by the case: $\widetilde{x} < x_H^* < x_H^* + (\mu - 1)(x_H^* - \widetilde{x}) < x_L^* < b$). Then all plaintiffs who settle would prefer to do it for $x_L^* - t_L^P$ rather than $\mu x_H^* - t_H^P - c$. Yet the offer $x_L^* - t_L^P$ would only be made if all types selected $e = L$, which is not an equilibrium.

Finally, and more importantly, asymmetric information induces overinvestment. Plaintiffs with weak cases ($x < \widetilde{x}$) do not invest when information is symmetric, while they do when it is asymmetric. This choice is rewarding as they earn an informational rent. Part of the high-types (those who settle: $\widetilde{x} < x < x_H^*$) also benefit from asymmetric information.

5 Equilibria when trial costs are low

We now turn to the complementary case $x_H^* < \widetilde{x}$. Under this condition, assume that all types decide to invest in case preparation. Then the defendant makes the offer $\mu x_H^* - t_H^P$, which is rejected by all types above x_H^*. But the plaintiffs whose type lie between x_H^* and $\widetilde{x}$ are better off, in court, with $e = L$ rather than with $e = H$ (see Figure 4). Therefore, it can no longer be an equilibrium for all types of plaintiff to invest.

We know from Lemma 1 that high types invest in case preparation. The following proposition goes a step further towards the characterization of the equilibrium in characterizing partially the plaintiff's strategy and completely the defendant's one.

Proposition 2. *Assume that $x_H^* < \widetilde{x}$. (i) In equilibrium, plaintiffs with strong cases ($x > \widetilde{x}$) invest and go to court; plaintiffs with weak cases ($x \leq \widetilde{x}$) are indifferent between investing or not. (ii) After $e = L$, the defendant makes a single offer, $\widehat{x} - t_L^P$, where $\widehat{x}$ lies between x_H^* and x_L^* and $\widehat{x} < \widetilde{x}$; after $e = H$, she offers $\mu\widehat{x} - t_H^P$ with probability $1/\mu$ and $\mu\widetilde{x} - t_H^P$ with probability $1 - 1/\mu$.*

Proof. To prove (i), we first observe that $v_L(\widetilde{x}) = \widetilde{x} - t_L^P$ (see the proof of Lemma 1). In Appendix D, we show, by using the convexity of the functions v_H and v_L, that the defendant, after observing $e = H$, makes no offer greater than $\mu\widetilde{x} - t_H^P$, which is equivalent to $v_H(\widetilde{x}) = v_L(\widetilde{x})$. It follows that $v_H(x) = \mu x - t_H^P - c > v_L(x) = x - t_L^P$ for $x > \widetilde{x}$, and that plaintiffs with strong cases invest and go to court.

We now turn to plaintiffs with weak cases, and show that $v_L = v_H$ on $[a, \widetilde{x}]$. We proceed by contradiction. Suppose that there exists $x < \widetilde{x}$ such that $v_L(x) \neq v_H(x)$, say, for instance, $v_L(x) < v_H(x)$. Since $v_L = v_H$ at $\widetilde{x}$, there exists x_1, with $x < x_1 \leq \widetilde{x}$ such that $v_L = v_H$ at x_1 and $v_L < v_H$ on $[x, x_1)$. We have $v_L < v_H$, $\sigma = 1$ and $f_L = 0$ on $[x, x_1)$.

Suppose first that $F_L(x) = 0$. Applying Lemma 3, we conclude that v_L is constant on $[a, x]$. But this is impossible as $v_H(x_1) = v_L(x_1)$, $v_H > v_L$ on a left neighborhood of x_1, and v_H is nondecreasing.

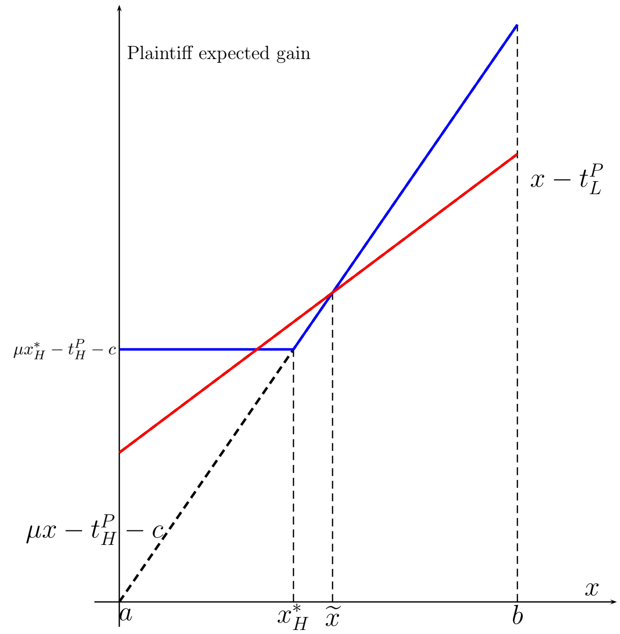

Figure 4: When $x_H^* < \widetilde{x}$, all types choosing $e = H$ is no longer an equilibrium

Next, suppose that $F_L(x)$ is positive. In this case, we have: $\pi'_L = -F_L + T_L f_L = -F_L < 0$ on $[x, x_1)$, so π_L does not attain its maximum in this interval, which, therefore, does not intersect the support of P_L. Applying Lemma 4, we conclude that v_L is affine on $[x, x_1)$. More generally, the argument shows that v_L is affine as long as it is below v_H and $F_L > 0$. Since v_H is convex and $v_H(x_1) = v_L(x_1)$, the only possibility is that $v_L < v_H$ and $\sigma = 1$, on the whole interval $[a, x_1]$, which contradicts $F_L(x) > 0$. It follows that $v_H \leq v_L$ on $[a, \widetilde{x}]$. The proof of $v_L \leq v_H$ is symmetric, which yields point (i) of the Proposition.

The key result in part (ii) of the Proposition is that the support of P_L must be a singleton. If the plaintiff has not invested, the defendant does not randomize.[17] In the other case, she randomizes between exactly two offers. Formally, if P_L is the singleton $\{\widehat{x}\}$, then the support of P_H must be the pair $\{\widehat{x}, \widetilde{x}\}$. The weights of $\widehat{x}$ and $\widetilde{x}$ are $1/\mu$ and $1 - 1/\mu$ respectively. Details can be found in Appendix E. □

Figure 5 shows the plaintiff's net gain in equilibrium, denoted by v^*, when $x_H^* < \widetilde{x}$. Point (i) of the Proposition 2 implies that weak plaintiffs ($x < \widetilde{x}$) gets the same expected utility irrespective of their investment decision: $v^* = v_H^* = v_L^*$. If $a \leq x \leq \widehat{x}$, the net expected payoff of the plaintiff of type x is $\widehat{x} - t_L^P$. It becomes $x - t_L^P$ when $\widehat{x} \leq x \leq \widetilde{x}$, and finally $\mu x - t_H^P - c$ for $\widetilde{x} \leq x \leq b$. Plaintiffs with very weak cases ($x < \widehat{x}$) never go to court, and earn an informational rent compared to the symmetric information case (see Figure 2). In contrast, types above $\widehat{x}$ have the same expected payoff under symmetric and

[17]This property follows from the log-concavity of the distribution of types, see section E.3.

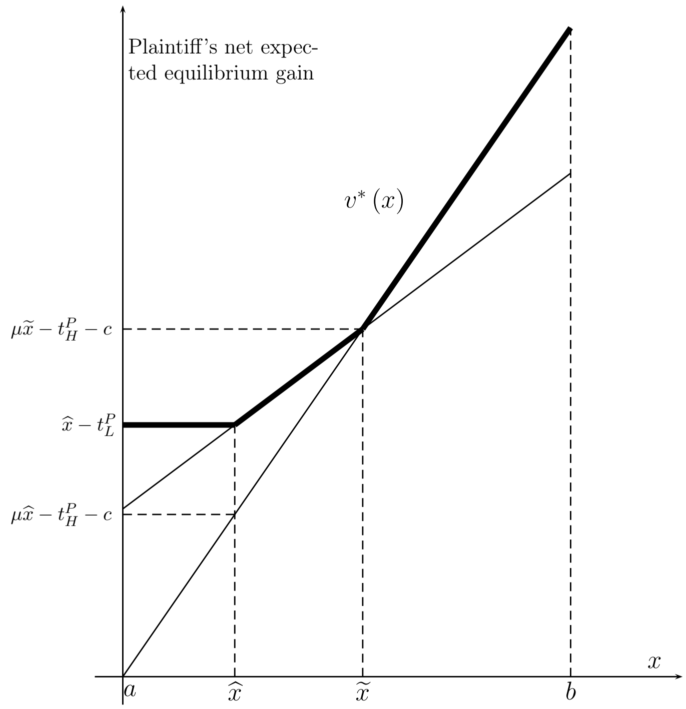

Figure 5: Equilibrium plaintiff's utility when $x_H^* < \widetilde{x}$

asymmetric information. Plaintiffs with intermediate cases ($\widehat{x} \leq x \leq \widetilde{x}$) go to court with probability one if they do not invest, with probability $1/\mu$ if they do. Plaintiffs with strong cases invest and go to court. We now exhibit a fully specified equilibrium (σ^*, P_H, P_L).

Proposition 3. *Assume that $x_H^* < \widetilde{x}$. If $T_L = T_H/\mu$, we set $\widehat{x} = x_L^* = x_H^*$, otherwise we define $\widehat{x}$ as the highest root to the equation*

$$\frac{f(\widetilde{x})}{f(x)} \exp\left[\frac{\mu}{T_H}(x_H^* - x)\right] = \frac{\tau(x) - T_L}{T_H/\mu - T_L}.$$

We define the defendant's strategy (P_H, P_L) as in Proposition 2. We define the plaintiff's strategy as follows.

$$\textit{If } T_L \geq T_H/\mu\textit{:} \quad \sigma^*(x) = \begin{cases} (f(\widetilde{x})/f(x_H^*)) \exp\left[\dfrac{\mu}{T_H}(x_H^* - \widetilde{x})\right] & \textit{for } a \leq x \leq x_H^* \\ (f(\widetilde{x})/f(x)) \exp\left[\dfrac{\mu}{T_H}(x - \widetilde{x})\right] & \textit{for } x_H^* \leq x \leq \widetilde{x} \\ 1 & \textit{for } \widetilde{x} \leq x \leq b \end{cases}$$

$$\text{and if } T_L < T_H/\mu\text{:}\quad \sigma^*(x) = \begin{cases} 1-(1-\sigma(\widehat{x}))\,(f(\widehat{x})/f(x_L^*))\exp\left[\dfrac{1}{T_L}(x_L^*-\widehat{x})\right] & \text{for } a \le x \le x_L^* \\ 1-(1-\sigma(\widehat{x}))\,(f(\widehat{x})/f(x))\exp\left[\dfrac{1}{T_L}(x-\widehat{x})\right] & \text{for } x_L^* \le x \le \widehat{x}, \\ (f(\widetilde{x})/f(x))\exp\left[\frac{\mu}{T_H}(x-\widetilde{x})\right] & \text{for } \widehat{x} \le x \le \widetilde{x}, \\ 1 & \text{for } \widetilde{x} \le x \le b, \end{cases}$$

then (σ^*, P_H, P_L) *is an equilibrium.*

The equilibrium of Proposition 3 is represented on Figures 6a and 6b. The formal check that the given strategies form an equilibrium is relegated in Appendix G.

The logic behind the shape of $\sigma^*(.)$ is the following. As stated in Proposition 2, the defendant makes a unique offer, $\widehat{x}$, after $e = L$, and she randomizes between $\widehat{x}$ and $\widetilde{x}$ after $e = H$. Accordingly, π_L has to attain its maximum at $x = \widehat{x}$, and π_H at both $\widehat{x}$ and $\widetilde{x}$. The function σ^* of Proposition 3 is such that π_H is constant between $\widehat{x}$ and $\widetilde{x}$. In Appendix G, we check that π_L is maximal at $\widehat{x}$ on the interval $[\widehat{x}, b]$. On $[a, \widehat{x}]$, $\sigma^*(.)$ is such that π_H and π_L are nondecreasing and such that enough low types choose to invest, in accordance with Equation (7) below. When $T_L \ge T_H/\mu$ (Figure 6a), this can be achieved by maintaining π_H flat between x_H^* and $\widehat{x}$, and keeping $\sigma^*(.)$ constant between a and x_H^*. When $T_L < T_H/\mu$ (Figure 6b), this choice of σ^* below $\widehat{x}$ is not consistent with the equilibrium requirements, as it would not give enough weight to $e = H$ and Equation (7) would be violated. As a consequence, $\sigma^*(.)$ has to decrease between a and $\widehat{x}$. A way to satisfy Equation (7) is to choose $\sigma^*(.)$ between x_L^* and $\widehat{x}$ such that π_L is flat and then to maintain $\sigma^*(.)$ constant between a and x_L^*. In Appendix G, it is checked that, in these circumstances, π_H is increasing on $[a, \widehat{x}]$.

Proposition 3 derives an equilibrium plaintiff's strategy σ. Such a strategy need not be unique, implying a multiplicity of equilibria. However, all equilibria share an important common feature. The fraction of weak plaintiffs ($x < \widetilde{x}$) who invest in case preparation is constant across equilibria. Investment by weak plaintiffs is tantamount to bluff: a weak plaintiff who invests is certain to regret his choice should he receive the low offer and proceed to court.

Proposition 4. *The fraction of plaintiffs with weak cases* ($x < \widetilde{x}$) *who invest in case preparation is given by*

$$\Pr(e = H \,|x < \widetilde{x}\,) = \frac{T_H}{\mu}\frac{f(\widetilde{x})}{F(\widetilde{x})}. \tag{7}$$

Proof. Since $\widetilde{x}$ belongs to the support of P_H, Lemma 5 implies that $\sigma(\widetilde{x}^+) \le \sigma(\widetilde{x}^-)$. As $\sigma(\widetilde{x}^+) = 1$ it means that $\sigma(.)$ is continuous at $\widetilde{x}$, with $\sigma(\widetilde{x}) = 1$. Therefore, the function π_H is differentiable at $\widetilde{x}$; since π_H is maximal at $\widetilde{x}$, we have $\pi'_H(\widetilde{x}) = 0$, which writes:

$$F_H(\widetilde{x}) = \frac{T_H}{\mu} f_H(\widetilde{x}) = \frac{T_H}{\mu}\frac{f(\widetilde{x})}{\Pr(e = H)},$$

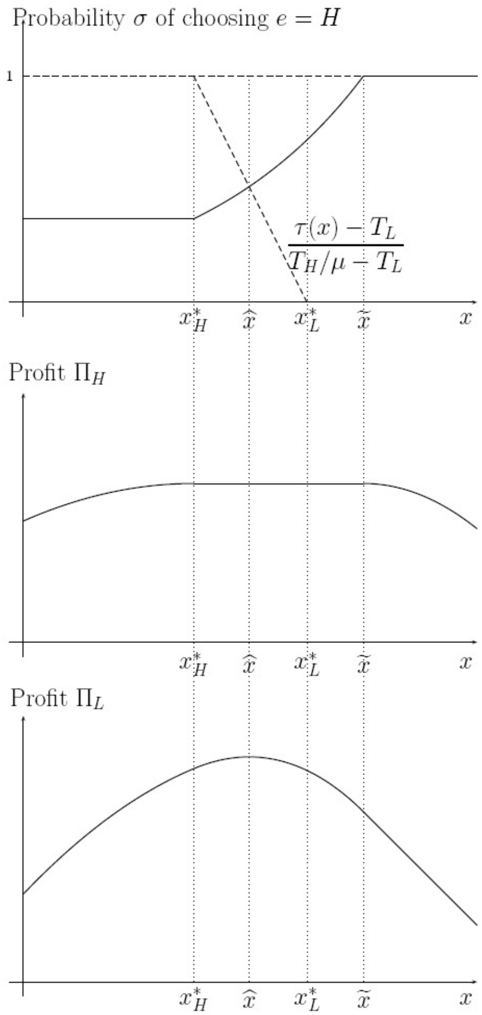

Figure 6a: $x_H^* < \widetilde{x}$ and $x_H^* < x_L^*$

Parameter values: F uniform on $[1, 3]$, $t_D = t_D = 0$, $t_H^P = 1.5$, $t_L^P = 1.25$, $c = 2.25$, and $\mu = 2$.

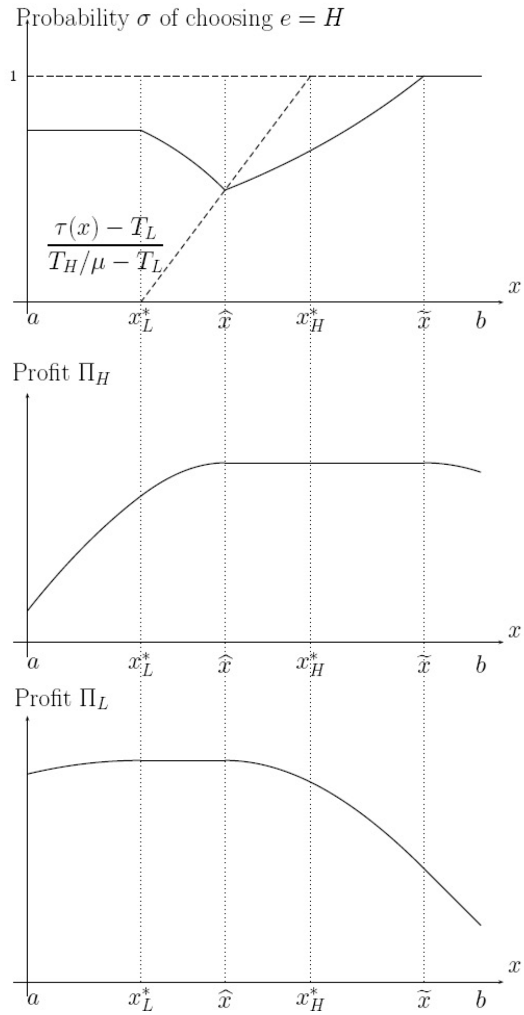

Figure 6b: $x_L^* < x_H^* < \widetilde{x}$

Parameter values: F uniform on $[1, 3]$, $t_D = t_D = 0$, $t_H^P = 2.5$, $t_L^P = 0.5$, $c = 0.75$, and $\mu = 2$.

which, combined with

$$\Pr(e = H \,|x < \widetilde{x}) = \frac{1}{F(\widetilde{x})} \int_a^{\widetilde{x}} \sigma(x) f(x) \mathrm{d}x = \frac{1}{F(\widetilde{x})} \Pr(e = H) F_H(\widetilde{x}),$$

yields (7). □

In the equilibrium of Proposition 3, all plaintiffs with weak cases ($x < \widetilde{x}$) resort to a mixed strategy. Obviously, one can construct equilibria where many low types play in pure strategy, provided that the probability of investment conditional on $x < \widetilde{x}$, given (7), is not affected. However, mixed strategies cannot not be entirely ruled out, as proved in

Appendix E.4, because the mixed strategy of the plaintiff of type $\widehat{x}$ must be non degenerate: $0 < \sigma^*(\widehat{x}) < 1$. As $\widehat{x}$ belongs to the supports of both P_L and P_H, σ^* is continuous at $\widehat{x}$ (Lemma 5), therefore a positive mass of plaintiffs plays in mixed strategy.[18]

6 Discussion

We now present qualitative properties of the equilibria and comparative statics results that illustrate the strategic effects at work. First, we show how the introduction of a publicly observable investment decision by the plaintiff alters Bebchuk's equilibrium pattern. Second, we examine the extent of overinvestment by plaintiffs with weak cases, and explain how it varies with the primitives of the model. Third, we study the welfare effects of introducing a new technology, starting from a one-technology world.

6.1 The trial probability may locally decrease with case strength

Figures 7a and 7b plot the probabilities of trial conditional on the investment choice as functions of the case strength x. As stated in Lemma A.1 (see Appendix A), these probabilities are nondecreasing in x. A plaintiff who expects large damages in court is less likely to settle. Yet Figure 7c shows that this property does not extend to the unconditional trial probability: a plaintiff with a stronger case can settle more often than a plaintiff with a weaker case.

Proposition 5. *Assume $x_H^* < \widetilde{x}$. At the equilibrium of Proposition 3, the unconditional trial probability decreases with case strength on $[x_H^*, \widetilde{x}]$.*

Proof. First, we prove that the investment probability σ^* given in Proposition 3 increases with case strength on $[x_H^*, \widetilde{x}]$. For $x_H^* \leq x_1 < x_2 \leq \widetilde{x}$, log-concavity yields

$$\ln F(x_2) \leq \ln F(x_1) + \frac{f(x_1)}{F(x_1)}(x_2 - x_1) \leq \ln F(x_1) + \mu/T_H(x_2 - x_1).$$

Taking the exponential and using $f(x_2)/f(x_1) \leq F(x_2)/F(x_1)$ yields $\sigma^*(x_2) \geq \sigma^*(x_1)$. It follows that the unconditional probability of trial, $1 - \sigma^*(x) + \sigma^*(x)/\mu$, decreases with case strength on $[x_H^*, \widetilde{x}]$. □

This result is in sharp contrast with Bebchuk (1984), as well as with Reinganum and Wilde (1986), where the probability of settlement is decreasing in x. In the present model, the endogenous investment decision entails a selection effect, which can delete monotonicity, as shown on Figure 7c. For a given investment decision, the trial probability increases with case strength. Ex ante, however, the opposite result may hold as plaintiffs with stronger cases are more likely to invest in case preparation.

[18] Admittedly, these results partly follow from our assumption that σ is a function of bounded variations (see footnote 16). Recall, however, that this restriction allows for an infinite number of discontinuity points.

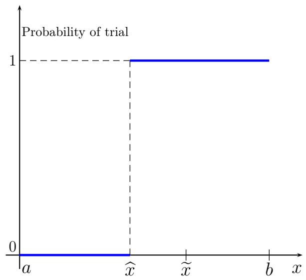

Figure 7a: Proba. of trial after $e = L$

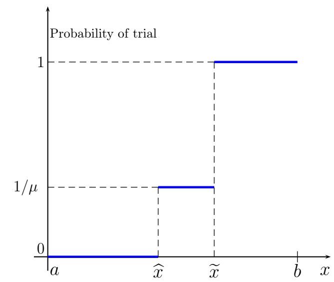

Figure 7b: Proba. of trial after $e = H$

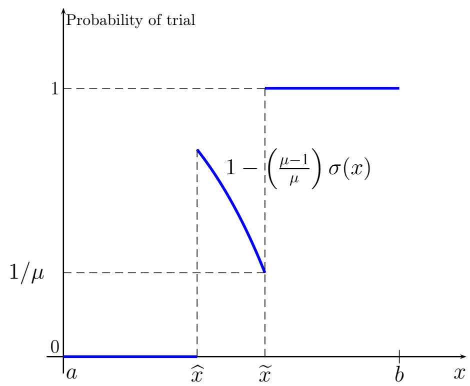

Figure 7c: Unconditional probability of trial

6.2 The extent of overinvestment

Plaintiffs with weak cases ($x < \widetilde{x}$) who decide to invest in case preparation can be called bluffers. If the defendant calls their bluff and brings them to court, they regret their choice. When $\widetilde{x} < x^*_H$, the proportion of bluffers is one, as all types invest, and the bluff is always successful as the defendant settles with all types below $x^*_H > \widetilde{x}$. When $x^*_H < \widetilde{x}$, however, this proportion, given by (7), is lower than one; moreover, the bluff is successful with probability $1 - 1/\mu$ only as the defendant randomizes between a low and a high offer after observing that the plaintiff has invested. The following Lemma describes how the extent of bluff varies with the primitives of the model.

Lemma 2 (Proportion of bluffers)**.** *In equilibrium, the proportion of bluffers becomes larger when (other things being equal):*
i) The defendant's trial cost, t_D, increases.
ii) The plaintiff's trial cost after e = L, t^P_L, increases.
iii) The sunk cost of investment, c, decreases.

Proof. i) An increase of t_D increases T_H, but not does not change $\widetilde{x}$. Moreover, the proportion of bluffers tends to one as t_D goes to $\mu\tau(\widetilde{x}) - t^P_H$, all other parameters being fixed. ii) and iii) If t^P_L increases (resp. c decreases), T_H/μ is not affected, while $\widetilde{x}$ decreases, therefore $\frac{T_H}{\mu}/\tau(\widetilde{x})$ increases. $\square$

6.3 The welfare effects of introducing a second technology

At the end of Section 2, we compared the expected litigation costs $\mathcal{C}^*_{\{H\}}$ and $\mathcal{C}^*_{\{L\}}$ in the two benchmark situations. We now investigate the welfare effects of introducing the costly technology, H, when only the basic one, L, is available, and, symmetrically, of introducing L when only H is available. The former exercise (introducing the H technology) is relevant when a costly, previously unavailable option (e.g. hiring an expert witness) becomes available. The latter exercise (introducing the L technology) might seem less intuitive: the plaintiff, who initially must resort to a costly technology, can now avoid it, at the cost of a reduced expected award. In practice, such a change could follow from a regulatory intervention that lowers the standard for legal representation. For instance, the government or the courts could allow self-representation: litigants could then choose to hire a licensed attorney or to represent themselves.

When $x^*_H \geq \widetilde{x}$, the effect is obvious as the equilibrium configuration is the same in $\{H\}$ and in $\{HL\}$. Hereafter, we focus on the case $x^*_H < \widetilde{x}$. The total litigation costs in $\{HL\}$ are given by:

$$\mathcal{C}^*_{\{HL\}} = \int_{\widehat{x}}^{\widetilde{x}} \left[\sigma(x)\frac{T_H}{\mu} + (1-\sigma(x))T_L\right] f(x)\mathrm{d}x + \left[1 - F(\widetilde{x})\right] T_H + c\Pr(e = H).$$

Plaintiffs with very weak cases ($x < \widehat{x}$) settle. Plaintiffs with intermediate cases ($\widehat{x} \leq x \leq \widetilde{x}$) invest with probability σ, then go to court with probability $1/\mu$, or do not invest (probability $1-\sigma$) and go to court with certainty; their contribution to the total costs is therefore $\sigma(x)\frac{T_H}{\mu} + (1-\sigma(x))T_L$. Finally, plaintiffs with strong cases ($x > \widetilde{x}$) invest and go to court, generating trial costs T_H.

Starting from a one-technology world $\{e\}$, the introduction of the case preparation technology has four effects on total costs. First, the critical case strength below which plaintiffs always settle is x^*_e in $\{e\}$ and $\widehat{x}$ in $\{HL\}$. The threshold $\widehat{x}$ is lower or higher than x^*_e depending on the ordering of T_L and T_H/μ. Second, in a (possibly empty) intermediate region $[\max(x^*_e, \widehat{x}), \widetilde{x}]$, T_e is replaced by $\sigma(x)\frac{T_H}{\mu} + (1-\sigma(x))T_L$, which is certainly smaller than T_H, and may be lower or higher than T_L depending, again, on the ordering of T_L and T_H/μ. Third, the introduction of L starting from $\{H\}$ does not change the contribution of plaintiffs with strong cases; the introduction of H starting from $\{L\}$ increases their contribution from T_L to T_H. Fourth, the introduction of a second technology modifies the overall investment probability $\Pr(e = H)$, and, in turn, the weight of the sunk cost c.

The overall effect is ambiguous in general, but can be determined in the particular case $T_H/\mu = T_L$. Under this assumption, the trial probability is the same in $\{H\}$ and in $\{L\}$, so the direct cost effect implies $\mathcal{C}^*_{\{H\}} > \mathcal{C}^*_{\{L\}}$. Furthermore, in $\{HL\}$, plaintiffs with weak cases generate the same expected trial costs, whether they invest (T_H/μ) or not (T_L).

Proposition 6 (Pure bluff effect)**.** *Assume that $x^*_H = x^*_L < \widetilde{x}$. Then*

i) From $\{L\}$, the introduction of H reduces the trial probability, raises the expected litigation costs, benefits the plaintiff, irrespective of his type, and harms the defendant.

ii) From $\{H\}$, the introduction of L reduces the trial probability and the expected litigation costs, is beneficial to the plaintiff as well as, if $x^*_L \geq T_L$, to the defendant.

Proof. See technical appendix H. □

The results of Proposition 6 are driven by a pure bluff effect. Under the assumption $x_H^* = x_L^* < \widetilde{x}$, the incentives to settle in $\{H\}$ and in $\{L\}$ are identical; the lower threshold for settlement, $\widehat{x}$, coincide with $x_H^* = x_L^*$. Yet a fraction of types above this threshold invests in the costly technology, and is rewarded by a generous settlement offer, thereby reducing the overall probability of trial.

Part ii of Proposition 6 shows that, if $T_L \leq x_L^* = x_H^* < \widetilde{x}$, the introduction of the basic technology, starting from the situation $\{H\}$, is Pareto-improving. The left inequality holds, for instance, when the distribution of case strength is uniform as $x_L^* = a + T_L$.

Under the assumptions of Proposition 6, the equilibrium total cost when both technologies are available lies between $\mathcal{C}_{\{H\}}^*$ and $\mathcal{C}_{\{L\}}^*$. This is not true in general. While overinvestment generates inefficient sunk costs, it may also trigger more settlements through the bluff effect, thereby reducing trial costs. The overall effect may be a reduction of the litigation costs. Starting from $\{L\}$, the introduction of H may reduce the expected litigation costs, even when H alone leads to higher costs than L alone. In other words, the ordering $\mathcal{C}_{\{HL\}}^* < \mathcal{C}_{\{L\}}^* < \mathcal{C}_{\{H\}}^*$ is possible.[19]

7 The game with continuous investment

So far, we have assumed that the investment in case preparation is a binary variable. We now provide some insights of what happens when the investment is continuous. The following model simply extends the previous one to the continuous case.

7.1 The model

The plaintiff chooses a continuous investment effort e. The expected award at trial is $\mu(e)x$, where μ increases in e. The effort entails a sunk cost $c(e)$ and a trial cost $t^P(e)$ that increase in e. Total trial costs after e are noted $T(e)$.

Under symmetric information, the defendant, after observing e, offers $\mu(e)x - t^P(e)$ to settle the case. The plaintiff accepts the offer and gets utility $\mu(e)x - t^P(e) - c(e)$. Accordingly, plaintiff x chooses effort to maximize

$$\max_e \mu(e)x - t^P(e) - c(e) \overset{\mathrm{d}}{\equiv} v_0(x).$$

The function v_0 is nondecreasing and convex in x. For simplicity, we assume that the maximum is attained at a unique value, that we note $e^\star(x)$. By the envelope theorem,

[19] Assume that x uniformly distributed on $[1, 3]$, and set: $c = 0.08$, $t_L^P = 0.5$, $t_D = 1.085$ so $T_L = 1.585$, $t_H^P = 0.54$, $t_D = 1.25$ so $T_H = 1.79$, and $\mu = 1.04$. Then we have: $x_L^* = 2.585$, $\widehat{x} = 2.699$, $x_H^* = 2.721$, $\widetilde{x} = 3$, and $\mathcal{C}_{\{HL\}}^* = 0.3259 < \mathcal{C}_{\{L\}}^* = 0.3288 < \mathcal{C}_{\{H\}}^* = 0.3295$. All plaintiffs have weak cases as $\widetilde{x} = b$, yet 86% of them invest in case preparation when both technologies are available. This excessive investment entails inefficient sunk costs, but comes with a higher settlement rate. The settlement probability is indeed 98% in $\{HL\}$ as opposed to 79% only in $\{L\}$. Accordingly, trial costs are reduced, which more than offsets the increase in preparation costs.

$v_0'(x) = \mu(e^\star(x))$. The complete information effort $e^\star(x)$ is nondecreasing in case strength x. For simplicity, we assume hereafter that $e^\star$ is continuous and increasing in x.[20]

Under asymmetric information, the defendant updates her belief on the distribution of types after observing effort e. Then she makes a unique settlement offer or randomizes across several offers, leaving utility $v(x|e)$ to plaintiff x. As explained in Section 3 and Lemma A.1, one can represent the defendant's strategy by the utility she leaves to the plaintiff. The latter chooses effort to maximize

$$v(x) = \max_e v(x|e). \tag{8}$$

Plaintiff x's strategy is a probability measure $\sigma(e|x)$. If he chooses a unique effort with certainty, $\sigma(e|x)$ is a mass point at the corresponding effort level; if he randomizes across several effort levels, the measure $\sigma(e|x)$ has a non degenerated support.

In the binary game, we have assumed that the optimal offers x_H^* and x_L^* are interior (Section 2.2). Similarly, we assume in the continuous game that, for all $x \in [a, b]$, the defendant problem if all plaintiffs choose $e^\star(x)$ admits an interior solution, i.e. the optimal offer in the Bebchuk continuation game after each relevant effort is interior.

7.2 Some properties of the equilibria

The results of this section follow from a line of reasoning similar to that previously used in the binary case. Details can be found in Appendix I. To begin with, we rule out semi-separating and pooling equilibria.

Proposition 7. *(i) For all but possibly one observed effort, the defendant cannot infer the plaintiff 's type (no semi-separating equilibria). (ii) The distribution of observed efforts has infinite support (no pooling equilibria).*

Point (i) of the Proposition states that at most one type separates in equilibrium. The intuition is as follows: if the defendant infers a plaintiff's type, x_0, after observing a particular effort, she offers the settlement amount that leaves this type indifferent with trial. This offer would attract lower types, thus putting constraints on v that make x_0 unique.

Point (ii) implies that the distribution of observed efforts cannot be a singleton: full pooling is impossible in the continuous model. To illustrate, suppose that all plaintiffs were to choose $e^\star(b)$, as shown on Figure 8. Since the corresponding Bebchuk offer is interior by assumption, the defendant would offer an amount smaller than $\mu(e^\star(b))b - t^P(e^\star(b))$, inducing agents x slightly below b to choose effort $e^\star(x)$ and go to court.[21] It is worth noticing the contrast with the binary game, where pooling occurs provided $\widetilde{x} < x_H^*$ (see Section 4). When $\widetilde{x}$ tends to b, the latter condition becomes more stringent, and is

[20] In other words, we assume here that v_0 is differentiable and strictly convex (see Figure 8). This contrasts with the binary case, where v_0 is the maximum of only two affine functions, and is therefore piecewise affine (see Figure 2).

[21] Pooling would occur only in the special case where the defendant always settles after all plaintiffs have chosen $e^\star(b)$, i.e. the continuation Bebchuk problem after $e^\star(b)$ admits a corner solution.

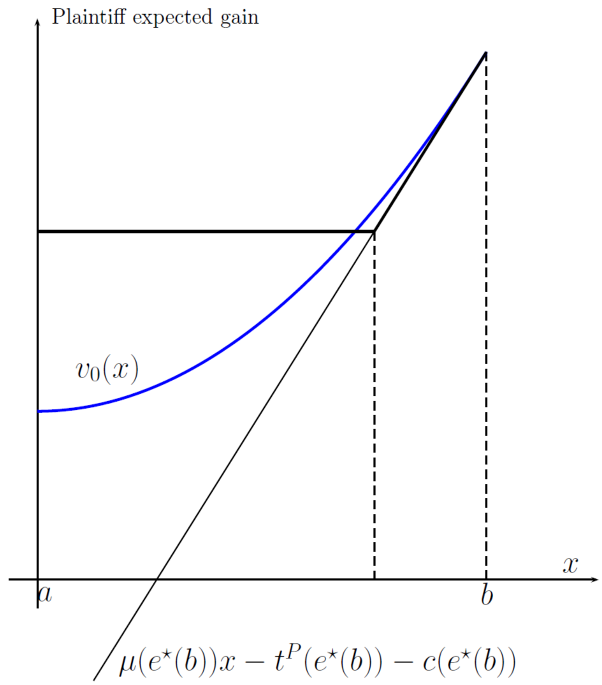

Figure 8: All types choosing $e^\star(b)$ cannot be an equilibrium

eventually incompatible with the Bebchuk offer after H being interior ($x_H^* < b$); this limit case helps understand why full pooling is impossible in the continuous game.

The next result that extends Proposition 4 to the continuous case establishes that the plaintiff necessarily bluffs in equilibrium.

Proposition 8. *In equilibrium, the probability that the defendant observes effort $e^\star(b)$ equals $f(b)T(e^\star(b))/\mu(e^\star(b))$, and is constant across equilibria.*

The proof of Proposition 8 shows that, in equilibrium, effort $e^\star(b)$ is optimal for any plaintiff, irrespective of his type, a property also true in the binary case. The proposition states that the aggregate probability that a plaintiff actually chooses effort $e^\star(b)$ is positive and constant across equilibria.

From Lemma I.1 in Appendix, we know that a plaintiff never chooses an effort below his full information effort. A natural candidate equilibrium in the continuous case would have plaintiffs x in some interval $[\widehat{x}, b]$ randomizing between the two values $e^\star(x)$ and $e^\star(b)$ and getting utility $v_0(x)$. Such a configuration would reproduce the equilibrium configuration of Proposition 2. However, the observation of efforts $e^\star(x)$, $x < b$, would allow the defendant to infer x, in violation of Proposition 7 (i). This suggests that extensive and complicated randomization from both parties is necessary to sustain an equilibrium with continuous investment.

7.3 How does the trial probability vary with case strength?

As shown in Section 3, the probability that agent x goes to court after choosing effort e is $v'(x|e)/\mu(e)$. By the envelope theorem, if agent x chooses effort e with positive probability ($\sigma(e|x) > 0$), we have: $v'(x|e) = v'(x)$. The conditional probability of trial for plaintiff x after effort e is therefore given by $v'(x)/\mu(e)$. This probability increases with x and decreases with e. The unconditional trial probability for plaintiff x is given by:

$$\text{Proba trial } (x) = \int_{e^\star(x)}^{e^\star(b)} \sigma(e|x)\frac{v'(x)}{\mu(e)}\text{d}e.$$

The unconditional probability of trial depends on case strength through two channels. After a given effort e, the conditional probability of trial increases with case strength, in accordance with the standard monotonicity result. The second channel is the selection effect: different plaintiffs make different efforts ($\sigma(e|x)$ depends on x).

The plaintiff with the highest type b chooses his perfect information effort $e^\star(b)$ with probability 1. The probability that he goes to court is therefore $v'(b)/\mu(e^\star(b))$. It follows that, for any type x

$$\frac{\text{Proba trial } (x)}{\text{Proba trial } (b)} = \frac{v'(x)}{v'(b)}.\int_{e^\star(x)}^{e^\star(b)} \sigma(e|x)\frac{\mu(e^\star(b))}{\mu(e)}\text{d}e. \tag{9}$$

The right-hand side of (9) is the product of two terms. The first one, $v'(x)/v'(b)$, is smaller than or equal to one by convexity. It tends to make the trial probability higher at b than at $x < b$, the usual monotonicity result. The second term reflects the selection effect. For any type x, this term is greater than or equal to one as μ increases in e, and thus plays in the opposite direction.[22] We do not know whether the selection effect can dominate when the investment is continuous, as is the case under a binary effort.

8 Concluding remarks

The above presentation has assumed that the informed party is the plaintiff. The model, however, allows for alternative interpretations. Consider for instance the following tax evasion situation, where the informed party is the defendant. After a preliminary inspection, the tax department has found that an agent (firm or individual) hid some transactions and that a certain amount of taxes has not been paid as a result. Yet thanks to skillful accounting practices, the agent can justify a fraction of this tax evasion. The amount the agent can justify is his private information. Once challenged by the authorities, the agent can hire a costly tax advisor, who is able to reduce the tax liability even further. After observing this choice, the tax department makes a settlement offer to save on inspection costs. If the settlement is rejected by the agent, a thorough inspection starts which is costly for both the tax department and the taxpayer.

[22] The second term equals one if and only if the plaintiff chooses effort $e^\star(b)$ with probability one. Since all plaintiffs choosing $e^\star(b)$ is not an equilibrium, the second term is *strictly* greater than one for a non negligible set of plaintiffs.

The model also applies to the following procurement issue. A firm or a government buys an input (e.g. a commodity) whose quality is variable. The price contractually depends on the estimated quality, but the quality audit is costly. To save on the evaluation costs, the buyer proposes a price to the supplier. If the latter refuses the proposed price, the audit is undertaken, and both the buyer and the supplier incur costs. The supplier can hire an engineer in charge of quality management, which entails a sunk cost but increases quality. The buyer makes her price offer after observing the supplier's effort in quality management.

More generally, the model applies to any context where a monetary transfer must be decided on the basis of unobserved characteristics of one party, and these characteristics can be revealed through a costly audit. To save on the auditing costs, the uninformed party makes an offer, but the informed party has the opportunity to move first, by investing to improve its position should the audit occur. Notice that the investment of the informed party could directly enhance the welfare through a real effect on attributes valued by the players (e.g. a quality improvement). In this paper, we have ruled out this possibility to focus exclusively on the signaling mechanism and the bluff effect.

Our model assumes that case preparation is more rewarding for stronger cases. But it accommodates easily to the reverse assumption that weaker cases benefit more from preparation. Bluff would then translate into underinvestment rather than overinvestment.

We have assumed that the informed party initiates the case, and makes his preparation decision before the opposite party can move. Depending on the circumstances, however, the uninformed party may be able to anticipate future litigation and to make an initial offer at the very beginning of the process. Accordingly, we could envisage successive sequences of investment decisions and settlement offers. It would be of interest to study the resulting dynamics, in the spirit of Spier (1992).

Finally, an obvious limitation of our framework is that we study the strategic effects of case preparation of one party only, leaving the pretrial efforts of the adversary exogenous. It is tempting to allow both parties to invest. Under symmetric information, this could result in an arms race. Depending on the costs and returns of the respective investments, both parties could invest and neutralize each other -a prisoner dilemma. The current framework emphasizes a different idea: plaintiffs with weak cases invest to manipulate the beliefs of the opposite party; overinvestment is inherent to asymmetric information.

References

Lucian Arye Bebchuk. Suing solely to extract a settlement offer. *Journal of Legal Studies*, 17(2): 437–450, June 1988.

Lucian Arye Bebchuk. Litigation and settlement under imperfect information. *RAND Journal of Economics*, 15(3):404–15, Autumn 1984.

Bernard Black, Charles Silver, David A. Hyman, and William M. Sage. Stability, not crisis: Medical malpractice claims in texas, 1988-2002. *Journal of Empirical Legal Studies*, 2:207–245, 2005.

Amitabh Chandra, Nundy Shantanu, and Seth A. Seabury. The growth of physician malpractice payments: Evidence from the national practitioner data bank. *Health Affairs*, 24:240–50, 2005.

Philippe Choné and Laurent Linnemer. Optimal litigation strategies with signaling and screening. *CREST Working paper*, 2008-16, 2008.

Robert D. Cooter and Daniel L. Rubinfeld. An economic model of legal discovery. *Journal of Legal Studies*, 23(1):435–463, January 1994.

Andrew F. Daughety and Jennifer F. Reinganum. Economic theories of settlement bargaining. *Annual Review of Law and Social Sciences*, (1):35–59, 2005.

Andrew F. Daughety and Jennifer F. Reinganum. Settlement negociations with two-sided asymmetric information: Model duality, information distribution, and efficiency. *International Review of Law and Economics*, 14:283–298, 1994.

Andrew F. Daughety and Jennifer F. Reinganum. Communicating quality: A unified model of disclosure and signaling. *RAND Journal of Economics*, 39(4):973–89, Winter 2008.

Andrew F. Daughety and Jennifer F. Reinganum. Hush money. *RAND Journal of Economics*, 30 (4):661–78, Winter 1999.

Amy Farmer and Paul Pecorino. Civil litigation with mandatory discovery and voluntary transmission of private information. *Journal of Legal Studies*, 34(1):137–160, January 2005.

Nuno Garoupa and Fernando Gomez-Pomar. Cashing by the hour: Why large firms prefer hourly fees over contingent fees. *Journal of Law Economics & Organization*, 24(2):458–475, 2008.

Bruce L. Hay. Effort, information, settlement, trial. *Journal of Legal Studies*, 24(1):29–62, January 1995.

David S. Kaplan, Joyce Sadka, and Jorge Luis Silva-Mendez. Litigation and settlement: New evidence from labor courts in mexico. *Journal of Empirical Legal Studies*, 5(2):309–350, 2008.

Robert Mnookin and Robert Wilson. A model of efficient discovery. *Games and Economic Behavior*, 25(2):219–250, November 1998.

Barry Nalebuff. Credible pretrial negotiation. *RAND Journal of Economics*, 18(2):198–210, Summer 1987.

Jennifer F. Reinganum and Louis L. Wilde. Settlement, litigation, and the allocation of litigation costs. *RAND Journal of Economics*, 17(4):557–66, Winter 1986.

Joel L. Schrag. Managerial judges: An economic analysis of the judicial management of legal discovery. *RAND Journal of Economics*, 30(2):305–23, Summer 1999.

Warren F. Schwartz and Abraham L. Wickelgren. Credible discovery, settlement, and negative expected value suits. *Working paper*, 2008.

Urs Schweizer. Litigation and settlement under two-sided incomplete information. *Review of Economic Studies*, 56(2):163–77, April 1989.

Steven Shavell. Sharing of information prior to settlement or litigation. *RAND Journal of Economics*, 20(2):183–95, Summer 1989.

Joel Sobel. An analysis of discovery rules. *Law and Contemporary Problems*, 52(1):133–59, Winter 1989.

Kathryn E. Spier. The dynamics of pretrial negotiation. *Review of Economic Studies*, 59(1):93–108, January 1992.

W. P. Ziemer. *Weakly Differentiable Functions: Sobolev Spaces and Functions of Bounded Variation.* Springer-Verlag, New York, 1989.

Appendix

Lemma 3. *Suppose that the plaintiff chooses technology e with positive probability. Let $x \geq a$ be such $F_e(x) = 0$. Then the defendant makes no offer smaller than or equal to $x - t_L^P$ ($\mu x - t_H^P$) after L (after H): $P_e = v_e' = 0$ on $[a, x]$.*

Proof. We write the proof for $e = L$. We proceed by contradiction. Suppose that the defendant makes an offer smaller than or equal to $x - t_L^P$: π_L attains its maximum at x. Since $F_L = 0$ on $[a, x]$, π_L is constant on $[a, x]$, and therefore, attains its maximum at a. This would imply

$$\int_a^y [-F_L(z) + T_L f_L(z)]\mathrm{d}z \leq 0 \text{ for all } y \in [a, b]\text{, or, noting } G_L(y) = \int_a^y F_L(z)\mathrm{d}z$$

$$e^{-y/T_L}\left[-\frac{1}{T_L}G_L(y) + F_L(y)\right] \leq 0.$$

The function $e^{-y/T_L}G_L(y)$ would be nonincreasing. Since it is zero at a, it would be everywhere nonpositive, implying that F_L and f_L would be identically 0. This is impossible as F_L is a probability distribution on $[a, b]$, the desired contradiction. □

Lemma 4. *Let $x \in (a, b)$ be such that π_e is not maximal at x. Then v_e is affine on a neighborhood of x. Convex kinks in v_e can occur only at points where π_e attains its maximum.*

Proof. Since π_e is continuous, it is not maximal in a neighborhood of x, which, therefore, does no intersect the support of P_e. Since, in (a, b), the support of P_e is the same as the support of v_e'' (see Appendix A), π_e is affine on that neighborhood. □

Lemma 5. *If x belongs to the support of P_H, then $\sigma(x^+) \leq \sigma(x^-)$. If x belongs to the support of P_L, then $\sigma(x^+) \geq \sigma(x^-)$. If x belongs to both supports, then σ is continuous at x.*

Proof. The result follows immediately from the fact that if x maximizes π_e, then we must have: $\pi_e'(x^-) \geq \pi_e'(x^+)$. □

The appendices referenced by a letter (Appendix A, Appendix A.1 . . .) are in a technical appendix which is available from the authors upon request.

Optimal Litigation Strategies with Observable Case Preparation

Philippe Choné and Laurent Linnemer

Technical appendix

A Representation of the defendant's strategy

Let K_L be the set of nondecreasing, convex functions v from $[a,b]$ to $\left[a-t_L^P, b-t_L^P\right]$ satisfying $v(b)=b-t_L^P$ and $0 \leq v' \leq 1$. Similarly, let K_H be the set of nondecreasing, convex functions v from $[a,b]$ to $\left[\mu a-t_H^P-c, \mu b-t_H^P-c\right]$ satisfying $v(b)=\mu b-t_H^P-c$ and $0 \leq v' \leq \mu$.

Lemma A.1. *There exists a one-to-one map between pairs (P_H, P_L) of probability distributions on $[a,b]$ and pairs $(v_H, v_L) \in K_H \times K_L$. Conditionally on the litigation technologies, the trial probabilities are given by:*

$$P_H(x \leq y) = v_H'(y)/\mu \quad \text{and} \quad P_L(x \leq y) = v_L'(y), \tag{10}$$

and are nondecreasing in case strength.

Proof. For all $x \in [a,b]$, the functions $v_{\{H\}}(.;x)$ and $v_{\{L\}}(.;x)$ belong to K_H and K_L respectively. Both sets are convex. The functions v_H and v_L given by (3) being convex combinations of the base functions $v_{\{H\}}(.;x)$ and $v_{\{L\}}(.;x)$, also belong to K_H and K_L. We have:

$$\begin{aligned} v_H(y) &= \left[\mu y - t_H^P\right] P_H(x \leq y) + \int_y^b [\mu x - t_H^P] \mathrm{d}P_L(x) - c & (11) \\ v_L(y) &= \left[y - t_L^P\right] P_L(x \leq y) + \int_y^b [x - t_L^P] \mathrm{d}P_L(x), & (12) \end{aligned}$$

where $P_e(x \leq y)$ is the probability that the plaintiff of type y, if he exerts effort e, goes to trial. We refer to this probability as the *conditional* trial probability. Being nondecreasing and convex, v_e is differentiable almost everywhere.[23] Differentiating (11) and (12) yields (10). Since v_L and v_H are convex, the conditional trial probabilities are nondecreasing in case strength.

We show now that the base functions $v_{\{H\}}(.;x)$ and $v_{\{L\}}(.;x)$, $a \leq x \leq b$, generate the convex sets K_H and K_L and explain how any function $v_e \in K_e$, $e = H, L$, can be written in the form (3).

[23] More precisely, v_e admits at every point a left- and a right-derivative, which coincide almost everywhere.

We have already shown that if P_e is a probability distribution on $[a,b]$, then the functions v_H and v_L given (3) belongs to K_H and K_L. Conversely, pick any $v_L \in K_L$. Integrating by parts and using $v_L(b) = b - t_L^P$, we have

$$\begin{aligned}
v_L(x) &= -\int_x^b v_L'(y)\mathrm{d}y + b - t_L^P \\
&= \int_x^b v_L''(y).[y - t_L^P]\mathrm{d}y + [1 - v_L'(b)].[b - t_L^P] + v_L'(x).[x - t_L^P] \\
&= \int_x^b v_L''(y).[y - t_L^P]\mathrm{d}y + \int_a^x v_L''(y).[x - t_L^P]\mathrm{d}y \\
&\quad +[1 - v'(b)].[b - t_L^P] + v_L'(a)[x - t_L^P] \\
&= \int_a^b v_L''(y)\max\{x - t_L^P, y - t_L^P\}\mathrm{d}y \\
&\quad +[1 - v_L'(b)](b - t_L^P) + [x - t_L^P]v_L'(a).
\end{aligned}$$

The same computation yields

$$\begin{aligned}
v_H(x) &= \int_a^b \frac{v_H''(y)}{\mu}\max\{\mu x - t_H^P, \mu y - t_H^P\}\mathrm{d}y \\
&\quad + \left[1 - \frac{v_H'(b)}{\mu}\right](\mu b - t_H^P) + [\mu x - t_H^P]\frac{v_H'(a)}{\mu} - c.
\end{aligned}$$

It follows that any $v_e \in K_e$ can be written according to (3) with P_e given by

$$\begin{aligned}
P_L(y) &= v_L'(a)\delta_a + v_L''(y) + [1 - v_L'(b)]\delta_b \\
P_H(y) &= \frac{v_H'(a)}{\mu}\delta_a + \frac{v_H''(y)}{\mu} + \left[1 - \frac{v_H'(b)}{\mu}\right]\delta_b,
\end{aligned}$$

where δ_x represents the mass point at x. Since v_e is a convex function, its first-order derivative is a nondecreasing function and its second-order derivative is a positive measure. We have

$$\int_a^b \mathrm{d}P_L(y) = v_L'(a) + \int_a^b v_L''(y)\mathrm{d}y + 1 - v_L'(b) = 1,$$

so the total mass of P_L is 1. The same result holds for P_H. So P_L and P_H are probability distributions on $[a,b]$.

Notice that P_e could, in theory, have mass points at a, at b and at any interior point in (a,b). An interior mass point corresponds, as already mentioned, to a convex kink of v_e (jump of v_e', mass in v_e''). $\square$

B Evaluation of the defendant's profit

Lemma B.1. *The conditional profits of the defendant given $e = H$ and $e = L$ can be expressed as functions of her strategy (v_H, v_L) in the following way:*

$$\begin{aligned}\Pi_H &= \int_a^b \left[-v_H(x) - \frac{T_H}{\mu} v'_H(x)\right] f_H(x)\mathrm{d}x - c \\ \Pi_L &= \int_a^b \left[-v_L(x) - T_L v'_L(x)\right] f_L(x)\mathrm{d}x.\end{aligned}$$

Proof. Integrating twice by parts and using the fact that $\pi_L(b) = -(b - t_L^P)$, we find that the defendant's profit, when she faces $e = L$, is given by

$$\begin{aligned}\int_a^b \left[-v_L(x) - T_L v'_L(x)\right] f_L(x)\mathrm{d}x &= \int_a^b v'_L(x)[F_L(x) - T_L f_L(x)]\mathrm{d}x - v_L(b) \\ &= -\int_a^b v'_L(x)\pi'_L(x) - [b - t_L^P] \\ &= \int_a^b \pi_L(x)v''_L(x)\mathrm{d}x - v'_L(b)\pi_L(b) + v'_L(a)\pi_L(a) - [b - t_L^P] \\ &= \int_a^b \pi_L(x)v''_L(x)\mathrm{d}x + [1 - v'_L(b)]\pi_L(b) + v'_L(a)\pi_L(a) \\ &= \int_a^b \pi_L(x)\mathrm{d}P_L(x).\end{aligned}$$

Similarly, the defendant's profit, when she faces $e = H$, is given by (use $\pi_H(b) = -(\mu b - t_H^P)$)

$$\int_a^b \left[-v_H(x) - \frac{T_H}{\mu} v'_H(x)\right] f_H(x)\mathrm{d}x - c =$$

$$\begin{aligned}&= \int_a^b v'_H(x)\left[F_H(x) - \frac{T_H}{\mu} f_H(x)\right] - v_H(b) - c \\ &= \frac{1}{\mu}\left(\int_a^b \pi_H(x)v''_H(x)\mathrm{d}x - v'_H(b)\pi_H(b) + v'_H(a)\pi_H(a)\right) - [\mu b - t_H^P] \\ &= \int_a^b \pi_H(x)v''_H(x)/\mu\,\mathrm{d}x + [1 - v'_H(b)/\mu]\pi_H(b) + \pi_H(a)v'_H(a)/\mu \\ &= \int_a^b \pi_H(x)\mathrm{d}P_H(x).\end{aligned}$$

□

C Proof of Proposition 1 (uniqueness part)

The proof of the uniqueness of the equilibrium configuration when $x_H^* > \widetilde{x}$ requires a preliminary result.

Lemma C.1. *Let E be an equilibrium such that $v_L(\widetilde{x}) = \widetilde{x} - t_L^P$. Let $\check{E}$ be the same configuration as E except that v_H is replaced by* $\max(v_L, v_H)$ *on* $[a, \widetilde{x}]$.

Then $\check{E}$ is an equilibrium. For both litigants, the payoffs are the same at E and $\check{E}$.

Proof. We have $v_H(\widetilde{x}) \geq v_L(\widetilde{x}) = \widetilde{x} - t_L^P$, which yields: $\check{v}_H(\widetilde{x}) = v_H(\widetilde{x})$. It follows immediately that $\check{v}_H$ belongs to K_H and that the change corresponds to an admissible defendant's strategy $\check{P}_H$. The only impact of the change is the following: the plaintiffs who strictly preferred L to H at E are indifferent between the two levels of effort at $\check{E}$. Those plaintiffs can therefore continue to choose $e = L$ with certainty ($\check{\sigma} = \sigma = 0$ is optimal for them).

We now use Lemma B.1 (see Appendix B) to show that the defendant's profit is maximal at $\check{E}$. Since $\check{v}_l = v_L$ and the plaintiff's strategy is the same at E and $\check{E}$ ($\check{\sigma} = \sigma$), the defendant's payoff when $e = L$ is the same at E and $\check{E}$. Therefore the strategy v_L still maximizes the defendant's payoff when $e = L$.

If the plaintiff invests in case preparation, the utility v_H changes only when $\sigma = 0$, so the integral $\int_a^b v_H(x) f_H(x) \mathrm{d}x$ is not affected. The derivatives v'_H and $\check{v}'_H$ coincide whenever $v_L \neq v_H$, but can be different at points where $v_H = v_L$ and $\sigma = \check{\sigma} > 0$ (corresponding to plaintiffs who randomize between the two technologies). The difference between v'_H and $\check{v}'_H$ at points where f_H is positive could, in principle, affect the defendant's payoff. In fact, this is not the case, since such a possibility can only occur on a negligible set.

Indeed, v_H and v_L are differentiable almost everywhere. If $v_H(x) = v_L(x)$ and $v'_H(x) = v'_L(x)$, then $\check{v}'_H(x) = v'_H(x)$, so v'_H and $\check{v}'_H$ coincide. If $v_H(x) = v_L(x)$ and $v'_H(x) \neq v'_L(x)$, then $\check{v}_H$ is not differentiable at x, which can occur only a negligible subset.

It follows that v'_H and $\check{v}'_H$ coincide at almost every x such that $\sigma > 0$ and that the defendant's payoff when $e = H$ is the same at E and $\check{E}$. We conclude that the change in the defendant's strategy (from v_H to $\check{v}_H$) does not affect her payoff.

□

To prove uniqueness, we proceed by contradiction. We suppose that there exists an equilibrium where the probability of $e = L$ is positive. From the proof of Lemma 1, we know that $v_L(\widetilde{x}) = \widetilde{x} - t_L^P$. Lemma C.1 applies: there exists another equilibrium with the same plaintiff's strategy, the same payoff for both litigants and $v_H \geq v_L$ on $[a, \widetilde{x}]$. We now work with this equilibrium and exhibit a contradiction. Since, by assumption, the probability of $e = L$ is positive and given that high types ($x > \widetilde{x}$) chooses $e = H$ (recall Lemma 1), we cannot have $v_H > v_L$ everywhere on $(a, \widetilde{x}]$. Let x_1, $a < x_1 \leq \widetilde{x}$, be the highest solution to $v_L = v_H$. By construction, all plaintiffs with type $x > x_1$ strongly prefer $e = H$ to $e = L$; in other words: $v_H > v_L$ and $\sigma = 1$ on $(x_1, b]$. It follows that $\sigma(x_1^+) = 1$.

We now use the property $v_H \geq v_L$ on $[a, \widetilde{x}]$ (coming from Lemma C.1) to prove that x_1 necessarily belongs to the support of P_H. Here again, we proceed by contradiction. Suppose that x_1 were not in the support of P_H. Then, by Lemma 4, v_H would be affine in a small interval around x_1. The function v_L would be below this affine function for $x \leq x_1$, would coincide with it at x_1 and would be *strictly* below for $x > x_1$. This would violate the convexity of v_L. We conclude that x_1 belongs to the support of P_H.

Yet we have:

$$\begin{aligned}
-\int_a^{x_1} \sigma(t)\mathrm{d}F(t) + \frac{T_H}{\mu}\sigma(x_1^+)f(x_1) &= -\int_a^{x_1} \sigma(t)\mathrm{d}F(t) + \frac{T_H}{\mu}f(x_1) \\
&\geq -F(x_1) + \frac{T_H}{\mu}f(x_1) \\
&= f(x_1).\,[T_H/\mu - \tau(x_1)] \\
&\geq f(x_1).\,[T_H/\mu - \tau(\widetilde{x})] \\
&> 0.
\end{aligned}$$

This implies that $\pi'_H(x_1^+) > 0$, which contradicts $x_1 \in \operatorname{supp} P_H \subset \operatorname{argmax} \pi_H$. This contradiction yields uniqueness. $\square$

D Proof of the first part of Proposition 2

We assume that $T_H/\mu < \tau(\widetilde{x})$ or $x_H^* < \widetilde{x}$. We already know that the defendant makes no offer greater than $\widetilde{x}$ after observing L, that is $v_L(\widetilde{x}) = \widetilde{x} - t_L^P$. In this appendix, we show that, after observing $e = H$, the defendant makes no offer greater than $\mu\widetilde{x} - t_H^P$ with positive probability, which is equivalent to $v_H(\widetilde{x}) = v_L(\widetilde{x})$.

We proceed by contradiction. Suppose that the defendant makes an offer $\mu x_2 - t_H^P$, $\widetilde{x} < x_2 \leq b$, with positive probability. We follow the same argument as in the proof of the uniqueness part of Proposition 1. Since $v_L(\widetilde{x}) = \widetilde{x} - t_L^P$, we can apply Lemma C.1: replacing v_H by $\max(v_L, v_H)$ on $[a, \widetilde{x}]$, leaving v_L and σ unchanged, we get another equilibrium where both litigants get the same payoff. We now work with this new equilibrium and exhibit a contradiction. Since, by assumption, the probability of L is positive and given that high types ($x > \widetilde{x}$) chooses $e = H$ by Lemma 1, we cannot have $v_H > v_L$ everywhere on $(a, \widetilde{x}]$. Let $a < x_1 < \widetilde{x}$ be the highest solution to $v_L = v_H$.[24] By construction, all plaintiffs with type $x > x_1$ strongly prefer $e = H$ to $e = L$; in other words: $v_H > v_L$ and $\sigma = 1$ on $(x_1, b]$.

We now use the property (coming from Lemma C.1) that $v_H \geq v_L$ on $[a, \widetilde{x}]$, to prove that x_1 belongs to the support of P_H. Again, we proceed by contradiction. Suppose that x_1 were not in the support of P_H. Then, by Lemma 4, v_H would be affine in a small interval around x_1. Since $v_L \leq v_H$ on $[a, \widetilde{x}]$, the function v_L would be below this affine function for $x \leq x_1$, would coincide with it at x_1 and would be *strictly* below for $x > x_1$. This would violate the convexity of v_L. We conclude that x_1 must belong to the support of P_H.

On $[x_1, b]$, we have $\sigma = 1$ and π'_H is equal (up to a positive multiplicative factor) to

$$-\int_a^x \sigma(t)f(t)\mathrm{d}t + \frac{T_H}{\mu}f(x) = \kappa - F(x) + \frac{T_H}{\mu}f(x),$$

where $\kappa = \int_a^{x_1}(1-\sigma)\mathrm{d}F > 0$. Since τ is nondecreasing, the function π_H on $[x_1, b]$ is first nondecreasing, then nonincreasing: it attains its maximum at some point(s) greater than

[24] Note that, in the proof of the uniqueness part of Proposition 1, we only knew that $x_1 \leq \widetilde{x}$.

x_H^*. We must therefore have: $x_H^* < x_1 < \widetilde{x} < x_2$. Since the function π_H is maximal at x_1 and at x_2, it must remain constant in between, which implies $\kappa - F(x) + \frac{T_H}{\mu} f(x) = 0$, or

$$\frac{T_H}{\mu}\frac{1}{\tau} = 1 - \frac{\kappa}{F}$$

on $[x_1, x_2]$. Since the left-hand side is decreasing and the right-hand side is increasing, it follows that both sides are constant, which implies that F is constant, and $f = 0$ on $[x_1, x_2]$, the desired contradiction. It follows that $v_L(\widetilde{x}) = v_H(\widetilde{x})$.

E Proof of the second part of Proposition 2

The proof is organized as follows. First, we show that after $e = H$, the defendant offers $\mu\widetilde{x} - t_H^P$ with probability $1 - 1/\mu$. Second, we establish a number of properties of offers made after $e = L$. Third, we exploit these properties to prove that, after $e = L$, the defendant makes only one offer (the support of P_L is the singleton). Finally, noting $\widehat{x} - t_L^P$ this single offer, we show that the indifferent plaintiff $\widehat{x}$ uses a non-degenerated mixed strategy: $0 < \sigma(\widehat{x}) < 1$.

E.1 After $e = H$, the defendant offers $\mu\widetilde{x} - t_H^P$ with probability $1 - 1/\mu$

Since $v_H = v_L$ on $[a, \widetilde{x}]$ and $v_L \in K_L$, the left-derivative of v_H at $\widetilde{x}$ is lower than or equal to 1. Since $v'_H(\widetilde{x}^+) = \mu$, the derivative of v_H/μ jumps upwards, with a jump greater than or equal to $1 - 1/\mu$. The probability of playing $\widetilde{x}$ after $e = H$ has to be greater (or equal to) than $1 - 1/\mu$.

It follows that π_H is maximal at $\widetilde{x}$. This implies that σ is continuous at $\widetilde{x}$. Indeed, if $\sigma(\widetilde{x}^-)$ were strictly smaller than $\sigma(\widetilde{x}^+) = 1$, we would have $\pi'_H(\widetilde{x}^-) < \pi'_H(\widetilde{x}^+) = 0$ and π_H would not be maximal at $\widetilde{x}$.

Since $-F_L(\widetilde{x}) + T_L f_L(\widetilde{x}) = -F_L(\widetilde{x}) < 0$, π_L is *not* maximal at $\widetilde{x}$ and v_L is affine on a neighborhood of $\widetilde{x}$. This implies $v'(\widetilde{x}^-) = 1$. So the jump of v'_H/μ at $\widetilde{x}$ is exactly $1 - 1/\mu$. We conclude that, after $e = H$, the defendant offers $\mu\widetilde{x} - t_H^P$ with probability $1 - 1/\mu$

E.2 Properties of offers made after $e = L$

From the proof of Lemma 1, we know that any offer made after $e = L$, say $\widehat{x} - t_L^P$, satisfies: $\widehat{x} \leq \widetilde{x}$. As mentioned above, π_L is not maximal at $\widetilde{x}$, so we have: $\widehat{x} < \widetilde{x}$. From Proposition 2, we know that $v_H = v_L$ on $[a, \widetilde{x}]$. From Lemma A.1 (see Appendix A), we deduce that the supports of P_L and P_H coincide on this interval. It follows that $\widehat{x}$ belongs to the support of both P_H and P_L, and that the functions π_H and π_L both attain their maximum at $\widehat{x}$. From Lemma 5, σ must be continuous at this point, and π_H and π_L are differentiable at $\widehat{x}$, so we have

$$\int_a^{\widehat{x}} \sigma \mathrm{d}F = \frac{T_H}{\mu}\sigma(\widehat{x})f(\widehat{x}) \tag{13}$$

$$\int_a^{\widehat{x}} (1-\sigma) \mathrm{d}F = T_L[1 - \sigma(\widehat{x})]f(\widehat{x}). \tag{14}$$

Adding up this two equations yields

$$F(\widehat{x}) = [T_H/\mu - T_L]\,\sigma(\widehat{x}) f(\widehat{x}) + T_L f(\widehat{x}),$$

which implies $\widehat{x} = x_H^* = x_L^*$ when $T_H/\mu = T_L$, and equation (15) otherwise:

$$\sigma(\widehat{x}) = \frac{\tau(\widehat{x}) - T_L}{T_H/\mu - T_L}. \tag{15}$$

As the probability $\sigma(\widehat{x})$ is between 0 and 1, and as τ is increasing, it follows that $\widehat{x}$ lies between x_L^* and x_H^*. More precisely, if $T_L < T_H/\mu$, then $x_L^* < \widehat{x} < x_H^*$ and if $T_H/\mu < T_L$, then $x_H^* < \widehat{x} < x_L^*$ (see Figures 6a and 6b).

Since π_H attains its maximum at $\widehat{x}$, we have

$$\int_{\widehat{x}}^{x} \left[-F_H(y) + \frac{T_H}{\mu} f_H(y)\right] \mathrm{d}y \le 0$$

for all y. Setting $G_H(x) = \int_{\widehat{x}}^{x} F_H(y)\mathrm{d}y$, we get

$$-G_H(x) + \frac{T_H}{\mu} F_H(x) \le \frac{T_H}{\mu} F_H(\widehat{x}). \tag{16}$$

Multiplying by $\exp(-\mu x/T_H)$ yields

$$\exp(-\mu x/T_H)\left[-\frac{\mu}{T_H} G_H(x) + F_H(x)\right] \le F_H(\widehat{x}) \exp(-\mu x/T_H).$$

Integrating between $\widehat{x}$ and $x \ge \widehat{x}$ yields

$$G_H(x) \le \frac{T_H}{\mu} F_H(\widehat{x}) \left[\exp\left(\frac{\mu}{T_H}(x - \widehat{x})\right) - 1\right]$$

Now using (16) yields, after simplification

$$F_H(x) \le F_H(\widehat{x}) \exp\left[\frac{\mu}{T_H}(x - \widehat{x})\right] \tag{17}$$

for every $x \ge \widehat{x}$. The same computation for the L-technology shows that

$$F_L(x) \le F_L(\widehat{x}) \exp\left[\frac{1}{T_L}(x - \widehat{x})\right]. \tag{18}$$

E.3 After $e = L$, the defendant makes only one offer

We now prove that P_L is a mass point. We proceed by contradiction. We assume that there exists x_1 and x_2, $a < x_1 < x_2 < \widetilde{x}$, both in the supports of P_L and P_H.

If $T_H/\mu = T_L$, we already know that this is impossible, since we must have $x_1 = x_2 = x_H^* = x_L^*$. We assume $T_L < T_H/\mu$, which implies (see the preceding section): $x_L^* < x_1 <$

$x_2 < x_H^*$.[25] Applying (17) for $\widehat{x} = x_1$ and $x = x_2$ we obtain (using the f.o.c. (13) for both x_1 and x_2):

$$\sigma(x_2)f(x_2) \leq \sigma(x_1)f(x_1)\exp\left[\frac{\mu}{T_H}(x_2 - x_1)\right]. \tag{19}$$

From (15), we have for $i = 1, 2$:

$$\sigma(x_i)f(x_i) = \frac{F(x_i) - T_L f(x_i)}{T_H/\mu - T_L} > 0$$

and as $T_L < T_H/\mu$, $F(x_i) - T_L f(x_i) > 0$. Therefore, inequality (19) can be rewritten as

$$F(x_2) - T_L f(x_2) \leq (F(x_1) - T_L f(x_1))\exp\left[\frac{\mu}{T_H}(x_2 - x_1)\right]$$

or taking the logarithm

$$\log\left(F(x_2) - T_L f(x_2)\right) - \log\left(F(x_1) - T_L f(x_1)\right) \leq \frac{\mu}{T_H}(x_2 - x_1). \tag{20}$$

Let $\Lambda(x) = \ln\left(F(x) - T_L f(x)\right) = \ln F(x) + \ln\left(1 - T_L/\tau(x)\right)$. Since F is log-concave and τ is increasing, we have $\tau(x_2) < \tau(x_H^*) = T_H/\mu$, and

$$\Lambda(x_2) - \Lambda(x_1) \geq \ln F(x_2) - \ln F(x_1) \geq \frac{1}{\tau(x_2)}.(x_2 - x_1) \geq \frac{\mu}{T_H}(x_2 - x_1), \tag{21}$$

which, combined to (20), yields $x_1 = x_2$, the desired contradiction. We must therefore have: $x_1 = x_2$.

E.4 The plaintiff $\widehat{x}$ randomizes: $0 < \sigma(\widehat{x}) < 1$

We now prove that the probability $\sigma(\widehat{x})$ is strictly between 0 and 1 as it satisfies

$$\frac{f(\widetilde{x})}{f(\widehat{x})}\exp\left[\frac{\mu}{T_H}(\widehat{x} - \widetilde{x})\right] \leq \sigma(\widehat{x}) \leq 1 - \frac{\Pr(e = L)}{T_L f(\widehat{x})}\exp\left[\frac{1}{T_L}(\widehat{x} - \widetilde{x})\right], \tag{22}$$

where $\Pr(e = L) = 1 - \Pr(e = H)$ is known from (7).

Indeed, using the first-order condition (13) and applying (17) at $x = \widetilde{x}$ yields

$$F_H(\widetilde{x}) = \frac{T_H}{\mu}\frac{f(\widetilde{x})}{\Pr(e = H)} \leq \frac{T_H}{\mu}\frac{\sigma(\widehat{x})f(\widehat{x})}{\Pr(e = H)}\exp\left[\frac{\mu}{T_H}(\widetilde{x} - \widehat{x})\right]$$

which yields the left inequality of (22). Using the first-order condition (14) and applying (18) at $x = \widetilde{x}$ yields

$$F_L(\widetilde{x}) = 1 \leq T_L\frac{(1 - \sigma(\widehat{x}))f(\widehat{x})}{\Pr(e = L)}\exp\left[\frac{1}{T_L}(\widetilde{x} - \widehat{x})\right]$$

which yields the right inequality of (22).

[25] The case $x_H^* < x_1 < x_2 < x_L^*$ is treated similarly starting with (18) instead of (17).

F The defendant's payoff

Assume that $x_H^* < \widetilde{x}$. The defendant's profit can be expressed as the difference between the total welfare and the utility left to the plaintiff (see Lemma B.1). Given that $v_L = v_H = v^*$, we have:

$$\begin{aligned} -\Pi^*_{\{HL\}} &= -\Pr(e = L)\Pi^*_L - \Pr(e = H)\Pi^*_H \\ &= \int_a^b \left\{ [v(x) + T_L v'(x)][1 - \sigma(x)] + \left[v(x) + \frac{T_H}{\mu} v'(x) \right] \sigma(x) \right\} f(x)\mathrm{d}x \\ &\quad + c\Pr(e = H). \end{aligned}$$

Using equation (7) and observing that $v' = 1$ on $[\widehat{x}, \widetilde{x}]$ yields:

$$\begin{aligned} -\Pi^*_{\{HL\}} &= (\widehat{x} - t_L^P) \int_a^{\widehat{x}} f(x)\mathrm{d}x + \int_{\widehat{x}}^{\widetilde{x}} [x + t_D]\, f(x)\mathrm{d}x + \int_{\widetilde{x}}^b [\mu x + t_D]\, f(x)\mathrm{d}x \\ &\quad + c\frac{T_H}{\mu} f(\widetilde{x}) + \left(\frac{T_H}{\mu} - T_L \right) \int_{\widehat{x}}^{\widetilde{x}} \sigma(x) f(x)\mathrm{d}x. \end{aligned} \tag{23}$$

G Construction of an equilibrium (Proof of Proposition 3)

We first check that the triplet (σ^*) We examine successively the case $T_H/\mu < T_L$, $T_H/\mu = T_L$ and $T_H/\mu > T_L$. Hereafter, we adopt the natural notations: $f_L^* = f(x_L^*)$, $F_L^* = F(x_L^*)$, $f_H^* = f(x_H^*)$, $F_H^* = F(x_H^*)$, $\tilde{f} = f(\widetilde{x})$, $\tilde{F} = F(\widetilde{x})$.

G.1 The case $T_H/\mu < T_L$

Using the log-concavity of F, we get, for $x_H^* < x_1 < x_2 < \widetilde{x}$:

$$\ln F(x_2) - \ln F(x_1) \leq \frac{1}{\tau(x_1)}(x_2 - x_1) \leq \frac{\mu}{T_H}(x_2 - x_1).$$

It follows that $-\ln F(x) + \frac{\mu}{T_H} x$ is nondecreasing on $[x_H^*, \widetilde{x}]$. Since $\tau = F/f$ is increasing, it follows that $-\ln f(x) + \frac{\mu}{T_H} x$ is increasing, and so is σ^* on $[x_H^*, \widetilde{x}]$. Since $T_H/\mu - T_L < 0$, the function $(\tau - T_L)/(T_H/\mu - T_L)$ is decreasing on $[x_H^*, \widetilde{x}]$, is equal to 1 at x_H^* and to 0 at x_L^*. By continuity, the equation

$$\sigma^*(x) = \frac{\tau(x) - T_L}{T_H/\mu - T_L} \tag{24}$$

has a unique solution $\widehat{x}$ in $[x_H^*, \min(\widetilde{x}, x_L^*)]$. Notice that we have: $\sigma^* \geq (\tau - T_L)/(T_H/\mu - T_L)$ on $[\widehat{x}, \widetilde{x}]$.

We have: $\sigma^*(\widetilde{x}) = 1$, so σ^* is continuous at $\widetilde{x}$. The function σ^* takes its value in $(0, 1]$ and is continuous on $[a, b]$. Since σ^* is constant on $[a, x_H^*]$, we have have

$$\begin{aligned} \int_a^{x_H^*} \sigma^*(x) f(x)\mathrm{d}x = \frac{\tilde{f}}{f_H^*} \exp\left[\frac{\mu}{T_H}(x_H^* - \widetilde{x}) \right] .F_H^* &= \frac{T_H}{\mu} \tilde{f} \exp\left[\frac{\mu}{T_H}(x_H^* - \widetilde{x}) \right] \\ &= \frac{T_H}{\mu} \sigma^*(x_H^*) f(x_H^*), \end{aligned}$$

which yields $\pi'_H(x^*_H) = 0$. From the definition of σ^* on $(x^*_H, \widetilde{x})$, it follows that π'_H is constant on that interval, which, given that $\pi'_H(x^*_H) = 0$, yields $\pi'_H = 0$ on that interval. On $[a, x^*_H]$, the quantity $-F_H + \frac{T_H}{\mu} f_H$ is equal, up to a multiplicative positive constant, to $-F + \frac{T_H}{\mu} f \geq 0$, which yields $\pi'_H \geq 0$ on $[a, x^*_H]$. Finally, on $[\widetilde{x}, b]$, we have $\sigma^* = 1$ and the quantity $-F_H + \frac{T_H}{\mu} f_H$ is equal, up to an additive constant, to $-F + \frac{T_H}{\mu} f$, which is decreasing on $[\widetilde{x}, b]$ (by log-concavity of F). This shows that $\pi'_H \leq 0$ on this interval. In sum, the function π_H is nondecreasing on $[a, x^*_H]$ and constant on $[x^*_H, \widetilde{x}]$ and nonincreasing (and concave) on $[x, b]$.

Using $\widehat{x} < x^*_L$, σ constant, and the log-concavity of F, it is easy to check that π_L is nondecreasing on $[a, \widehat{x}]$. On $[\widehat{x}, \widetilde{x}]$, we have, by using $\pi'_H = 0$:

$$\begin{aligned} -\int_a^x (1-\sigma^*(t))f(t)\mathrm{d}t + T_L(1-\sigma^*(x))f(x) &= -F(x) + \int_a^x \sigma^*(t)f(t)\mathrm{d}t + T_L(1-\sigma^*(x))f(x) \\ &= -F(x) + \left(\frac{T_H}{\mu} - T_L\right)\sigma^*(x)f(x) + T_L f(x) \\ &= (T_H/\mu - T_L)f(x)\left[\sigma^*(x) - \frac{\tau(x) - T_L}{T_H/\mu - T_L}\right] \\ &\leq 0 \end{aligned}$$

which yields $\pi'_L \leq 0$. It follows that π_L is nonincreasing on $[\widehat{x}, \widetilde{x}]$. On $[\widetilde{x}, b]$, we have $\sigma^* = 1$, and π_L is affine and decreasing. In sum, π_L is nondecreasing on $[a, \widehat{x}]$ and nonincreasing on $[\widehat{x}, b]$.

We have shown that π_H attains its maximum at $\widehat{x}$ and $\widetilde{x}$ and that π_L attains its maximum at $\widehat{x}$. Therefore the corresponding strategy (P_H, P_L) is optimal for the defendant. Any plaintiff with type below $\widetilde{x}$ is indifferent between investing or not, and may randomize according to the proposed probability σ^*.

□

G.2 The case $T_H/\mu = T_L$

Using σ constant and the log-concavity of F, it is easy to check that π_H and π_L are nondecreasing on $[a, \widehat{x}]$. As above, π_H is constant on $[x^*_H, \widetilde{x}]$ and decreasing and concave on $[\widetilde{x}, b]$. On $[\widehat{x}, \widetilde{x}]$, the same computation as above shows that

$$-\int_a^x (1-\sigma^*(t))f(t)\mathrm{d}t + T_L(1-\sigma^*(x))f(x) = -F(x) + T_L f(x) = -F(x) + \frac{T_H}{\mu}f(x) \leq 0,$$

implying that π_L is nonincreasing on $[\widehat{x}, \widetilde{x}]$. On $[\widetilde{x}, b]$, π_L is still affine and decreasing.

□

G.3 The case $T_H/\mu > T_L$

As shown above, the function $\left(\tilde{f}/f\right) \exp\left[\frac{\mu}{T_H}(x - \widetilde{x})\right]$ is increasing on $[x^*_H, \widetilde{x}]$, and is equal to 1 at $\widetilde{x}$. Thus the value of the function at x^*_H is smaller than or equal 1. The function

$(\tau - T_L)/(T_H/\mu - T_L)$ is also increasing on $[x_L^*, x_H^*]$, is equal to 0 at x_L^*, to 1 at x_H^* and is greater than 1 at $\widetilde{x}$. It follows that the equation

$$\frac{\tilde{f}}{f(x)} \exp\left[\frac{\mu}{T_H}(x - \widetilde{x})\right] = \frac{\tau(x) - T_L}{T_H/\mu - T_L} \tag{25}$$

has at least one solution in $[x_L^*, \min(x_H^*, \widetilde{x})]$. We define $\widehat{x} < \widetilde{x}$ as the highest root in this interval.

By the same reasoning as above (using the log-concavity of F), we obtain that σ^* is decreasing on $[x_L^*, \widehat{x}]$. Since σ^* is constant on $[a, x_L^*]$, σ^* takes its values in $[0, 1]$, and we have: $\sigma^* \geq (\tau - T_L)/(T_H/\mu - T_L)$ on $[a, \widehat{x}]$ and $\sigma^* \leq (\tau - T_L)/(T_H/\mu - T_L)$ on $[\widehat{x}, \widetilde{x}]$.

We have

$$\begin{aligned}
\int_a^{x_L^*} (1 - \sigma^*(x)) f(x) \mathrm{d}x &= (1 - \hat{\sigma}) \left(\hat{f}/f_L^*\right) . \exp\left[\frac{1}{T_L}(x_L^* - \widehat{x})\right] .F_L^* \\
&= T_L(1 - \hat{\sigma})\hat{f} \exp\left[\frac{1}{T_L}(x_L^* - \widehat{x})\right] \\
&= T_L\,(1 - \sigma(x_L^*))\, f(x_L^*),
\end{aligned}$$

which yields $\pi_L'(x_L^*) = 0$. Since, by construction of σ^*, π_L' is constant on $[x_L^*, \widehat{x}]$, we also have: $\pi_L' = 0$ on that interval. In particular: $\pi_L'(\widehat{x}) = 0$. The definition of $\widehat{x}$ then implies: $\pi_H'(\widehat{x}) = 0$. Now, by construction of σ^*, π_H' is constant on $[\widehat{x}, \widetilde{x}]$, which, in turn, yields: $\pi_H' = 0$ on that interval. As in case (i) and (ii), we have $\sigma^* = 1$ on $[\widetilde{x}, b]$, and π_H is concave and nonincreasing on that interval.

Again, using σ constant on $[a, x_L^*]$, it is easy to check that π_L and π_H are nondecreasing on that interval. On $[x_L^*, \widehat{x}]$, we have, by using $\pi_L' = 0$:

$$\begin{aligned}
-\int_a^x \sigma^* f(t)\mathrm{d}t + \frac{T_H}{\mu}\sigma^*(x)f(x) &= -F(x) + \int_a^x (1 - \sigma^*)f(t)\mathrm{d}t + \frac{T_H}{\mu}\sigma^*(x)f(x) \\
&= -F(x) + T_L(1 - \sigma^*(x))f(x) + \frac{T_H}{\mu}\sigma^*(x)f(x) \\
&= -F(x) + T_L f(x) + \sigma^*(x)f(x)\left(\frac{T_H}{\mu} - T_L\right) \\
&= f(x)(T_H/\mu - T_L)\left[\sigma^*(x) - \frac{\tau(x) - T_L}{T_H/\mu - T_L}\right] \\
&\geq 0,
\end{aligned}$$

implying that π_H is nondecreasing on $[x_L^*, \widehat{x}]$. By construction, π_H is constant on $[\widehat{x}, \widetilde{x}]$, and is concave and nonincreasing on $[\widetilde{x}, b]$. It follows that π_H attains its maximum at $\widehat{x}$ and $\widetilde{x}$.

As to π_L, we already know that it is nondecreasing on $[a, x_L^*]$ and constant on $[x_L^*, \widehat{x}]$. Now on $[\widehat{x}, \widetilde{x}]$, the same computation as in case (i) shows, by using $\pi_H' = 0$, that

$$-\int_a^x (1 - \sigma^*)f(t)\mathrm{d}t + T_L(1 - \sigma^*(x))f(x) = (T_H/\mu - T_L)f(x)\left[\sigma^*(x) - \frac{\tau(x) - T_L}{T_H/\mu - T_L}\right] \leq 0$$

implying that π_L is nonincreasing on $[\widehat{x}, \widetilde{x}]$. Finally on $[\widetilde{x}, b]$, we know that $\sigma^* = 1$, so π_L is affine and decreasing, which shows that π_L attains its maximum at $\widehat{x}$.

□

H Proof of Proposition 6

If $x_H^* = x_L^*$, we have, by Proposition 2: $\widehat{x} = x_H^* = x_L^*$. In the two benchmark situations, plaintiffs settle if and only if their type is below $\widehat{x}$, and the probability of settlement is $F(x_H^*) = F(x_L^*) = F(\widehat{x})$. When both technologies are available, plaintiffs with type $x \leq \widehat{x}$ continue to settle; but plaintiffs with type $x \in [\widehat{x}, \widetilde{x}]$ now settle when they invest and receive the high offer $\mu\widetilde{x} - t_H^P$, which occurs with probability $\sigma(x).[1 - 1/\mu]$. The probability of settlement is, therefore, increased by $(1 - 1/\mu) \int_{\widehat{x}}^{\widetilde{x}} \sigma(x) f(x) \mathrm{d}x$, which is positive, since $\sigma(\widetilde{x}) = 1$ and σ is continuous at $\widetilde{x}$ (see Appendix E).

Under the assumptions of the Proposition, the total costs simplify into $\mathcal{C}_{\{HL\}}^* = c \Pr(e = H) + [F(\widetilde{x}) - F(\widehat{x})]T_L + [1 - F(\widetilde{x})]T_H$. Comparing with (2) yields $\mathcal{C}_{\{L\}}^* < \mathcal{C}_{\{HL\}}^* < \mathcal{C}_{\{H\}}^*$.

It is straightforward to check that the plaintiff's payoff in $\{HL\}$, which is represented on Figure 5, is uniformly greater than his payoffs $v_{\{L\}}(x; \widehat{x})$ and $v_{\{H\}}(x; \widehat{x})$ in the one-technology worlds. It follows that the plaintiff, whatever his type, prefers $\{HL\}$ to both $\{H\}$ and $\{L\}$.

Starting from $\{L\}$, the introduction of the costly technology reduces total welfare and benefits the plaintiff (irrespective of his type); it must therefore harm the defendant. Starting from $\{H\}$, the introduction of the basic technology raises total welfare; so it might benefit both parties.

Finally we prove item ii of the proposition: when $T_L \leq x_L^* = x_H^* < \widetilde{x}$, the defendant prefers $\{HL\}$ to $\{H\}$. Equation (23) of Appendix F yields the expected payoff of the defendant in equilibrium when both technologies are available:

$$\Pi_{\{HL\}}^* = -(\widehat{x} - t_L^P) \int_a^{\widehat{x}} f(x) \mathrm{d}x - \int_{\widehat{x}}^{\widetilde{x}} [x + t_D] f(x) \mathrm{d}x - \int_{\widetilde{x}}^{b} [\mu x + t_D] f(x) \mathrm{d}x - T_L f(\widetilde{x}) c.$$

When only the H technology is available, the defendant's profit is

$$\Pi_{\{H\}}^* = -(\mu\widehat{x} - t_H^P) \int_a^{\widehat{x}} f(x) \mathrm{d}x - \int_{\widehat{x}}^{b} [\mu x + t_D] f(x) \mathrm{d}x$$

whence

$$\begin{aligned} \Pi_{\{HL\}}^* - \Pi_{\{H\}}^* &= \left(\mu\widehat{x} - t_H^P - (\widehat{x} - t_L^P)\right) \int_a^{\widehat{x}} f(x) \mathrm{d}x + (\mu - 1) \int_{\widehat{x}}^{\widetilde{x}} x f(x) \mathrm{d}x - T_L f(\widetilde{x}) c \\ &+ (t_D - t_D) \left(F(\widetilde{x}) - F(\widehat{x})\right) \\ &= (\mu - 1) (\widehat{x} - T_L) \int_a^{\widehat{x}} f(x) \mathrm{d}x + (\mu - 1) \int_{\widehat{x}}^{\widetilde{x}} x f(x) \mathrm{d}x - T_L f(\widetilde{x}) c \\ &+ (t_D - t_D) F(\widetilde{x}) \end{aligned}$$

Using $\widetilde{x} = (t_H^P - t_L^P + c)/(\mu - 1) = T_L + c/(\mu-1) - (t_D - t_D)/(\mu - 1)$ and $F(\widetilde{x}) > T_L f(\widetilde{x})$, we get

$$\begin{aligned}
\frac{\Pi^*_{\{HL\}} - \Pi^*_{\{H\}}}{\mu - 1} &\geq (\widehat{x} - T_L)\, F(\widehat{x}) + \int_{\widehat{x}}^{\widetilde{x}} x f(x) \mathrm{d}x - T_L f(\widetilde{x})(\widetilde{x} - T_L) \\
&= (\widetilde{x} - T_L)\, F(\widetilde{x}) + \int_{\widehat{x}}^{\widetilde{x}} [-F(x) + T_L f(x)] \mathrm{d}x - T_L f(\widetilde{x})(\widetilde{x} - T_L) \\
&= \int_{\widehat{x}}^{\widetilde{x}} [-F(x) + T_L f(x)]\, \mathrm{d}x + (\widetilde{x} - T_L)\, [F(\widetilde{x}) - T_L f(\widetilde{x})]\,.
\end{aligned}$$

Using the log-concavity of F, it is easy to check that the function $-F + T_L f$ is nonincreasing on $[\widehat{x}, \widetilde{x}]$, which yields (using the assumption $\widehat{x} = x_L^* > T_L$)

$$\begin{aligned}
\frac{\Pi^*_{\{HL\}} - \Pi^*_{\{H\}}}{\mu - 1} &\geq (\widetilde{x} - \widehat{x})[-F(\widetilde{x}) + T_L f(\widetilde{x})] + (\widetilde{x} - T_L)\, [F(\widetilde{x}) - T_L f(\widetilde{x})] \\
&= (\widehat{x} - T_L)\, [F(\widetilde{x}) - T_L f(\widetilde{x})] \\
&\geq 0.
\end{aligned}$$

I The continuous model

I.1 Proof of Proposition 7

(i) Assume that the defendant infers a plaintiff's type x_0 from the observed effort, that is, only x_0 is choosing this effort in equilibrium. She therefore offers (with probability one) the settlement amount that leaves plaintiff x_0 indifferent between accepting the offer and going to trial. If this plaintiff has chosen effort $e^\star(x)$, he therefore gets utility $\mu(e^\star(x))x_0 - t^P(e^\star(x)) - c(e^\star(x))$, which is lower than or equal to $v_0(x_0)$, with equality only for $x = x_0$. It follows that x_0 can separate only by investing $e^\star(x_0)$. The settlement offer after $e^\star(x_0)$ would thus attract all types between a and x_0. For x_0 to separate with $e^\star(x_0)$, it is then necessary that all types between a and x_0 earn exactly $v_0(x_0)$. However, as $v(x)$ is convex, such a flat can occur only once. Consequently, only one type (if any) is able to separate in equilibrium.

(ii) We need the following result that extends Lemma 1 to the continuous case.

Lemma I.1. *The plaintiff never chooses an effort below his perfect information effort.*

Proof. We first show that, after observing effort e, the defendant makes no offer greater than $\mu(e)(e^\star)^{-1}(e) - t^P(e)$. Formally: $v(x|e^\star(x)) = v_0(x)$ for all x.

To this aim, we use an unraveling argument. Note first that $v(x|e^\star(x)) = v_0(x)$ is obvious for $x = b$ as the defendant has no incentives to offer more than $\mu(e^\star(b))b - t^P(b)$. Now suppose that for some $x < b$, $v(x|e^\star(x)) > v_0(x)$. Then the functions $v(.|e^\star(x))$ and $v_0(.)$ would intersect at some point x_0, with $x < x_0 < b$. The defendant would know for sure that types above x_0 do not choose effort $e^\star(x)$. Then she could reduce the utility

$v(.|e^\star(x))$ by the constant amount $v(x_0|e^\star(x)) - [\mu(e^\star(x))x_0 - t^P(e^\star(x)) - c(e^\star(x))] > 0$, thereby increasing her profit.

Now $v(y|e^\star(x)) = \mu(e^\star(x)).y - t^P(e^\star(x)) - c(e^\star(x)) < v_0(y)$ for all $x < y$, showing that plaintiff y never chooses effort x. □

To prove (ii) of Proposition 7, we proceed by contradiction. Consider any finite set $\{e_1, \ldots, e_N\}$ and suppose that all plaintiffs pick their effort in this set. We define x_i by $e_i = e^\star(x_i)$. From the proof of Lemma I.1, we know that the defendant makes no offer greater than $\mu(e_i)x_i - t^P(e_i)$ after observing e_i.

First, we notice that $x_N = b$, otherwise agents whose type x lies between x_N and b would choose effort $e^\star(x)$ and go to court. For the same reason, all agents between x_{N-1} and $x_N = b$ choose effort $e_N = e^\star(b)$, implying $\sigma(e_N|.) = 1$ in $(x_{N-1}, b]$. It follows that the posterior density $f(.|e^\star(b))$ is proportional to the prior distribution f on that interval. At $x = b$, the derivative of the defendant's profit after observing $e_N = e^\star(b)$ is therefore proportional to that of the defendant profit in the Bebchuk continuation game where all plaintiffs would choose effort $e^\star(b)$. The latter derivative is negative by the assumption that the Bebchuk solution after $e^\star(b)$ is interior. It follows that the defendant's profit after $e^\star(b)$ is not maximal in a neighborhood of b: there exists $y < b$ such that the defendant makes no offer greater than $\mu(e^\star(b))y - t^P(e^\star(b))$ after observing $e^\star(b)$. It follows the plaintiffs in $(y, b]$ choose their perfect information effort (and go to court), the desired contradiction.

I.2 Proof of Proposition 8

The proof requires two intermediary lemmas.

Lemma I.2. *For any plaintiff type x, effort $e^\star(b)$ solves the program (8). In other words, $v(x) = v(x|e^\star(b))$ for all x.*

Proof. The proof is identical to that of Proposition 2. Note first that $v(b) = v(b|e^\star(b)) = v_0(b)$. Now suppose that $v(x|e^\star(b)) < v(x)$ for some $x < b$. Since the functions $v(.|e^\star(b))$ and $v(.)$ coincide at b, there exists $x_1 \leq b$ such that they coincide at x_1 and v is above $v(.|e^\star(b))$ on $[x, x_1)$. Agents whose type y lies in that interval do not choose $e^\star(b)$: $\sigma(e^\star(b)|y) = 0$. Noting $f(.|e)$ the density of the posterior distribution (after the defendant's Bayesian revision), we therefore have $f(.|e^\star(b)) = 0$ on $[x, x_1)$. Now we consider two cases, depending on $F(x|e^\star(b))$.

Suppose first that $F(x|e^\star(b)) = 0$. It follows that $F(x_1|e^\star(b)) = 0$. Then we know from Lemma 3 that $v(.|e^\star(b))$ is constant on $[a, x_1)$. But this is impossible as $v(x_1|e^\star(b)) = v(x_1)$, $v(.|e^\star(b)) < v(.)$ on a left neighborhood of x_1, and v is nondecreasing.

Consider now the other case: some agents below x choose effort $e^\star(b)$ with positive probability and $F(x|e^\star(b)) > 0$. We have:

$$-F(y|e^\star(b)) + \frac{T(e^\star(b))}{\mu(e^\star(b))} f(y|e^\star(b)) = -F(x|e^\star(b)) < 0$$

for $y \in [x, x_1)$. The above expression is proportional to the derivative of the defendant's profit after observing $e^\star(b)$. Therefore the defendant's profit after $e^\star(b)$ cannot be maximal at y and $v(.|e^\star(b))$ is affine on $[x, x_1)$. More generally, the argument shows that $v(.|e^\star(b))$ is affine as long as it is below $v(.)$ and $F(.|e^\star(b))$ is positive. It follows that v is everywhere above $v(.|e^\star(b))$ on $[a, x_1)$, which contradicts $F(x|e^\star(b)) > 0$. □

Lemma I.3. *The defendant's profit after* $e^\star(b)$ *is maximal at* b *and its derivative at* b *is zero.*

Proof. It follows from Lemma I.2 that the defendant's profit after $e^\star(b)$ is maximal at b. Indeed if it were not the case, that profit would not be not maximal on some interval $(y, b]$, with $y < b$. The function $v(z|e^\star(b))$ would therefore with the segment $\mu(e^\star(b))z - t^P(e^\star(b)) - c(e^\star(b))$ on $(y, b]$, which would contradict Lemma I.2, as this segment is strictly below v_0.

Since the defendant's profit after $e^\star(b)$ is maximal at b, its derivative at this point is nonnegative. We now suppose that the derivative is strictly positive and get a contradiction.

If the derivative is positive, the defendant's profit after $e^\star(b)$ is not maximal in a left neighborhood of b and, consequently, the function $v(.|e^\star(b))$ is affine on this neighborhood. This implies that $v(.|e^\star(b))$ is strictly above v_0 (by Lemma I.2).

We now show that, in such a configuration, any type in that neighborhood necessarily chooses effort $e^\star(b)$ with certainty. Suppose indeed that some plaintiff x in that neighborhood chooses effort $e^\star(y)$, $x \leq y < b$, with positive probability. We would have: (i) $v(x|e^\star(y)) = v(x|e^\star(b))$; (ii) $v(y|e^\star(y)) = v_0(y) < v(y|e^\star(b))$; (iii) $v(.|e^\star(b)) \geq v(.|e^\star(y))$ everywhere by Lemma I.2; (iv) $v(.|e^\star(b))$ affine around x. These four properties are incompatible with the convexity of $v(.|e^\star(y))$. It follows that all plaintiffs in a neighborhood of b choose $e^\star(b)$ with probability one.

But this, again, is impossible, as the defendant profit after $e^\star(b)$ would have the same derivative at b as the Bebchuk profit after $e^\star(b)$, implying that b cannot be optimal after $e^\star(b)$ (we have assumed that the Bebchuk game has no corner solution). It follows that the derivative of the defendant's profit after $e^\star(b)$ evaluated at b is zero. □

From Lemma I.3, we have:

$$-F(b|e^\star(b)) + \frac{T(e^\star(b))}{\mu(e^\star(b))} f(b|e^\star(b)) = 0.$$

Using Bayes rule and $\sigma(e^\star(b)|b) = 1$ yields:

$$-1 + \frac{T(e^\star(b))}{\mu(e^\star(b))} \frac{f(b)}{\text{Proba}(e^\star(b))} = 0.$$

This shows that the probability of observing $e^\star(b)$ is $f(b)T(e^\star(b))/\mu(e^\star(b))$, which is lower than one from the assumption that the Bebchuk offer after $e^\star(b)$ is interior.