The Appeals Process and Incentives to Settle

Wohlschlegel, Ansgar

14 February 2014

The Appeals Process and Incentives to Settle

Ansgar Wohlschlegel*
University of Portsmouth

February 14, 2014

Abstract

This paper analyzes asymmetrically informed litigants' incentives to settle when they anticipate the possibility of appeals. It identifies a strategic effect, which induces a litigant to negotiate *pretrial* so as to optimize her *posttrial* bargaining position, and an information effect, which means that litigants will take into account *pretrial* how the information revealed by the trial court's verdict will translate into *posttrial* equilibrium payoffs. The paper's main contribution is twofold: First, it establishes a workhorse model of settlement and litigation in the shadow of appeals which may be used in future research to analyze specific issues of litigation and legal reform. Second, the importance of including the possibility of appeals in the litigation model is highlighted by an example in which some results contradict the immediate intuition: It is shown that (i) more accurate trial courts may actually attract less cases and (ii) cases may go to trial court with a larger ex-ante probability for higher legal costs in the appeals stage.

JEL Classification: K41; K13; D82

1 Introduction

Litigants often settle out of court in order to save on legal costs. However, informational asymmetries may result in a breakdown of settlement negotiations: If a litigant has private information which makes her confident to have a strong case, she will only settle out of court if this yields her a high payoff, whereas a lower payoff is required if she is

*I am grateful for discussions with and suggestions by Eberhard Feess, Daniel Göller, Michael Hewer, Thomas Kittsteiner, Elisabeth Schulte, Urs Schweizer, Petros Sekeris and participants of the research seminar in Aachen and of the annual meetings of Verein für Socialpolitik, Association Française de Science Economique and the German Law & Economics Association.

pessimistic regarding the strength of her case. Hence, the opposing party may find it worthwhile not to settle if that litigant is optimistic in order to be able to settle at more favorable terms with a less confident litigant.[1]

This basic tradeoff exists at any stage of a legal dispute. However, the terms at which litigants negotiate for out-of-court settlement will change in the course of the litigation process: Information revealed in court may remove the informational asymmetry to pave the way for agreement, which some litigants reach as late as while the jury comes back to the courtroom to announce its verdict. Similarly, litigants may use the very fact that the opposing party rejected previous settlement offers to update their assessment of the strength of their case.

Of course, rational litigants will take into account these changes in the bargaining environment when making decisions at earlier stages. In general, the better informed litigant will avoid to settle early if, given his private information, he anticipates his bargaining position to improve over time. Furthermore, the less informed litigant will seek to settle early with the opposing party if the latter appears to have observed information which makes him unprofitable to negotiate with at later stages.

This is the first paper to analyze how anticipating future stages of appeal and settlement negotiation influences litigants' decisions at earlier stages of the legal process. In particular, a model is considered in which a plaintiff may make a take-it-or-leave-it settlement demand to a privately informed defendant. If the defendant rejects, the case goes to trial court. If appeal is filed for the trial court's verdict, the plaintiff may make another take-it-or-leave-it settlement demand, the rejection of which will bring the case to the appeals court, which is assumed to be the final stage of the litigation process.

Two main effects are identified which drive litigants' decisions in equilibrium: First, the plaintiff will anticipate that her posttrial equilibrium payoff will depend on the private information that a defendant who has rejected the pretrial settlement demand may have. When choosing the pretrial settlement demand, the plaintiff will therefore take into account for which private information the defendant will reject or accept it, and how this affects posttrial payoffs. In other words, the plaintiff will choose the pretrial settlement demand so as to optimize the strategic environment in which posttrial settlement negotiations take place. I shall label this effect the 'strategic effect'.

For instance, if the trial court's decision is purely random, it depends on the ex-ante probability distribution of which private information the defendant observes whether the equilibrium probability that settlement ever occurs with two levels of jurisdiction is higher or lower than with just one level. Specifically, the threshold private information beyond which a defendant will refuse to settle both pretrial and posttrial is shifted towards lower

[1] See Bebchuk (1984), Nalebuff (1987) and the literature following these papers.

densities of this distribution: A lower density of the marginal 'type' of defendant implies that the plaintiff has a lower marginal benefit of settling with a higher probability, which induces her to make a higher posttrial settlement demand.

The second effect is that reducing the probability of settlement by making a tougher pretrial settlement demand will improve the plaintiff's average case that actually goes to trial and thus increase the plaintiff's probability of winning in trial. Hence, the plaintiff will prefer a higher pretrial settlement demand if the difference in her equilibrium payoff after winning and losing in trial court is higher. In other words, litigants will take into account *pretrial* how the information revealed by the trial court's verdict will translate into *posttrial* equilibrium payoffs. Hence, this effect will be referred to as the 'information effect'. A typical feature of the information effect is that the plaintiff's basic cost-benefit tradeoff in posttrial settlement negotiation is less sensitive to the settlement demand if she has lost in trial court. Hence, parameter changes that *increase* the plaintiff's posttrial payoffs will *reduce* the difference in her posttrial payoffs after winning and losing in trial court and, therefore, *increase* the plaintiff's incentives to settle pretrial.

If the information effect dominates, it will thus result in counter-intuitive comparative statics. For instance, in the example discussed in Subsection 6.2, higher legal costs in the appeals stage will increase the ex-ante probability that a case goes to trial court in equilibrium. Furthermore, a trial court that predicts the appeal court's eventual judgment more accurately may be used with lower probability, as the defendant anticipates to earn a lower information rent posttrial and will, thus, tend to accept higher pretrial settlement demands.

The existing economic literature on appeals has mainly focused on how the possibility of an appeal affects judges' incentives, especially when they have career concerns.[2] The basic idea of this literature is that judges, and decision makers in general, prefer their decisions not to be reversed by an appeals instance. Hence, in Shavell (2006, 2007) and Iossa and Palumbo (2007) the threat of appeals serves as a disciplining device preventing opportunistic judges from deviating too much from the socially preferred outcome.[3] Daughety and Reinganum (2000) and Levy (2005) analyze Bayesian updating of imperfectly informed judges which seek to avoid reversal of their judgment in the first, and to impress an imperfectly informed expert in the second paper.

[2]Notable exceptions are Shavell (1995, 2010), in which social costs and benefits of having more or less levels of jurisdiction, of the accuracy of these levels and of discretionary review versus direct appeals are analyzed. Bütler and Hauser (2000) analyze settlement incentives of *symmetrically* informed litigants under the specific rules of the WTO dispute settlement system.

[3]In a similar vein, Spitzer and Talley (2000) analyze a game of judicial review when judges at both levels care about the distance of the final decision from their personal political position and about legal cost.

While none of these papers consider settlement incentives of asymmetrically informed litigants, I take a different approach by focusing on exactly this aspect and treating courts as stochastic dummy players. As a consequence, litigants in my model only care about how the appeals court will eventually judge, and the information that the trial court's verdict and the better informed litigant's actions reveal regarding this issue.

Another line of related literature deals with settlement in dynamic contexts. Robson and Skaperdas (2008) discuss the case of litigants fighting over an initially undefined property right. Since information is symmetric in that paper, the two main effects identified in the present paper cannot occur. However, posttrial bargaining will occur in equilibrium although the trial court defines binding property rights, because either litigant's individually most preferred choice of how to use the property is different than the joint surplus maximizing choice, whereas in my paper litigants just bargain for a joint-welfare neutral transfer. Although confined to pretrial settlement, Spier (1992) is closely related to the present paper as she also allows for multiple rounds of negotiation between asymmetrically informed litigants. Indeed, her two-period case with $c > 0$ is equivalent to the analysis of the purely strategic effect in Section 5 of this paper. However, as the focus of her paper is on the timing of settlement when there are multiple stages, she confines that basic analysis of the two-period case to deriving the result that there will be settlement with some positive probability in each period. By contrast, my focus is on the impact of the second stage and, in the more general part of the paper, of the information revealed by the trial court's verdict on the ex-ante settlement probability.

Last, the signalling argument that a privately informed player may delay a mutually beneficial agreement in the hope for an even better offer is well-known from the literature on sequential bargaining with asymmetric information. While incentives for delaying agreement are wiped out as offers can be made more frequently,[4] delay does occur in equilibrium whenever the time between offers is substantial, as it is the case in the present paper. For instance, Hart (1989) shows in such a setting the intuitively plausible result that higher cost of delay increase incentives to agree early. The present paper's contribution to this literature is to introduce the possibility that players observe an informative public signal between the rounds of negotiation, and to show that the well-known relationship between cost of delay and timing of agreement may reverse in such a model.

The remainder of the paper will be organized as follows: The timing and payoffs in the model will be presented in Section 2. Section 3 discusses assumptions regarding the stochastic relationship among the defendant's private information and the trial and appeals courts' verdicts in detail. Section 4 consists of a general analysis of the sub-

[4]See, for instance, Gul, Sonnenschein, and Wilson (1986)

game after the plaintiff has made her pretrial settlement demand. When analyzing the full game, I will first discuss in Section 5 the case in which the trial court's verdict is completely unrelated to how the appeals court will eventually judge, which allows me to isolate the pure strategic effect. Section 6 then deals with strictly informative trial courts. After presenting some general results, I will consider a specific class of signal technologies, which allows me to derive clear-cut comparative static results. Section 7 concludes and discusses qualifications and potential extensions of the model.

2 The Model

Consider a case in which a plaintiff sues a defendant for damages of an undisputed size D.[5] After filing suit, the plaintiff may make a take-it-or-leave-it settlement demand S_T, which I will refer to as pretrial settlement demand. The defendant then observes some private information x which allows him to update the probability that the appeals court will eventually judge in favor of the plaintiff, and chooses whether to accept the demand. To be more specific on the informational structure, let us assume that the private signal x is distributed with a full-range distribution function $F(x)$ with density $f(x)$ and monotonically increasing hazard rate $\frac{f(x)}{1-F(x)}$.

If the defendant accepts the pretrial settlement demand, he pays S_T to the plaintiff, and the game ends. If the defendant rejects the settlement demand, the case goes to trial court, which imposes litigation costs c_T^p on the plaintiff and c_T^d on the defendant.[6] The trial court awards damages equal to D or zero. Let $l_T \in \{0, 1\}$ denote the trial court's decision, where $l_T = 1$ means that damages D are awarded.

The defendant decides whether to appeal. If he does not appeal, the plaintiff may appeal.[7] If any of the litigants has appealed, the plaintiff may make another settlement demand S_A, which I refer to as posttrial settlement demand, and which the defendant may accept or reject. If he rejects, the case goes to the appeals court, which imposes additional litigation costs c_A^p on the plaintiff and c_A^d on the defendant. The damages awarded by the appeals court, if any, are again D.

[5]For instance, in an accident case, the size of the victim's harm and the injurer's negligence may be undisputed, and the case is about finding out whether the plaintiff's negligence was causal for the accident.

[6]In reality, the plaintiff may choose whether to indeed go to trial or to back down if her demand has been rejected. This raises credibility issues of settlement demands discussed in Nalebuff (1987), which are known to result in an upper bound to settlement demands. In order to avoid the case distinctions associated with the possible boundary solutions, we make this simplifying assumption.

[7]Note that this timing of the right to appeal is wlog, as it will turn out that the losing party will file appeal anyway.

How the appeals court will eventually judge is unknown to litigants, but the defendant's private information and the trial court's decision are potentially informative signals thereon. I will explain the specific assumptions on the litigants' prior information and the signal technology in Section 3. Summing up, the timing of the game is as follows:

(i) Plaintiff makes a settlement demand S_T.

(ii) Defendant privately observes x and decides whether to accept or reject the demand.

(iii) If the defendant has rejected the demand, the case goes to the trial court, and the trial court's verdict $l_T \in \{0, 1\}$ is announced.

(iv) Defendant may appeal.

(v) If defendant has not appealed, Plaintiff may appeal.

(vi) Upon appeal, plaintiff may make posttrial settlement demand S_A.

(vii) Defendant decides whether to accept the demand.

(viii) If the defendant has rejected the demand, the case goes to the appeals court, and the appeals court's verdict $l_A \in \{0, 1\}$ is announced.

The equilibrium concept used throughout is perfect Bayesian, and I will focus on pure-strategy equilibria. More specifically, note that the game has a proper subgame beginning after the plaintiff has made the pretrial settlement demand S_T. For every given S_T, we can characterize pure-strategy perfect Bayesian equilibria of the subsequent subgame by means of (i) the set of realizations of x for which the defendant rejects S_T, (ii) both litigants' appeals decisions should they have lost in trial court, (iii) the plaintiff's beliefs on what the defendant has observed given the trial court verdict l_T, (iv) the plaintiff's posttrial settlement demand S_A^r given the trial court's verdict $l_T = r$, $r = 0, 1$, and (v) the set of realizations of x for which the defendant rejects $S_A^r(S_T)$, $r = 0, 1$. Anticipating that such a perfect Bayesian equilibrium will be played in the subsequent subgame, the plaintiff will then choose S_T so as to maximize her expected payoff.

The model set out here is an extension of the standard (one-stage) litigation model, in which an unsettled case goes directly to the final instance, adding steps (iii)-(vii) to that model. Hence, it may be useful to recapitulate the result of the standard model, which is presented without proof. To this end, denote by c^p (c^d) the plaintiff's (defendant's) costs of litigating at the single court.[8]

[8] Note that in the original version of this proposition, Bebchuk (1984) makes some additional assumptions to rule out the boundary solution $x^* = 0$.

Proposition 1 (Bebchuk (1984)) *Consider a litigation model with just one round of litigation, which consists of steps (i)-(ii) and (viii) of the game set out above. This version of the game has a unique subgame perfect equilibrium, in which the case goes to court if and only if the defendant observes some* $x < x^*$ *implicitly given by*

$$D = \frac{f(x^*)}{1 - F(x^*)}(c^p + c^d) \tag{1}$$

if $D \geq \frac{f(0)}{1-F(0)}(c^p + c^d)$*, and* $x^* = 0$ *otherwise.*

The intuitive trade-off that the plaintiff faces when making a settlement demand is that reducing the set of types of defendant with which a settlement is reached increases the amount at which the case is settled but comes at additional expected litigation costs represented by the right-hand side of (1). Note that the assumption of an increasing hazard rate $\frac{f(x)}{1-F(x)}$ is sufficient to guarantee uniqueness of the interior equilibrium given by (1).

3 The Signal Technology

In this section, I will be more specific about how the trial and the appeals courts' outcomes and the defendant's private signal x are related to each other. Let the set of all states of nature be partitioned into two mutually exclusive and jointly exhaustive events l_A which refer to the appeals court's eventual judgement: In the event denoted by $l_A = 1$ ($l_A = 0$), the appeals court will deterministically judge in favor of the plaintiff (defendant). The state of nature is unknown to litigants, who just know ξ, the unconditional ex-ante probability of event $l_A = 1$.

Two noisy signals on the state of nature may be observed in the course of the game: First, after the plaintiff has submitted her pretrial settlement demand, the defendant privately observes a real number $x \in [0, 1]$ as a noisy signal on the true state of nature. Let the informativeness of this signal be such that, conditional on the signal x being observed, the defendant can update the probability that the true state of nature is in event $l_A = 1$ from ξ to x.[9] Assuming that the ex-ante probability that the defendant observes signal x is distributed continuously on $[0, 1]$ with density $f(\cdot)$, consistency requires that $\xi = \int_0^1 x f(x) dx$.

Second, the trial court's verdict $l_T \in \{0, 1\}$ is a noisy public signal on the state of nature and, therefore, on how likely each event l_A is. In order to be able to analyze litigants' Bayesian updating after observing l_T, I need to define its accuracy in predicting

[9] For convenience, I will sometimes refer to a defendant who has observed the private signal x as a 'type-x defendant'.

l_A. As I want to accommodate the case in which the two signals are correlated even conditional on the event l_A in which the true state of nature is, the notation must allow for the public signal's accuracy to be a function of x, and to vary across events l_A. In particular, let the accuracy of the public signal when the defendant's private signal is x and the true state of the world is in event $l_A = r$ be $p_r(x)$. That is to say, if $l_A = r$, the probability that the trial court's verdict correctly anticipates the appeals court's when the defendant has observed x is $p_r(x)$. In order to rule out trivial signal technologies, assume that for all r and x, $0 < p_r(x) < 1$.

This informational structure implies litigants' posterior beliefs upon observing the signals: Recall that the plaintiff's ex-ante beliefs for the distribution of the defendant's private signal x has density $f(\cdot)$. After observing the public signal, she may update the density of the defendant's types, as the public signal may be correlated with the defendant's private signal. In particular, her updated belief on the defendant's private signal when observing the public signal $l_T = r$ have density $y_r(x)f(x)$, where

$$y_r(x) := Prob(l_T = r \mid x) \tag{2}$$

denotes the overall (i.e., unconditional on l_A) probability of a public signal (trial court verdict) $l_T = r$ given the defendant's private information x. Using Bayes' rule, we get

$$\begin{aligned} y_1(x) &= p_1(x)x + (1 - p_0(x))(1 - x) && (3)\\ y_0(x) &= p_0(x)(1 - x) + (1 - p_1(x))x = 1 - y_1(x). && (4) \end{aligned}$$

The defendant may also use the observed public signal to update his beliefs on the probabilities of events l_A. In particular, let

$$z_r(x) := Prob(l_A = 1 \mid x, l_T = r) \tag{5}$$

denote the probability of the true state of nature being in event $l_A = 1$ conditional on the realizations of the private signal x and the public signal $l_T = r$. Then $z_r(x)$ is a type-x defendant's posterior of $l_A = 1$ after observing $l_T = r$. Using Bayes' rule, we can express $z_r(\cdot)$ in terms of the accuracy $p_r(\cdot)$:

$$\begin{aligned} z_1(x) &= \frac{Prob(l_A = 1 \wedge l_T = 1 \mid x)}{Prob(l_T = 1 \mid x)} = \frac{p_1(x)x}{p_1(x)x + (1 - p_0(x))(1 - x)} && (6)\\ z_0(x) &= \frac{Prob(l_A = 1 \wedge l_T = 0 \mid x)}{Prob(l_T = 0 \mid x)} = \frac{(1 - p_1(x))x}{p_0(x)(1 - x) + (1 - p_1(x))x}. && (7) \end{aligned}$$

Having defined a signal technology in a most general way, I shall now restrict generality in two respects: First, for simplicity, I will focus on signal technologies with continuous and differentiable functions $p_r(\cdot)$, which implies that $y_r(\cdot)$ and $z_r(\cdot)$ are also continuous and differentiable. Second, the following plausible properties of the signal technology are assumed:

Assumption 1 *(a)* $y_1(\cdot)$ *non-decreasing: Defendants who observed lower* x *are no more likely to win in the trial court than those who observed higher* x*.*

(b) $z_r(\cdot)$ *strictly increasing,* $r = 0, 1$*: Given any verdict of the trial court, defendants who observed higher* x *are strictly more likely to win in the appeals court than those who observed lower* x*.*

(c) $\forall x : p_1(x) + p_0(x) \geq 1$*: This assumption is equivalent to* $z_0(x) \leq z_1(x)$ *for all* x*, which means that given the defendant's* private *signal, the* public *signal is informative in the sense that a defendant who has lost in trial court can never expect to be more likely to win in the appeals court than if he had won in trial court.*

The signal technologies that satisfy Assumption 1 include a number of prominent special cases some of which I will now briefly discuss by formalizing the public signal's accuracy and then using (3), (4), (6) and (7) to derive parties' posterior beliefs. The most common case in the literature on aggregating informative signals, such as Ottaviani and Sørensen (2001) or Gerardi and Yariv (2008), is that the signals are drawn independently. In this case, the public signal's accuracy neither depends on the defendant's private signal x nor on the true event $l_A = r$:

Example 1 (Independent Signals) *If signals are drawn independently, then* $p_1(x) = p_0(x) = \rho > \frac{1}{2}$*, which is some constant. It follows that*

$$\begin{aligned} y_1(x) &= \rho x + (1-\rho)(1-x) \\ y_0(x) &= \rho(1-x) + (1-\rho)x \\ z_1(x) &= \frac{\rho x}{\rho x + (1-\rho)(1-x)} \\ z_0(x) &= \frac{(1-\rho)x}{\rho(1-x) + (1-\rho)x}. \end{aligned}$$

From the plaintiff's perspective, the public signal $l_T = 1$ may be true, which occurs with ex-ante probability ρx, or false, which occurs with probability $(1-\rho)(1-x)$. The defendant updates his private information by dividing the probability that the public signal $l_T = 1$ is true by the total ex-ante probability of the signal $l_T = 1$.

A polar case is that the trial court just rolls dice, i.e. where its verdict does not contain any information on the true state of nature. This implies that any signal $l_T = r$ is sent with some constant probability which is neither related to the defendant's private signal nor the event l_A in which the true state of nature is.

Example 2 (Uninformative Public Signal) *If the public signal is completely uninformative, then* $p_1(x) = 1 - p_0(x) = \rho$*, which is some constant. It follows that*

$$
\begin{aligned}
y_1(x) &= \rho \\
y_0(x) &= 1 - \rho \\
z_1(x) &= x = z_0(x).
\end{aligned}
$$

As the probability that a particular public signal $l_T = r$ is sent neither depends on the true event l_A nor on the defendant's private information x, no party can use the signal to update information.

A third example is a case which I will refer to as the public signal being based on the private signal but otherwise random, that is to say, the trial court judges with exactly those probabilities given by the defendant's private signal, but completely randomizes given these probabilities. In this example, the defendant cannot infer any new information from the public signal, but, subject to this restriction, the public signal is as informative as possible for the plaintiff. This is the case whenever the public signal $l_T = 1$ is sent with probability x independent of the true event l_A.

Example 3 (Randomizing Based on Private Signal) *If the public signal is randomly drawn on the basis of the defendant's private signal, then* $p_1(x) = 1 - p_0(x) = x$*. It follows that*

$$
\begin{aligned}
y_1(x) &= x \\
y_0(x) &= 1 - x \\
z_1(x) &= x = z_0(x).
\end{aligned}
$$

Just like in the case of the completely uninformative public signal, the probability of each signal l_T does not depend on the true event, i.e. $p_1(x) = 1 - p_0(x)$. Hence, the defendant's posterior beliefs on the probability distribution of the events l_A is the same as before observing the public signal. However, as the trial court sends the public signal with the correct probabilities privately known by the defendant, the plaintiff can update her beliefs on the defendant's private information. In particular, the posterior density function with which the plaintiff believes the defendant's types to be distributed after observing the public signal $l_T = r$ is $y_r(x)f(x)$, which is clearly different from her ex-ante beliefs $f(x)$.

4 General Analysis

As the game has a proper subgame starting after the pretrial settlement demand S_T has been made, I will analyze the game using backward induction. Hence, most of the following analysis will be performed for some given S_T. Let us start with the plaintiff's choice of posttrial settlement demand S_A after a pretrial settlement demand S_T has been made by the plaintiff and rejected by the defendant, the trial court has made a verdict $l_T = r$, $r = 0, 1$, and appeal has been filed. The plaintiff believes that the defendant's private information is distributed on $[0, 1]$ with some density $\mu_r(x)$ with support $M_r \subseteq [0, 1]$. Furthermore, the defendant may have used the trial court's verdict to update his private information x, so that he expects to be held liable by the appeals court with probability $z_r(x)$.

The concept of perfect Bayesian equilibrium then requires that, given these updated expectations, players maximize their payoffs. Hence, the analysis of the posttrial settlement negotiations is similar to the standard screening model of litigation: If a type-x defendant rejects the settlement demand, his expected payoff is $-z_r(x)D - c_T^d - c_A^d$, which is strictly decreasing in x, whereas accepting a posttrial settlement demand S_A gives him $-S_A - c_T^d$, which is constant in x. Hence, if any type of defendant accepts the settlement demand, it is someone who has observed a high probability x of losing in the appeals court. The plaintiff anticipates the defendant's optimal strategy when choosing her posttrial settlement demand. The following Lemma states that in any perfect Bayesian equilibrium and for each outcome r of the trial, there exists a type of defendant x_A^r, $r = 0, 1$, who is indifferent between accepting and rejecting the equilibrium posttrial settlement demand, and whose equilibrium strategy has been to reject the pretrial demand S_T.

Lemma 1 *In any pure-strategy perfect Bayesian equilibrium, the plaintiff's posttrial settlement demand after a trial court's verdict $l_T = r$, $r = 0, 1$ is $S_A^r = z_r(x_A^r)D + c_A^d$ with $x_A^r < \sup M_r$ and*

$$x_A^r \in arg \max_{x_A \in M_r} \int_{M_r \cap [0, x_A]} (z_r(x)D - c_A^p)\mu_r(x)dx + (z_r(x_A)D + c_A^d) \int_{M_r \cap [x_A, 1]} \mu_r(x)dx. \quad (8)$$

Proof. As $z_r(\cdot)$ are strictly increasing in x due to Assumption 1, there exists, for every S_A and r, a unique $\tilde{x}_A^r(S_A)$ such that $S_A = z_r(\tilde{x}_A^r(S_A))D + c_A^d$. Suppose that this $\tilde{x}_A^r(S_A) \notin M_r$, and define $M_r^{acc}(S_A) := \{x \in M_r : x > \tilde{x}_A^r(S_A)\}$ the set of all private signals x such that a defendant who has observed x appears in the posttrial settlement negotiation stage with positive probability in equilibrium and accepts the posttrial settlement offer S_A.

If $M_r^{acc}(S_A) \neq \emptyset$, then $S_A' = z_r(\inf M_r^{acc}(S_A))D + c_A^d$ yields the plaintiff a strictly higher expected payoff than S_A, as it will be accepted with the same probability and $S_A' > S_A$. If, on the other hand, $M_r^{acc}(S_A) = \emptyset$, this means that S_A will be rejected with probability 1. Then there exists a sufficiently small $\varepsilon > 0$ such that $[\sup M_r - \varepsilon, \sup M_r) \subset M_r$ and $S_A'' = z_r(\sup M_r - \varepsilon)D + c_A^d$ yields the plaintiff a strictly higher expected payoff than S_A.

The plaintiff's expected payoff with this S''_A is

$$\int_{M_r} (z_r(x)D - c_A^p)\mu_r(x)dx + \int_{\sup M_r - \varepsilon}^{\sup M_r} [c_A^d + c_A^p - (z_r(x) - z_r(\sup M_r - \varepsilon))D]\mu_r(x)dx,$$

whereas her payoff with a never-accepted settlement demand is just the first summand thereof, $\int_{M_r}(z_r(x)D - c_A^p)\mu_r(x)dx$. Note that this also proves $x_A^r < \sup M_r$. ■

Lemma 1 greatly simplifies the subsequent analysis by allowing to transform the plaintiff's problem of optimizing the posttrial settlement demand into one of optimizing the marginal type of defendant who will reject or accept the posttrial settlement demand, just like the literature on the single-stage model usually proceeds. However, that this is possible in this model is not trivial, as there may be density holes in the plaintiff's beliefs, i.e. the support of the plaintiff's beliefs M_r may not be an interval. Furthermore, Lemma 1 proves that it cannot be optimal for the plaintiff to make a settlement demand that is rejected with certainty: Making a settlement demand that will be accepted only just by the highest type of defendant from the set M_r yields the plaintiff a settlement payment that is only marginally smaller than what expected damages from these types would have been, whereas the litigation cost savings are substantial.

This latter implication of Lemma 1, that it is optimal for the plaintiff to settle with some types of defendant posttrial, also implies that the highest types of defendant which the plaintiff believes to be facing always pay less than what they expect to pay after a verdict by the appeals court. To be more specific, given a trial court's verdict $l_T = r$, all types $x \geq x_A^r$ have exactly the same payoff in the posttrial stage. The next Lemma shows that this property implies that a losing defendant always files appeal: If the plaintiff believes that some type $x \geq x_A^r$ of defendant has filed appeal, she must do so for all types of defendant who had turned down the pretrial settlement demand.

The following Lemma shows that this is true also for a losing plaintiff, although this latter result is an artefact of the simplifying assumption that the case directly proceeds to court after the defendant has rejected a settlement demand (which means that the settlement demand is always credible): The plaintiff always prefers to file appeal and then appropriate the defendant's cost savings in the settlement.

Lemma 2 *In any pure-strategy perfect Bayesian equilibrium, the litigant who has lost in trial files appeal.*

Proof. Suppose that $l_T = 1$. Without appeal the defendant's payoff is $-D - c_T^d$ with certainty. With Lemma 1, the defendant's expected payoff from filing appeal is $-z_1(\min\{x, x_A^1\})D - c_A^d - c_T^d$. Hence, the defendant files appeal if and only if

$$z_1(\min\{x, x_A^1\})D + c_A^d \leq D. \tag{9}$$

In a perfect Bayesian equilibrium in which, in the posttrial settlement negotiation, the plaintiff believes that the defendant has observed $x \in M_1$, consistency of beliefs requires that condition (9) is satisfied for all $x \in M_1$. Furthermore, with Lemma 1, $x_A^1 \in M_1$. Hence, all $x \in [0,1]$ satisfy (9).

Suppose now that $l_T = 0$. If appeal is not filed, the plaintiff's payoff is $-c_T^p$ with certainty. Upon filing appeal, she can always secure herself a payoff $c_A^d - c_T^p$ by setting $x_A = \inf M_0$. Hence, filing appeal must be optimal also under the optimal posttrial settlement strategy. ■

An immediate implication of Lemma 2 is that the plaintiff's equilibrium beliefs when her settlement demand is due do not depend on the trial court's outcome, i.e. $M_0 = M_1 =: M$. Hence, when restricting attention to pure-strategy equilibria, consistent beliefs on x have density

$$\mu_r(x) := \frac{y_r(x)f(x)}{\int_{x' \in M} y_r(x')f(x')dx'} \tag{10}$$

if $x \in M$.

Consider now a defendant's decision of whether to accept the plaintiff's pretrial settlement demand S_T. In a perfect Bayesian equilibrium, the defendant will take the plaintiff's equilibrium beliefs with density $\mu(\cdot)$ and the resulting equilibrium posttrial settlement demands S_A^r as given. Recall from Lemma 1 that the posttrial settlement demands S_A^r can be expressed in terms of the marginal types of defendant x_A^r accepting this demand. Hence, every pure-strategy Bayesian equilibrium of the subgame following a pretrial settlement demand S_T can be completely characterized by the triple (x_A^0, x_A^1, M).

More specifically, recall that the defendant anticipates pretrial to be held liable in trial court with probability $y_1(x)$. Hence, defining

$$S(x) := \sum_r y_r(x) z_r\left(\min\{x, x_A^r\}\right), \tag{11}$$

the defendant's expected payoff when going to trial is

$$\Pi^d(x) = -S(x)D - c_A^d - c_T^d. \tag{12}$$

The defendant will reject the pretrial settlement demand S_T if and only if $-S_T < \Pi^d(x)$. The plaintiff's beliefs are consistent if and only if $M = \{x : -S_T < \Pi^d(x)\}$. Furthermore, Lemma 1 requires that $x_A^0, x_A^1 \in M$. These observations imply that any pure-strategy perfect Bayesian equilibrium must take one of the three forms set out in the following proposition:

Proposition 2 *Consider the subgame after the plaintiff has made a settlement demand S_T. In any pure-strategy perfect Bayesian equilibrium one of the following statements is true:*

(i) $z_0(x_A^0) < z_1(x_A^1)$ *and there is some* $x_T > \max\{x_A^0, x_A^1\}$ *such that the defendant rejects* S_T *if and only if* $x \leq x_T$.

(ii) $z_0(x_A^0) = z_1(x_A^1)$, $S(x_A^0)D + c_A^d + c_T^d \leq S_T$ *and* $[0, x_A^0] \subseteq M$.

(iii) $z_0(x_A^0) > z_1(x_A^1)$, $[0, x_A^1] \subset M$ *and* $[x_A^0, 1] \subset M$.

Proof. Using (4), (6) and (7), we can write (11) as

$$S(x) = \begin{cases} x, & \text{if } x \leq \min\{x_A^0, x_A^1\}; \\ (1 - y_1(x))z_0(x_A^0) + y_1(x)z_1(x), & \text{if } x_A^0 < x < x_A^1; \\ (1 - y_1(x))z_0(x) + y_1(x)z_1(x_A^1), & \text{if } x_A^1 < x < x_A^0; \\ z_0(x_A^0) + y_1(x)(z_1(x_A^1) - z_0(x_A^0)), & \text{if } x \geq \max\{x_A^0, x_A^1\}. \end{cases} \tag{13}$$

Part (i): If $z_0(x_A^0) < z_1(x_A^1)$, then $S(\cdot)$ is strictly increasing in x: This is obvious for the first and the last case in (13); if $x_A^0 < x < x_A^1$ then[10] $S'(x) = y_1(x)z_1'(x) + y_1'(x)(z_1(x) - z_0(x_A)) > y_1(x)z_1'(x) + y_1'(x)(z_1(x_A^0) - z_0(x_A^0)) \geq 0$; and if $x_A^1 < x < x_A^0$ then $S'(x) = (1 - y_1(x))z_0'(x) + y_1'(x)(z_1(x_A^1) - z_0(x)) > (1 - y_1(x))z_0'(x) + y_1'(x)(z_1(x_A^1) - z_0(x_A^0)) > 0$.

Hence, $\Pi^d(\cdot)$ is strictly decreasing in x, which implies that if any type of defendant rejects S_T, it will be those who observed low x. Finally, recall that defendants who observed x_A^0 or x_A^1 must reject S_T in equilibrium, which completes the proof by showing that there are some types at all that reject S_T.

Part (ii): Note first that $z_0(x_A^0) = z_1(x_A^1)$ implies $x_A^1 \leq x_A^0$. Hence, the second case in (13) is the empty set. Furthermore, $S(\cdot)$ is flat for $x \geq \max\{x_A^0, x_A^1\} = x_A^0$. Last, if $x_A^1 < x < x_A^0$, then $S'(x) = (1 - y_1(x))z_0'(x) + y_1'(x)(z_1(x_A^1) - z_0(x)) > (1 - y_1(x))z_0'(x) + y_1'(x)(z_1(x_A^1) - z_0(x_A^0)) > 0$. Summing up, $S(\cdot)$ is strictly decreasing in x if $x < x_A^0$ and constant otherwise.

As a consequence, if $S(x_A^0)D + c_A^d + c_T^d > S_T$, then $x_A^0 \notin M$, a contradiction to consistency of beliefs. Hence, $S(x_A^0)D + c_A^d + c_T^d \leq S_T$.

Part (iii): $z_0(x_A^0) > z_1(x_A^1)$ implies $x_A^1 < x_A^0$. Hence, $S(\cdot)$ is strictly increasing in x if $x \leq \min\{x_A^0, x_A^1\} = x_A^1$ and strictly decreasing in x if $x \geq \max\{x_A^0, x_A^1\} = x_A^0$. With $x_A^0, x_A^1 \in M$, this implies that $S_T \geq \max\{S(x_A^0), S(x_A^1)\}D + c_A^d + c_T^d$. ■

If the trial court's verdict contains any information on how the appeals court will eventually judge, case (i) is intuitively most plausible: Winning in trial court indicates to the plaintiff that the defendant's private information is likely to be in the plaintiff's favour, and that the appeals court is likely to rule for the plaintiff *given* any private information of the defendant. Hence, an intuitive implication of this case is that, if the less informed litigant has won in trial court, she will be tougher in posttrial bargaining

[10] Recall that $S(\cdot)$ is continuous and differentiable in x.

and agreement is less likely to be reached than if the better informed litigant had prevailed in trial court.

However, depending on the exact functional forms implied by the signal technology, the counterintuitive case (iii), where prevailing in the trial court is seen as bad news by the plaintiff, cannot be ruled out in general. Hence, it is necessary to deal with case (iii) for the sake of completeness, and the analysis of specific signal technologies in Sections 5 and 6.2 will involve identifying which of these three cases may ever occur in equilibrium.

5 The Strategic Effect

The aim of this section is to isolate the strategic effect of anticipating posttrial settlement negotiations by assuming that the trial court's verdict does not reveal any information to litigants on the strength of their case. Referring to the discussion in Section 3, the signal technology is characterized by an ex-ante probability that the plaintiff wins in trial court that is independent of x, $y_1(x) = \rho$, and by a type-x defendant's posterior probability of being held liable in the appeals court of $z_1(x) = z_0(x) = x$ that is independent of the trial court's verdict and equal to the ex-ante probability. In this case, the trial court's verdict does not matter for the litigants' expected payoffs and the plaintiff's beliefs. Hence, case (ii) of Proposition 2 applies:

Lemma 3 *Assume that the trial court's verdict is completely uninformative. In any pure-strategy perfect Bayesian equilibrium (x_A^0, x_A^1, M) of the subgame after the plaintiff has made a settlement demand S_T, $z_0(x_A^0) = z_1(x_A^1)$, $x_A^0 = x_A^1$, $x_A^0 D + c_A^d + c_T^d \leq S_T$ and $[0, x_A^0] \subseteq M$.*

Proof. With the discussion of Example 2 in Section 3, we have, for every x, $z_0(x) = z_1(x)$. Hence, the plaintiff's posttrial optimization problem (8) is identical for both potential outcomes of the trial up to the constant probability that the plaintiff wins in trial court. Hence, $x_A^0 = x_A^1$, and $z_0(x_A^0) = z_1(x_A^1)$. The remaining claims made in the Lemma then follow immediately from Proposition 2. ■

Lemma 3 simplifies the analysis considerably: It implies that a defendant with private information $x_A := x_A^0 = x_A^1$ anticipates being indifferent between accepting and rejecting the equilibrium posttrial settlement demand later on. A defendant with private information $x < x_A$ anticipates to reject the equilibrium posttrial settlement demand and earn an even higher payoff. Hence, if a defendant who has observed $x = x_A$ rejects the *pretrial* demand S_T, so will a defendant who has observed $x < x_A$.

In a perfect Bayesian equilibrium, the plaintiff's beliefs, which are characterized by M, must be consistent with this strategy. On the other hand, x_A must solve the plaintiff's posttrial problem (8). Hence, the range of the first integral in (8) is simply $[0, x_A]$. Furthermore, the second integral in (8) is the probability that x is in M but larger than x_A. Hence, denoting the probability that x is in M as $\lambda := \int_{x' \in M} f(x')dx'$, (8) can be written as

$$x_A \in \arg\max_{x'_A \in M} \int_0^{x'_A} (xD - c_A^p) \frac{f(x)}{\lambda} dx + (x'_A D + c_A^d) \left(1 - \frac{F(x'_A)}{\lambda}\right). \tag{14}$$

The objective function is continuous and differentiable in x'_A, and, due to the monotonically increasing hazard rate of $F(\cdot)$, there is a unique maximum given by the first-order condition

$$D = \frac{f(x_A)}{\lambda - F(x_A)} (c_A^p + c_A^d) \tag{15}$$

if $D \geq \frac{f(0)}{\lambda - F(0)}(c_A^p + c_A^d)$, and $x_A = 0$ otherwise.

Note the similarity to the first-order condition (1) of the standard model of litigation with just a single stage - indeed, the conditions are identical for $\lambda = 1$, $c_A^p = c^p$ and $c_A^d = c^d$. Of course, the intuitive tradeoff carries over from the single-stage model: If an indifferent defendant's private information x_A is more favorable, this allows the plaintiff to settle for a larger amount $x_A D$. However, increasing x_A comes at the cost of litigation with the marginal type of defendant, represented by the right-hand side of (15).

The following lemma states that, for every S_T, the equilibrium cutoff type of defendant for posttrial settlement, $x_A(S_T)$, and the equilibrium probability (unconditional on x) of rejection of the pretrial demand, $\lambda(S_T)$, are unique:

Lemma 4 *Let $\overline{x}_A$ be the unique x_A that solves* (14) *for $\lambda = 1$. For every $S_T \geq c_A^d + c_T^d$, there exist unique $x_A(S_T)$ and $\lambda(S_T)$ such that for every pure-strategy perfect Bayesian equilibrium (x_A^0, x_A^1, M) of that subgame, $x_A(S_T) = x_A^0 = x_A^1 = \min\left\{\frac{S_T - c_A^d - c_T^d}{D}, \overline{x}_A\right\}$, $\lambda(S_T) = \int_M f(x)dx$ is the ex-ante probability that the defendant rejects S_T, and the plaintiff's expected payoff in that subgame is*

$$\begin{aligned} \Pi^p(S_T) = & \int_0^{x_A(S_T)} (xD - c_A^p - c_T^p) f(x) dx \\ & + (\lambda(S_T) - F(x_A(S_T)))(x_A(S_T)D + c_A^d - c_T^p) + (1 - \lambda(S_T))(x_A(S_T)D + c_A^d + c_T^d). \end{aligned} \tag{16}$$

Proof. In any perfect Bayesian equilibrium of the subgame following a pretrial settlement demand S_T, x_A is a solution to the plaintiff's posttrial optimization problem (14) given λ, $\lambda \in [F(x_A), 1]$ if $x_A D + c_A^d + c_T^d = S_T$, and $\lambda = 1$ if $x_A D + c_A^d + c_T^d < S_T$. Note first that the solution x_A to (14) is increasing in λ due to the increasing hazard rate of $F(\cdot)$. Hence, the largest x_A that can ever be a solution to (14) is $\overline{x}_A$, which solves (14) for $\lambda = 1$. Hence, if $S_T > \overline{x}_A D + c_A^d + c_T^d$, there is no $\lambda \leq 1$ such that $x_A = \frac{S_T - c_A^d - c_T^d}{D}$ solves

(14), so that $S_T > x_A D + c_A^d + c_T^d$. However, this implies $M = [0,1]$ and, therefore, $\lambda = 1$, in which case the unique solution to (14) is $x_A = \overline{x}_A$.

Suppose now that $S_T \leq \overline{x}_A D + c_A^d + c_T^d$. In this case, the equilibrium $x_A(S_T)$ is uniquely given by $x_A(S_T) = \frac{S_T - c_A^d - c_T^d}{D}$: If x_A were below that, this would imply $\lambda = 1$, in which case the unique solution to (14) would be $\overline{x}_A > x_A$, a contradiction. Furthermore, there is a unique λ such that $x_A(S_T) = \frac{S_T - c_A^d - c_T^d}{D}$ solves (14).

To complete the proof, the plaintiff's equilibrium payoff given by (16) is obtained by summing up the posttrial payoff (14) with probability $\lambda(S_T)$ and the payoff from pretrial settlement S_T with probability $1 - \lambda(S_T)$. ■

If $x_A(S_T)D + c_A^d + c_T^d < S_T$, then the defendant will reject S_T no matter what his private information is, which means that $\lambda(S_T) = 1$ is unique. The unique equilibrium settlement demand in this case induces $x_A(S_T) = \overline{x}_A$. An equilibrium in which some types of defendant accept S_T is possible only if $x_A(S_T)D + c_A^d + c_T^d = S_T$, the left-hand side of which is strictly increasing in x. Hence, the $x_A(S_T)$ that satisfies this condition is unique. On the other hand, the solution x_A to (14) is strictly increasing in λ, so that the equilibrium $\lambda(S_T)$ is also unique.

Lemma 4 is important because it establishes that, before making a pretrial settlement demand S_T, the plaintiff can anticipate a unique equilibrium payoff in the respective subgame following each choice of S_T. An equilibrium pretrial settlement demand therefore maximizes (16) and can be intuitively characterized by the cutoff type of defendant x_A^* which is indifferent between accepting and rejecting the equilibrium posttrial settlement demand:

Proposition 3 *Assume that the trial court's verdict is completely uninformative. In any pure-strategy perfect Bayesian equilibrium, a case goes all the way to the appeals court if and only if $x < x_A^*$, where $x_A^* \leq \overline{x}_A$. If $0 < x_A^* < \overline{x}_A$, then*

$$D = \frac{f(x_A^*)}{1 - F(x_A^*)}(c_T^p + c_T^d + c_A^p + c_A^d) + \frac{f'(x_A^*)}{1 - F(x_A^*)} \frac{c_A^p + c_A^d}{D}(c_T^p + c_T^d). \qquad (17)$$

Proof. Note that, due to uniqueness of the equilibrium payoffs in the subgame following any settlement demand S_T, the plaintiff's problem of choosing S_T boils down to choosing some $x_A \leq \overline{x}_A$ so as to maximize (16), where λ satisfies the posttrial first-order condition (15) for that x_A. Hence, the upper bound $x_A^* \leq \overline{x}_A$ follows immediately from Lemma 4, and an interior solution can be obtained by taking the derivative of (16),

$$\frac{d\Pi^p}{dx_A} = (1 - F(x_A))D - f(x_A)(c_A^p + c_A^d) - \frac{d\lambda}{dx_A}(c_T^p + c_T^d).$$

Using the total differential of (15), substituting for $\frac{d\lambda}{dx_A} = f(x_A) + f'(x_A)\frac{c_A^p + c_A^d}{D}$ yields (17). ■

The left-hand side and the first summand on the right-hand side of (17) again constitute the well-known tradeoff from the literature on single-stage litigation systems. In addition to that, the second summand on the right-hand side of (17) shows the effect introduced by the possibility to appeal, which depends on the allocation of parties' total litigation costs on the trial and the appeal stage, and on the first derivative of the density of the defendant's private signal. Intuitively, this term reflects that the result of pretrial settlement will influence the plaintiff's cost of increasing x_A in posttrial settlement bargaining, $f(x_A)(c_A^p + c_A^d)$, which is litigation costs in the appeal stage times the marginal probability of the defendant to prefer settlement.

The similarity of the equilibrium conditions of the single-stage and two-stage models, (1) and (17), makes the two models easy to compare. The following Proposition analyzes the effect of appeals on the set of cases that are eventually settled in some stage, and on the defendant's expected equilibrium payoff.

Proposition 4 *Let $D \geq \frac{f(0)}{1-F(0)}(c_T^p + c_T^d + c_A^p + c_A^d)$. If $f'(x_A^*) < 0$ ($f'(x_A^*) > 0$), then the ex-ante probability that the case will be settled in some stage is lower (higher), and the defendant's ex-ante expected payoff is smaller (larger) than in a one-stage litigation system with total legal costs $c^p + c^d = c_T^p + c_T^d + c_A^p + c_A^d$.*

Proof. The proposition compares the present model with a one-stage model just consisting of stages (i), (ii) and (viii) and legal costs $c^p + c^d = c_T^p + c_T^d + c_A^p + c_A^d$. The condition $D \geq \frac{f(0)}{1-F(0)}(c_T^p + c_T^d + c_A^p + c_A^d)$ implies that the unique equilibrium of the latter model is given by (1). Furthermore, recall the definition of $\overline{x}_A$, $D = \frac{f(\overline{x}_A)}{1-F(\overline{x}_A)}(c_A^p + c_A^d)$, which is strictly smaller than $\frac{f(\overline{x}_A)}{1-F(\overline{x}_A)}(c^p + c^d)$. Summing up, the unique equilibrium of the single-stage model satisfies

$$0 < x^* < \overline{x}_A. \tag{18}$$

Suppose first that $f'(x_A^*) < 0$. If $0 < x_A^* < \overline{x}_A$, x_A^* satisfies (17), which implies $D < \frac{f(x_A^*)}{1-F(x_A^*)}(c^p + c^d)$ and, with Proposition 1, $x^* < x_A^*$. If $x_A^* = \overline{x}_A$, then $x^* < x_A^*$ is trivially implied by (18). $x_A^* = 0$ cannot occur, as a necessary condition for this would be that $D < \frac{f(0)}{1-F(0)}(c_T^p + c_T^d + c_A^p + c_A^d)$, a contradiction to a condition of the Proposition.

Suppose now that $f'(x_A^*) > 0$. If $0 < x_A^* < \overline{x}_A$, x_A^* satisfies (17), which implies $D > \frac{f(x_A^*)}{1-F(x_A^*)}(c^p + c^d)$ and, with Proposition 1, $x^* > x_A^*$. If $x_A^* = 0$, then $x^* > x_A^*$ is trivially implied by (18). $x_A^* = \overline{x}_A$ cannot occur, as a necessary condition for this would be that $D > \frac{f(\overline{x}_A)}{1-F(\overline{x}_A)}(c_A^p + c_A^d)$, a contradiction to the definition of $\overline{x}_A$. ■

Proposition 4 shows that the direction in which the appeals system affects equilibrium crucially depends on the distribution from which the defendant's private information is drawn: In a system with appeals, the plaintiff's pretrial settlement demand also seeks to optimize the cost-benefit tradeoff that governs posttrial bargaining. As the marginal cost of increasing x_A^* is positively related to the marginal probability that the defendant prefers settlement, introducing the additional level of jurisdiction has shifted the defendant's equilibrium cut-off type in the direction of lower density.

6 The Information Effect

6.1 Some General Results

In this section, it will be analyzed how the litigants' equilibrium settlement behavior is affected when they can use the trial court's outcome to update their expectations on how the appeals court will eventually judge. Depending on the exact nature of the signal technology, the plaintiff may use the trial court's verdict to update her beliefs on the private information x that the defendant has observed, and both litigants may update their expectations of how the appeals court will eventually judge for given x. Hence, winning (losing) in trial court makes the plaintiff more (less) confident of winning in the appeals court, and therefore it is plausible for her to demand a higher (lower) amount in the posttrial settlement negotiation. Recalling Lemma 2, this latter conclusion is equivalent to $z_1(x_A^1) > z_0(x_A^0)$, which is exactly how case (i) of Proposition 2 is defined. Hence, the intuitively most appealing case of that Proposition is the first one, which is why this Subsection is devoted to characterizing equilibria for general signal technologies under the assumption that equilibrium satisfies this intuitively appealing condition.

Note, however, that contrary to this intuitive argument, the plaintiff's optimal choice of posttrial settlement demand will depend on the *marginal* effect of the trial court's verdict on litigants' expectations of how the appeals court may judge. Hence, it is not possible to rule out the third case of Proposition 2 for general signal technologies, so that it is necessary to rule it out whenever one discusses a specific signal technology. I will analyze such a specific signal technology in the next Subsection. It will turn out that any perfect Bayesian equilibrium must be according to case (i) of Proposition 2. Furthermore, this specific signal technology will serve as an example for the possibility of some counter-intuitive comparative statics regarding the impact of legal costs and the trial court's accuracy on litigants' incentives to settle.

Under case (i) of Proposition 2 and using (10), the plaintiff's posttrial objective function (8) can be simplified to

$$x_A^r \in \arg\max_{x_A \in [0, x_T]} \int_0^{x_A} (z_r(x)D - c_A^p) y(x) f(x) dx + (z_r(x_A)D + c_A^d) \int_{x_A}^{x_T} y(x) f(x) dx, \quad (19)$$

where x_T is the threshold type of defendant defined in case (i) of Proposition 2 above which the defendant rejects the pretrial settlement demand. The first-order condition is

$$z_r'(x_A^r)D = \frac{y_r(x_A^r)f(x_A^r)}{\int_{x_A^r}^{x_T} y(x)f(x)dx}(c_A^p + c_A^d). \tag{20}$$

The following Lemma presents some results on the plaintiff's optimal choice of posttrial settlement demand.

Lemma 5 *Consider a pure-strategy perfect Bayesian equilibrium of the subgame after the plaintiff's pretrial settlement demand that is governed by case (i) of Proposition 2.*

(i) If $y_1(0) = 0$, then $x_A^1 > 0$. Furthermore, $x_A^0 = 0$ for sufficiently flat $z_0(\cdot)$ and $y_1(0) < 1$.

(ii) If the function $y(x)f(x)$ exhibits an increasing hazard rate and $z_r(\cdot)$ is weakly concave on $(0,1)$, there is at most one x_A that satisfies the first-order condition (20). In this case, the equilibrium choice of x_A^r is the unique solution of (20) and strictly increasing in x_T if it exists, and $x_A^r = 0$ otherwise.

Proof. Note first being in case (i) of Proposition 2 immediately rules out $x_T = 0$, as in this case the trivial posttrial equilibrium would be $x_A^0 = x_A^1 = 0$. Hence, $x_T > 0$.

Part (i): The first derivative of the objective function in (19) w.r.t. x_A is

$$z_r'(x_A)\int_{x_A}^{x_T} y(x)f(x)dxD - y_r(x_A)f(x_A)(c_A^p + c_A^d),$$

which is strictly positive for $x_A = 0$ if $y_r(0) = 0$. Hence, $x_A^1 > 0$ whenever $y_1(0) = 0$. Furthermore, for $y_0(0) > 0$ (which is equivalent to $y_1(0) < 1$) and sufficiently small $z_r'(\cdot)$, the derivative of the objective function will be negative for all $x_A \in [0,1]$, which implies $x_A^0 = 0$.

Part (ii): $y(x)f(x)$ exhibiting an increasing hazard rate implies that the right-hand side of (20) is strictly increasing in x_A, and $z_r(\cdot)$ being weakly concave implies that the left-hand side of (20) is strictly decreasing in x_A. Hence, there can be at most one x_A that satisfies (20). If a solution of (20) does not exist, $x_A^r = x_T$ is ruled out by Lemma 1, which leaves $x_A^r = 0$ as the only alternative.

Suppose now that $x_A^r > 0$ satisfies (20). Taking the total differential of (20) yields

$$\frac{dx_A^r}{dx_T} = \frac{y_r(x_T)f(x_T)z_r'(x_A^r)D}{\left[y_r(x_A^r)f(x_A^r)z_r'(x_A^r) - z_r''(x_A^r)\int_{x_A^r}^{x_T} y(x)f(x)dx\right]D + \left[y_r'(x_A^r)f(x_A^r) + y_r(x_A^r)f'(x_A^r)\right](c_A^p + c_A^d)}$$

which is strictly positive whenever $z_r''(x) \leq 0$. ■

If the defendant's private signal is sufficiently informative also for predicting the trial court's decision, a case will go to the appeals court with strictly positive probability after the plaintiff has won in trial court. The plaintiff infers having a weak case from the fact that the defendant has rejected the pretrial demand, but infers having a strong case from the trial court's decision. Hence, it is sufficiently unlikely that the defendant has observed a very low x as to make it worthwhile for the plaintiff to make a settlement demand that these low-x types of defendant will reject.

On the other hand, this is not necessarily true if the defendant has won in trial court: In this case, both the fact that the defendant is obviously confident and the trial court outcome make the plaintiff believe that the defendant is sufficiently likely to have observed a very low x that she prefers to settle even with these low types of defendant.

Part (ii) of Lemma 5 translates the uniqueness of the solution of the standard, single-instance model's first-order condition (1) to posttrial bargaining. If the distribution of the plaintiff's beliefs satisfies the increasing-hazard-rate condition, which is usually imposed on the ex-ante distribution of types in the single-stage model, then the right-hand side, divided by the integral, is increasing in x_A^r. Hence, if the left-hand side is non-increasing (which is always the case in (1), as it is independent of x^*), the result follows immediately.

However, note that this latter condition may not hold in posttrial bargaining, especially for the case that $l_T = 0$ in which the defendant will typically be more optimistic and revise his expectations of losing in the appeals court to below x: For instance, if signals are independent (see Example 1 in Section 3), $z_0(x)$ is convex, which means that the defendant's revision of expectations is larger for intermediate levels of observed x. In this sense, part (ii) of Lemma 5 points out another potential complication that may arise in the analysis of posttrial bargaining after the defendant has won in trial court, as opposed to settlement negotiation in the single-instance model.

The second result presented in Part (ii) of Lemma 5 is that, under the same set of assumptions that guarantees an interior posttrial solution, posttrial equilibrium choices are monotonic in pretrial choices. That is to say, if the ex-ante probability that the defendant rejects the pretrial settlement demand is high (which means that the pretrial demand was high), then the ex-ante probability that the case will go all the way to the appeals court is also high, which also means that the posttrial settlement demand is high.

Assuming that, for every pretrial settlement demand S_T, there is a pure-strategy perfect Bayesian equilibrium characterized by case (i) of Proposition 2, the plaintiff will anticipate equilibrium S_A^r, x_A^r and $x_T > \max\{x_A^0, x_A^1\}$ when choosing S_T. Using the results from Section 4, the plaintiff's objective is equivalent to choosing x_T so as to maximize

$$
\begin{aligned}
\Pi^p(x_T) = \sum_r & \left[\int_0^{x_A^r} (z_r(x)D - c_A^p) y_r(x) f(x) dx + (z_r(x_A^r)D + c_A^d - c_T^p) \int_{x_A^r}^{x_T} y_r(x) f(x) dx \right] \\
& + \left[\sum_r (y_r(x_T) z_r(x_A^r)) D + c_A^d + c_T^d \right] \int_{x_T}^1 f(x) dx.
\end{aligned}
\tag{21}
$$

The following Proposition presents the first-order condition of an interior optimum and specifies conditions under which it will be satisfied in perfect Bayesian equilibrium.

Proposition 5 *Assume that there is a pure-strategy perfect Bayesian equilibrium of the subgame after making the pretrial settlement demand that is governed by case (i) of Proposition 2, and let the conditions of Part (ii) of Lemma 5 be satisfied and $y_1(0)$ be sufficiently small as to guarantee $x_A^1 > 0$ for every x_T.*

If D is sufficiently large and $c_A^p + c_A^d$ and $f(0)$ are sufficiently small, then there is a perfect Bayesian equilibrium such that the case is settled pretrial if and only if $x \geq x_T^*$, where x_T^* satisfies the first-order condition

$$
\left\{ y_1'(x_T^*)(z_1(x_A^1) - z_0(x_A^0)) + \sum_r \left(z_r'(x_A^r) \frac{dx_A^r}{dx_T} y_r(x_T^*) \right) \right\} D = \frac{f(x_T^*)}{1 - F(x_T^*)} (c_T^d + c_T^p). \tag{22}
$$

Proof. The first-order condition (22) is obtained by taking the first derivative of the objective function (21) w.r.t. x_T. Let us denote this objective function $\Pi_T(x_T)$. A sufficient set of conditions for x_T to satisfy this first-order condition in a perfect Bayesian equilibrium is that (i) $\Pi_T(x_T)$ is indeed the relevant objective function, (ii) the first derivative of this objective function satisfies $\lim_{x_T \searrow 0} \Pi_T'(0) > 0$ and $\Pi_T'(1) < 0$, and (iii) the first derivative of the objective function $\Pi_T'(\cdot)$ is continuous except for a finite set (x_d^i), where for every i, $\lim_{x_T \nearrow x_d^i} \Pi_T'(x_T) < \lim_{x_T \searrow x_d^i} \Pi_T'(x_T)$.

The first condition is guaranteed by assumption. As for the third condition, continuity of $\Pi_T'(\cdot)$ depends on the impact of x_T on the posttrial choices $x_A^r(x_T)$: Low x_T may imply a boundary solution $x_A^r(x_T) = 0$ for some $r \in \{0, 1\}$. Due to Part (ii) of Lemma 5, x_A^r are increasing in x_T whenever they are positive. Hence, there may be up to two discontinuities at $x_d^r := \max\{x_T : x_A^r(x_T) = 0\}$. Furthermore,

$$
\lim_{x_T \searrow x_d^i} \Pi_T'(x_T) - \lim_{x_T \nearrow x_d^i} \Pi_T'(x_T) = z_r'(x_A^r) \frac{dx_A^r}{dx_T} y_r(x_d^r) D(1 - F(x_d^r)) > 0
$$

again due to Part (ii) of Lemma 5.

As for the second condition, note first that $1 - F(1) = 0$ as $F(\cdot)$ is a probability density function, and recall that $z_r(\cdot)$ and $y_r(\cdot)$ are differentiable by assumption. Hence,

$\Pi'_T(1) = -f(1)(c_T^d + c_T^p) < 0$. Furthermore, $\lim_{x_T \searrow 0}(z_1(x_A^1) - z_0(x_A^0)) = 0$, so that

$$\lim_{x_T \searrow 0} \Pi'_T(0) = \sum_r \left(z'_r(0) \frac{dx_A^r(0)}{dx_T} y_r(0) \right) D(1 - F(0)) - f(0)(c_T^d + c_T^p),$$

which is increasing in D and decreasing in $(c_T^d + c_T^p)$ and $f(0)$, as long as x_A^r is not a boundary solution for at least $r = 1$, which is guaranteed by the assumption of sufficiently large $y_1(0)$. ■

Intuitively, in order to establish the interior solution as an equilibrium, it is necessary to rule out the boundary solutions $x_T = 0$ and $x_T = 1$ as well as potentially those points at which the objective function may not differentiable. It turns out that the assumptions made in Part (ii) of Lemma 5 already rule out $x_T = 1$ and the non-differentiable points if they exist. Hence, if, in addition to these assumptions, the potential gain D for the plaintiff is sufficiently large compared to the legal costs and the defendant is sufficiently unlikely to privately know to have a very strong case as to rule out that the plaintiff wants to settle pretrial with certainty ($x_T = 0$), the equilibrium x_T^* will satisfy the first-order condition (22).

Let us now examine this first-order condition (22) in more detail: The right-hand side is the well-known marginal cost of making a tougher settlement demand that already appears in the first-order condition (1) of the standard, single-instance model: The defendant will accept such a higher settlement demand with lower ex-ante probability, so that the litigation costs of the trial court $c_T^d + c_T^p$ will be incurred with higher probability, represented by the hazard rate $\frac{f(x_T)}{1-F(x_T)}$. The difference introduced by the possibility of an appeal is related to the expected marginal benefit of making a higher settlement demand, which is on the left-hand side of the first-order condition. While this marginal benefit is just D in the standard model (see again equation (1)), there are two effects in the two-stage model: First, there is a strategic effect similar to that introduced in Section 5, captured by the second summand in the expression in curly brackets on the left-hand side of (22), $\sum_r \left(z'_r(x_A^r) \frac{dx_A^r}{dx_T} y_r(x_T) \right) D$. This effect is just the average impact of settling with lower probability pretrial on the probability of settling posttrial in equilibrium, and therefore also on the plaintiff's equilibrium posttrial payoff. Intuitively, the plaintiff chooses the pretrial settlement demand so as to optimize her strategic position in the eventual posttrial settlement negotiation.

The second effect, captured by the first summand in the expression in curly brackets on the left-hand side of (22), $y'_1(x_T)(z_1(x_A^1) - z_0(x_A^0))D$, is related to the litigants anticipating the way in which the information revealed by the trial court's verdict will eventually translate into posttrial payoffs. This is why I will refer to this effect as the *information effect*. The information effect is made up of the difference in equilibrium posttrial settlement payment $S_A^1 - S_A^0 = (z_1(x_A^1) - z_0(x_A^0))D$ after the plaintiff versus the defendant having won in trial court, and the impact $y_1'(x_T)$ of settling with less types of defendant pretrial on the probability of winning the marginal case in the trial court. Intuitively, making a tougher pretrial settlement demand is more attractive for the plaintiff if (i) this increases the probability $y_1(x_T)$ of winning the marginal case in trial court to a larger extent, and (ii) good news in trial court translate more heavily into posttrial equilibrium payoffs.

Note that, due to the increasing-hazard-rate assumption for $F(\cdot)$, it is sufficient to compare the left-hand side of (22) under different sets of assumptions to compare x_T and thus the probabilities of a pretrial settlement in these cases: If the equilibrium marginal type $\tilde{x}_T$ of defendant that settles pretrial under a certain set of assumptions causes the left-hand side of (22) to be larger than the right-hand side under some other set of assumptions, then the equilibrium marginal type x_T under this latter set of assumptions will be larger than $\tilde{x}_T$, which means that a case will go to trial court with a larger probability in the latter case.

6.2 Example: Identical Conditional Probabilities

In the preceding Subsection I have analyzed equilibrium assuming it to satisfy certain conditions. Whether these assumptions indeed hold, and in which direction the strategic and the information effects move equilibrium choices, will need to be assessed for specific given signal technologies. In this Subsection I will, therefore, illustrate some typical features of the information effect for a special class of signal technologies in which, conditional on the defendant's private signal x, the interim probabilities for the plaintiff to win in the trial court and for her to win in the appeals court are both equal to x, which means that $y_1(x) = x = 1 - y_0(x)$. In order to capture the effect of the trial court's accuracy, I assume that with probability ρ the trial court perfectly anticipates the appeals court's eventual verdict ($l_T = l_A$), and with probability $1 - \rho$ it just randomizes between each outcome using the probabilities x and $1 - x$, and that the 'type' of trial court is unobservable to litigants. Hence, ρ is a proxy for the trial court's accuracy: For high values of ρ, the appeals court is very unlikely to overturn the trial court's decision. If, on the other hand, ρ is low, observing the trial court's verdict is still useful for the plaintiff to update her beliefs on the defendant's private information, whereas litigants cannot learn much new on the appeals court's eventual verdict for given x.

Following the discussion in Section 3, the signal technologies analyzed in this section are characterized by $p_r(x)$, which are equal to 1 with probability ρ, and equal to those given in Example 3 with probability $1 - \rho$, i.e. $p_1(x) = \rho + (1 - \rho)x$ and $p_0(x) =$

$\rho + (1-\rho)(1-x)$. It follows that

$$\begin{aligned} z_1(x) &= \rho + (1-\rho)x \\ z_0(x) &= (1-\rho)x. \end{aligned}$$

Furthermore, I will simplify the analysis by assuming that the defendant's private information is ex-ante uniformly distributed. This assumption also allows me to better focus on the information effect, as it rules out some of the purely strategic effect analyzed in section 5.[11] The following lemma characterizes equilibrium of the subgame following the plaintiff's pretrial settlement demand S_T.

Lemma 6 *Assume that the trial court perfectly anticipates the appeals court's eventual decision with probability ρ, and judges according to Example 3 with probability $1-\rho$, and that the defendant's private information is ex-ante uniformly distributed. Then:*

(i) There is no pure-strategy perfect Bayesian equilibrium that is characterized by cases (ii) or (iii) of Proposition 2.

(ii) The unique pure-strategy perfect Bayesian equilibrium of the subgame following any pretrial settlement demand $S_T > c_T^d + c_A^d$ is characterized by x_A^0, x_A^1 and x_T which satisfy:

$$\begin{aligned} S_T &= \left[x_T(\rho + (1-\rho)x_A^1) + (1-x_T)(1-\rho)x_A^0\right] D + c_A^d + c_T^d & (23) \\ x_A^1 &= -\frac{Z_A}{1-\rho} + \sqrt{\frac{Z_A^2}{(1-\rho)^2} + x_T^2} & (24) \\ x_A^0 &= \max\left\{1 - \frac{Z_A}{1-\rho} - \sqrt{\frac{Z_A^2}{(1-\rho)^2} + (1-x_T)^2}, 0\right\}, & (25) \end{aligned}$$

where $Z_A := \frac{c_A^d + c_A^p}{D}$ is the ratio of total litigation costs in the appeals court and the amount in dispute.

Proof. For all proofs of this Subsection see the Appendix. ■

Taking the partial derivative of (24) and (25) with respect to ρ confirms the straightforward intuition that, for a given set of cases that go to trial court in equilibrium, a case will be settled posttrial with a higher probability if the appeals court is more costly and trial court more accurate: If it is known that the appeals court will probably judge in

[11] Recall from Proposition 4 that with uniformly distributed x, litigants' anticipation of appeals has no impact in the absence of the information effect.

the same way as the trial court, there is no need to incur the additional legal costs of the appeals court.

Furthermore, due to $y(0) = 0$ in this example, Part (i) of Lemma 5 applies which states that posttrial settlement will, in equilibrium, be governed by an interior solution after the plaintiff has won in trial court, i.e. the case will be settled posstrial and go to the appeals court with strictly positive probabilities. However, if the defendant has won in trial court, there may be a boundary solution posttrial: (25) implies that such a case will always be settled posttrial if $(1-x_T)^2 \geq 1 - \frac{2Z_A}{1-\rho}$, which is satisfied if the probability of the case not being settled pretrial was already low, the litigation costs in the appeals stage are sufficiently high relative to the potential damages, and the trial court is sufficiently accurate.

Let us now turn to the plaintiff's choice of pretrial settlement demand. Part (i) of Lemma 6 confirms that, for any choice of S_T, the subsequent equilibrium will be of the type described in case (i) of Proposition 2, i.e. the defendant's decision of whether to accept such a settlement demand depends in a strictly monotonic way on his private information: If this private information indicates that he is going to win in the appeals court with a high probability, then his expected payoff from the trial court and posttrial settlement negotiation will also be high. Hence, he will reject S_T if and only if x is below a certain threshold x_T.

The first-order condition for the plaintiff's optimal choice of pretrial settlement demand can therefore be obtained by using Lemma 6 to substitute for the subsequent equilibrium choices in (22). The following Proposition specifies under which conditions this first-order condition is also sufficient:

Proposition 6 *If the signal technology is as described in Lemma 6 and* $\rho \geq \frac{c_T^p + c_T^d}{D}$*, there is some* x_T^* *such that in any pure-strategy perfect Bayesian equilibrium the case is settled pretrial if and only if* $x \geq x_T^*$*, where* x_T^* *satisfies the first-order condition*

$$\left\{\rho + (1-\rho)\left(-1 + \sqrt{4{x_T^*}^2 + \frac{\frac{Z_A^4}{(1-\rho)^4}}{\frac{Z_A^2}{(1-\rho)^2} + {x_T^*}^2}} + \sqrt{4(1-x_T^*)^2 + \frac{\frac{Z_A^4}{(1-\rho)^4}}{\frac{Z_A^2}{(1-\rho)^2} + (1-x_T^*)^2}}\right)\right\} D = \frac{c_T^p + c_T^d}{1 - x_T^*} \tag{26}$$

if $(1-x_T^*)^2 < 1 - \frac{2Z_A}{1-\rho}$*, and*

$$\left\{\rho + (1-\rho)\left(-\frac{Z_A}{1-\rho} + \sqrt{4{x_T^*}^2 + \frac{\frac{Z_A^4}{(1-\rho)^4}}{\frac{Z_A^2}{(1-\rho)^2} + {x_T^*}^2}}\right)\right\} D = \frac{c_T^p + c_T^d}{1 - x_T^*} \tag{27}$$

if $(1-x_T^*)^2 > 1 - \frac{2Z_A}{1-\rho}$.

Specifically, if $\left(\frac{c_T^p + c_T^d}{\rho D}\right)^2 < 1 - \frac{2Z_A}{1-\rho}$*, then* $x_T^* > 1 - \frac{c_T^p + c_T^d}{D}$*.*

In line with the findings for general signal technologies presented in Proposition 5, Proposition 6 shows that there will be an interior equilibrium settlement demand (in the sense that it will be accepted and rejected with positive probabilities) if $\rho \geq \frac{c_T^p + c_T^d}{D}$, that is to say, if the trial court is sufficiently accurate (high ρ), the stakes D are sufficiently large and trial costs $c_T^d + c_T^p$ sufficiently low.

Taking a closer look at the first-order conditions, it turns out that the left-hand side of (26) is always larger than D and the left-hand side of (27) always larger than ρD. The right-hand sides are weakly smaller than D whenever $x_T \leq 1 - \frac{c_T^p + c_T^d}{D}$ and weakly smaller than ρD whenever $x_T \leq 1 - \frac{c_T^p + c_T^d}{\rho D}$. Hence, an equilibrium x_T^* that satisfies the first-order condition (26) [(27)] must be above $1 - \frac{c_T^p + c_T^d}{D}$ $[1 - \frac{c_T^p + c_T^d}{\rho D}]$.

Furthermore, if the lower bound $1 - \frac{c_T^p + c_T^d}{\rho D}$ for the interior solution given by the first-order condition (27) (which is the relevant condition for the case of the posttrial boundary solution $x_A^0 = 0$) is above the domain that is relevant for this condition, this immediately excludes the possibility of the posttrial boundary solution $x_A^0 = 0$ to occur in any perfect Bayesian equilibrium of the entire game. This is exactly the case that is highlighted by the last claim of Proposition 6: If $\left(\frac{c_T^p + c_T^d}{\rho D}\right)^2 < 1 - \frac{2Z_A}{1-\rho}$, then the objective function is increasing throughout the domain $(1 - x_T)^2 > 1 - \frac{2Z_A}{1-\rho}$ under which the posttrial boundary solution $x_A^0 = 0$ occurs, which implies that equilibrium will always satisfy (26), for which case we have just established that $x_T > 1 - \frac{c_T^p + c_T^d}{D}$.

Proposition 6 allows us to analyze the impact of litigation costs on litigants' incentives to settle. While the following proposition shows the expected positive effect of litigation costs in the trial stage on settlement incentives, it establishes the somewhat surprising result that, if equilibrium satisfies the first-order condition 26, higher litigation costs in the appeal stage actually increase the probability that a case goes to the trial court:[12]

Proposition 7 *Let the signal technology be as described in Lemma 6, assume* $\left(\frac{c_T^p + c_T^d}{\rho D}\right)^2 <$ $1 - \frac{2Z_A}{1-\rho}$*, and consider a marginal change in the total legal cost of the trial stage* $c_T^d + c_T^p$ *(of the appeals stage* $c_A^d + c_A^p$*). Then there is at least one perfect Bayesian equilibrium before and after the change such that the equilibrium probability that the case goes to the trial court* x_T^* *has decreased (increased) due to the change.*

Higher litigation costs in the trial stage increase the right-hand side of the first-order condition (26), which implies that x_T^* must be reduced. Hence, just like in the well-known

[12]Note that, while it is readily established that, under the conditions formulated in Proposition 7, x_T^* must satisfy the first-order condition (26) in any perfect Bayesian equilibrium, multiplicity of equilibrium cannot be ruled out. Hence, the comparative statics result presented in Proposition 7 holds only for continuous changes in the equilibrium x_T^*. However, such a continuously changing equilibrium always exists for every marginal parameter change.

single-stage model summarized in Proposition 1, higher litigation costs encourage out-of-court settlement. In order to understand why increasing the litigation costs of appeal has the opposite effect, it will be useful to look at both effects identified in the discussion of the first-order condition (22) separately: The marginal types of defendant who settle posttrial, x_A^1 and x_A^0 given by (24) and (25) are increasing in x_T, but this effect is smaller for higher litigation costs in the appeals stage. Hence, the impact of settling pretrial with lower probability on the plaintiff's posttrial payoff, which we have labelled the 'strategic effect', gets smaller as $c_A^d + c_A^p$ increases which would imply a decreasing equilibrium x_T^*.

However, the 'information effect', which captures how litigants anticipate the later use of the information revealed by the trial court, works in the opposite direction and turns out to dominate the strategic effect: Although, in the case of interior equilibria posttrial implied by the condition $(1 - x_T^*)^2 < 1 - \frac{2Z_A}{1-\rho}$, equilibrium posttrial settlement payments S_A^1 and S_A^0 for both possible trial outcomes are falling in the litigation costs, those after the defendant has won in trial court are more heavily affected by the litigation costs, so that the difference $S_A^1 - S_A^0$ is increasing in $c_A^d + c_A^p$. This makes it more attractive for the plaintiff to win in trial court and thus reduces her incentives to settle pretrial.

The intuitive reason for why $S_A^1 - S_A^0$ is increasing in $c_A^d + c_A^p$ is that losing in trial court makes the plaintiff also pessimistic on the appeals court's eventual judgment. Hence, losing in trial court reduces the impact $y_r'(\cdot)$ of the posttrial settlement demand S_A^r on the plaintiff's marginal cost of settling with lower probability posttrial (see the right-hand side of (20)). Consequently, for a given increase in legal costs, the plaintiff will reduce equilibrium posttrial settlement demand after losing in trial court more than after winning in trial court. In other words, the plaintiff's choice of posttrial settlement demand is less sensitive to parameter changes if she has lost in trial court. Hence, if a parameter change reduces equilibrium posttrial settlement demands after any outcome in the trial court, as the increase in the legal cost of the appeals stage does, it will increase the difference $S_A^1 - S_A^0$ in the plaintiff's posttrial payoffs after winning and losing in trial court. As this mechanism is just based on the identity $y_0(x) \equiv 1 - y_1(x)$ and therefore $y_0'(x) = -y_1'(x)$, it seems safe to argue that the information effect being countervailing to the strategic effect is a typical feature of interior equilibria of the game.

Another interesting question is whether a trial court that is more accurate in predicting the appeals court's eventual decision will attract more cases in the first place. Let us start by considering two extreme cases, a perfectly accurate and a purely randomizing trial court. As the perfectly accurate trial court is characterized by $\rho = 1$, posttrial settlement bargaining after the defendant has prevailed in trial court is given by the boundary solution $x_A^0 = 0$ for every x_T, which, together with $\rho = 1 > \frac{c_T^d + c_T^p}{D}$, implies that the equilibrium x_T^* is given by (27). Substituting for $\rho = 1$ in (27) yields the first-order condition $x_T^* = 1 - \frac{c_T^d + c_T^p}{D}$.

As for the other extreme case in which the trial court only randomizes, note first that this case is *not* equivalent to the case $\rho = 0$ in the example discussed in this Subsection - even this lowest possible accuracy reveals some information on the appeals court's eventual judgment to the plaintiff, as the trial court is known to decide for the plaintiff with the probability x privately observed by the defendant. Instead, we must go back to the analysis in Section 5, where (15) and (17) together imply that if the defendant's private information is ex ante uniformly distributed, as it is assumed throughout this Subsection, the probability that the case goes to trial court is $\lambda = 1 - \frac{c_T^d + c_T^p}{D}$.

The surprising result of this exercise is that a perfectly accurate trial court will be used with exactly the same ex-ante probability as a purely randomizing court. An immediate conclusion is that, unless equilibrium choices are completely independent of the trial court's accuracy, it is always possible to find two signal technologies such that the less accurate trial court will be used with higher ex-ante probability in equilibrium. Indeed, it is shown in Proposition 6 that, as long as stakes are sufficiently high relative to the legal costs of the trial court and ρ is not too high as to ensure an interior solution given by the first-order condition (26), a case will go to trial court with a higher probability than in either of the extreme cases just discussed. The following proposition summarizes this result and is presented without proof:

Proposition 8 *Consider a signal technology as described in Lemma 6. If* $\left(\frac{c_T^p + c_T^d}{\rho D}\right)^2 < 1 - \frac{2Z_A}{1-\rho}$*, such an intermediately accurate trial court will be used with higher ex-ante probability than a perfectly accurate trial court* ($\rho = 1$)*, which in turn will be used with identical ex-ante probability as the completely uninformative trial court discussed in Section 5.*

Intuitively, a very inaccurate trial court's decision won't influence posttrial equilibrium payoffs much. Due to the aforementioned effect that smaller differences in the plaintiff's posttrial equilibrium payoffs across trial court outcomes increase her incentive to settle pretrial, a very inaccurate trial court will be used with rather low probability. Hence, making the trial court slightly more accurate than that will increase the probability that it is used. However, as the trial court's accuracy increases further, the defendant's informational advantage in posttrial settlement negotiation vanishes. Anticipating his lower information rent in posttrial bargaining, the defendant will be easier to convince of settling pretrial, which brings the probability that the court is used back down again.

7 Conclusions

This paper has identified two effects through which the possibility of an appeal will influence asymmetrically informed litigants' incentives to settle: The strategic effect makes litigants consider in pretrial negotiations the strategic environment in which posttrial settlement negotiations will eventually take place and follows a similar intuition as the literature on sequential bargaining with asymmetric information. The second effect is the information effect which makes litigants anticipate how the information revealed by the trial court's verdict will influence equilibrium posttrial payoffs.

The main lesson from this paper is that taking into account the trial court's verdict as a public signal on the appeals court's eventual decision may yield the following surprising results: It turns out that the information effect implies higher incentives to settle pretrial if posttrial equilibrium payoffs are insensitive to the trial court's verdict, which is typically the case if legal costs are high. Furthermore, a very accurate trial court reduces the defendant's posttrial information rent and makes him more willing to accept a given pretrial settlement demand.

A policy discussion to which these results may make an important contribution is that on the optimal design of the legal process. In general, a social planner may decide on whether to invest in more or less levels of jurisdictions, and whether to invest more or less in the accuracy of the existing courts. For instance, Shavell (1995) compares costs and benefits of adding a level of jurisdiction, arguing that due to convex costs of avoiding judicial errors at each level two imperfectly accurate levels of court are socially preferable to a single, more accurate court. However, Shavell assumes that litigants are perfectly informed but cannot settle. The present paper relaxes these assumptions and suggests that litigants' settlement behavior may affect social welfare in different ways as a result of such a legal reform.

Of course, my results depend on various simplifying assumptions which may be relaxed in future research. First, a potential way of relating my analysis to the main line of the economic literature of appeals cited in the introduction would be to extend the model to allow for courts to act strategically or to update their information using a potentially better informed litigant's actions. Another group of strategic players that has not been addressed in my model is solicitors, whose incentives may not be aligned with their clients' depending on the compensation scheme in use. As for the timing of the game, I have made the simplifying assumption that the case always goes to court if settlement negotiations break down, which rules out those credibility issues addressed by Nalebuff (1987). Allowing for these credibility issues may create interesting countervailing effects and, therefore, be a worthwhile task for future research. Finally, when using the model presented in this paper for a welfare analysis of legal reform, it should be kept in mind that changes in the equilibrium of the litigation game may be interrelated with the incentives for the underlying actions before litigation takes place like, for instance, an injurer's decision to take precautions.[13]

A feature of my model that practitioners may feel uncomfortable with is that I do not address the appeals court's accuracy in finding the 'truly' correct decision at all. The reason why I ignored this question is that the paper's focus is on deriving generally valid results on litigants' settlement incentives. All that rational, monetary payoff maximizing litigants care about is the highest court's eventual decision. The results that I derive in this model are therefore valid whether or not the appeals court's decision is correct. Having said this, the issue of the appeals court's accuracy may be important when analyzing welfare effects, and if litigants suffer non-monetary preference costs when the legal system errs to their disadvantage.

Appendix: Proofs for Subsection 6.2

A Proof of Lemma 6

Suppose there is an equilibrium that is characterized by case (iii) of Proposition 2. Then, in the interval $x_A^1 < x < x_A^0$, $S(x) = [\rho + (1-\rho)(1 - x + x_A^1)]x$ is strictly convex. Hence, in any perfect Bayesian equilibrium characterized by case (iii) of Proposition 2 there must be unique x_T and x_T^0 such that $S(x_T)D + c_A^d + c_T^d = S(x_T^0)D + c_A^d + c_T^d = S_T$, $x_A^1 < x_T < x_T^0 < x_A^0$ and $M = [0, x_T] \cup [x_T^0, 1]$. Furthermore, by symmetry of the quadratic function $S(\cdot)$,

$$x_T^0 = \frac{1}{1-\rho} + x_A^1 - x_T. \tag{28}$$

Let $\lambda_r := \text{Prob}(x \in M \wedge l_T = r) = \int_M y_r(x)dx$. When choosing a posttrial settlement demand after a trial court's verdict $l_T = r$, the plaintiff's objective function (8) becomes, in our example,

$$\Pi_A^r(x_A) = \begin{cases} \int_0^{x_A} (z_r(x)D - c_A^p) \frac{y_r(x)}{\lambda_r} dx + (z_r(x_A)D + c_A^d) \int_{[x_A, x_T] \cup [x_T^0, 1]} \frac{y_r(x)}{\lambda_r} dx, & \text{if } x_A \le x_T; \\ \int_{[0, x_T] \cup [x_T^0, x_A]} (z_r(x)D - c_A^p) \frac{y_r(x)}{\lambda_r} dx + (z_1(x_A)D + c_A^d) \int_{x_A}^1 \frac{y_r(x)}{\lambda_r} dx, & \text{if } x_A \ge x_T^0, \end{cases}$$

the first derivative of which is

$$\Pi_A^{r\prime}(x_A) = \begin{cases} -\frac{y_r(x_A)}{\lambda_r}(c_A^d + c_A^p) + z_r'(x_A)D \int_{[x_A, x_T] \cup [x_T^0, 1]} \frac{y_r(x)}{\lambda_r} dx, & \text{if } x_A \le x_T; \\ -\frac{y_r(x_A)}{\lambda_r}(c_A^d + c_A^p) + z_r'(x_A)D \int_{x_A}^1 \frac{y_r(x)}{\lambda_r} dx, & \text{if } x_A \ge x_T^0, \end{cases} \tag{29}$$

[13]See, for instance, P'ng (1987), Spier (1994), Shavell (1999), Hylton (2002) and Landeo, Nikitin, and Baker (2007).

As both parts of $\Pi_A^{r\prime}(\cdot)$ are strictly decreasing in x_A, local maxima for either part are given by the first-order conditions.

Due to our initial supposition of being in case (iii) of Proposition 2, we have $x_A^1 \leq x_T$, which is given by the first-order condition based on the first case of (29):

$$-x_A^1(c_A^d + c_A^p) + (1-\rho)D\frac{1}{2}\left(x_T^2 - {x_A^1}^2 + 1 - {x_T^0}^2\right) = 0,$$

which, with $Z_A = \frac{c_A^d + c_A^p}{D}$, implies that

$$x_A^1 = -\frac{Z_A}{1-\rho} + \sqrt{\frac{Z_A^2}{(1-\rho)^2} + 1 - {x_T^0}^2 + x_T^2}. \tag{30}$$

Similarly, Proposition 2 requires that $x_T^0 \leq x_A^0$, so that x_A^0 is given by the first-order condition based on the second case of (29):

$$-(1-x_A^0)(c_A^d + c_A^p) + (1-\rho)D\frac{1}{2}(1-x_A^0)^2 = 0,$$

which implies that

$$x_A^0 = 1 - \frac{2Z_A}{1-\rho}. \tag{31}$$

However, this interior solution is not an element of the relevant domain $(x_T^0, 1]$:

$$\begin{aligned}
x_T^0 &= \frac{1}{1-\rho} + x_A^1 - x_T = 1 - \frac{2Z_A}{1-\rho} + \frac{\rho + Z_A}{1-\rho} - x_T + \sqrt{\frac{Z_A^2}{(1-\rho)^2} + 1 - {x_T^0}^2 + x_T^2} \\
&> 1 - \frac{2Z_A}{1-\rho} + \frac{Z_A}{1-\rho} - x_T + \sqrt{\frac{Z_A^2}{(1-\rho)^2} + x_T^2} \\
&> 1 - \frac{2Z_A}{1-\rho} = x_A^0.
\end{aligned}$$

Hence, for all $x_A > x_T^0$, $\Pi_A^0(x_A) < \Pi_A^0(x_T^0)$. As $\Pi_A^0(x_T^0) < \Pi_A^0(x_T)$, it follows that the plaintiff's optimal choice after $l_T = 0$ is some $x_A^0 \leq x_T$, a contradiction to case (iii) of Proposition 2.

Consider therefore an equilibrium characterized by cases (i) or (ii) of Proposition 2. Define $\lambda_r := \text{Prob}(x \in M \wedge l_T = r) = \int_M y_r(x)dx$. As $[0, \max\{x_A^0, x_A^1\}] \subseteq M$, we can rewrite the plaintiff's posttrial objective function (8) as

$$\Pi_A^r(x_A) = \int_0^{x_A} (z_r(x)D - c_A^p)\frac{y_r(x)}{\lambda_r}dx + (z_r(x_A)D + c_A^d)\left(1 - \int_0^{x_A}\frac{y_r(x)}{\lambda_r}dx\right).$$

Taking the partial derivative w.r.t. x_A yields the first-order conditions

$$\begin{aligned}
-x_A^1(c_A^d + c_A^p) + \frac{1}{2}(1-\rho)D(2\lambda_1 - {x_A^1}^2) &= 0 \\
-(1-x_A^0)(c_A^d + c_A^p) + \frac{1}{2}(1-\rho)D\left((1-x_A^0)^2 - (1-2\lambda_0)\right) &= 0,
\end{aligned}$$

which imply

$$x_A^1 = -\frac{Z_A}{1-\rho} + \sqrt{\frac{Z_A^2}{(1-\rho)^2} + 2\lambda_1} \tag{32}$$

$$x_A^0 = \max\left\{1 - \frac{Z_A}{1-\rho} - \sqrt{\frac{Z_A^2}{(1-\rho)^2} + (1-2\lambda_0)}, 0\right\}, \tag{33}$$

Let us compare x_A^1 and x_A^0. Note first that $x_A^1 > 0$. Hence, if $x_A^0 = 0$, then $x_A^0 < x_A^1$. If $x_A^0 > 0$, then

$$\begin{aligned} x_A^1 - x_A^0 &= \sqrt{\frac{Z_A^2}{(1-\rho)^2} + 2\lambda_1} + \sqrt{\frac{Z_A^2}{(1-\rho)^2} + (1-2\lambda_0)} - 1 \\ &= \sqrt{\frac{Z_A^2}{(1-\rho)^2} + 2\int_M x dx} + \sqrt{\frac{Z_A^2}{(1-\rho)^2} + 1 - 2\int_M (1-x) dx} - 1. \end{aligned}$$

The integrand in the first square root is smaller than that in the second one if and only if $x < \frac{1}{2}$. Hence, the above expression is minimized by $M = \left[0, \frac{1}{2}\right]$, in which case $2\int_M x dx = 1 - 2\int_M (1-x)dx = \frac{1}{4}$, which implies that $x_A^1 - x_A^0 \geq 2\sqrt{\frac{Z_A^2}{(1-\rho)^2} + \frac{1}{4}} - 1 > 0$. However, this is a contradiction to case (ii) of Proposition 2, which proves that equilibrium in this example must be characterized by case (i) of that Proposition.

Now that it is known that equilibrium is characterized by case (i) of Proposition 2, we can use $M = [0, x_T]$ to substitute for $2\lambda_1 = x_T^2$ and $2\lambda_0 = 1 - (1-x_T)^2$ in (32) and (33) and thereby obtain (24) and (25). ■

B Proof of Proposition 6

(26) and (27) are obtained by using (24) and (25) to substitute for x_A^0, x_A^1, $\frac{dx_A^0}{dx_T}$ and $\frac{dx_A^0}{dx_T}$ in (22): If $x_A^0 > 0$, the expression in curly brackets in (22) becomes

$$\begin{aligned} & \rho + (1-\rho)(x_A^1 - x_A^0) + (1-\rho)\left(\frac{x_T^2}{\sqrt{\frac{Z_A^2}{(1-\rho)^2} + x_T^2}} + \frac{(1-x_T)^2}{\sqrt{\frac{Z_A^2}{(1-\rho)^2} + (1-x_T)^2}}\right) \\ = \; & \rho + (1-\rho)\left(-1 + \frac{\frac{Z_A^2}{(1-\rho)^2} + 2x_T^2}{\sqrt{\frac{Z_A^2}{(1-\rho)^2} + x_T^2}} + \frac{\frac{Z_A^2}{(1-\rho)^2} + 2(1-x_T)^2}{\sqrt{\frac{Z_A^2}{(1-\rho)^2} + (1-x_T)^2}}\right) \end{aligned}$$

which is equal to the expression in curly brackets in (26). Note furthermore from (25) that $x_A^0 > 0$ if and only if $(1-x_T)^2 < 1 - \frac{2Z_A}{1-\rho}$.

If $(1-x_T)^2 \geq 1 - \frac{2Z_A}{1-\rho}$, then $x_A^0 \equiv 0$ and, therefore, $\frac{dx_A^0}{dx_T} = 0$. Hence, the expression in curly brackets in (22) becomes

$$\begin{aligned} & \rho + (1-\rho)x_A^1 + (1-\rho)\frac{x_T^2}{\sqrt{\frac{Z_A^2}{(1-\rho)^2} + x_T^2}} \\ = \quad & \rho + (1-\rho)\left(-\frac{Z_A}{1-\rho} + \frac{\frac{Z_A^2}{(1-\rho)^2} + 2x_T^2}{\sqrt{\frac{Z_A^2}{(1-\rho)^2} + x_T^2}}\right) \end{aligned}$$

which is equal to the expression in curly brackets in (27).

It remains to show that the optimum indeed satisfies the first-order condition. Note first that the conditions listed in the first paragraph of Proposition 5 are satisfied: As shown in Lemma 6, equilibrium is governed by case (i) of Proposition 2. Furthermore, $y(x)f(x) = x$ exhibits an increasing hazard rate and both $z_r(\cdot)$ are differentiable and weakly concave on $(0, 1)$. According to the proof of Proposition 5, under these assumptions the equilibrium pretrial choice x_T^* satisfies the first-order condition (26) or (27) if and only if the first derivative of the objective function at $x_T = 0$ is positive.

To check whether this is the case, note that $1 > 1 - \frac{2Z_A}{1-\rho}$, which implies that $\lim_{x_T \searrow 0} \frac{dx_A^0}{dx_T} = 0$, so that the first derivative of the plaintiff's objective function for x_T sufficiently close to zero is

$$\Pi_T'(x_T) = \left\{\rho + (1-\rho)\left(-\frac{Z_A}{1-\rho} + \sqrt{4x_T^2 + \frac{\frac{Z_A^4}{(1-\rho)^4}}{\frac{Z_A^2}{(1-\rho)^2} + x_T^2}}\right)\right\}(1-x_T)D - (c_T^p + c_T^d).$$

Hence, $\lim_{x_T \searrow 0} \Pi_T'(0) = \rho D - (c_T^p + c_T^d)$, which is positive if and only the condition in the proposition holds. ■

C Proof of Proposition 7

Proposition 6 implies that, under the conditions of this proposition, x_T^* satisfies (26) in any perfect Bayesian equilibrium. There may be multiple local maxima which satisfy (26), but as the objective function is continuous and differentiable in the set of parameters given by the conditions of this proposition, it is always possible to find equilibria that are continuous in a given, sufficiently small interval of the parameter space. Hence, it is always possible to find an equilibrium such that the comparative statics derived in this proposition hold. Let us therefore restrict to the impact of marginal changes in parameters on continuous changes in the solution to the first-order condition (26).

If $c_T^d + c_T^p$ increases, then the right-hand side of (26) increases. As the second-order condition for an interior maximum is satisfied by assumption, the equilibrium x_T^* will fall as a result.

Again due to the second-order condition, the sign of the effect of an increase in $c_A^d + c_A^p$ on x_T^* is equal to the sign of the partial derivative of the left-hand side of (26) w.r.t. Z_A, which in turn is equal to the sign of the partial derivative of any function $h(x) = \frac{x^2}{x+A}$ with $A > 0$. As $h'(x) = \frac{x(x+2A)}{(x+A)^2} > 0$, we can conclude that x_T^* is increasing in $c_A^d + c_A^p$. ■

References

Bebchuk, L. (1984): "Litigation and Settlement under Imperfect Information," *RAND Journal of Economics*, 15(3), 404–415.

Bütler, M., and H. Hauser (2000): "The WTO Dispute Settlement System: A First Assessment from an Economic Perspective," *Journal of Law, Economics, and Organization*, 16(2), 503–533.

Daughety, A., and J. Reinganum (2000): "Appealing Judgments," *RAND Journal of Economics*, 31(3), 502–525.

Gerardi, D., and L. Yariv (2008): "Information Acquisition in Committees," *Games and Economic Behavior*, 62(2), 436–459.

Gul, F., H. Sonnenschein, and R. Wilson (1986): "Foundations of Dynamic Monopoly and the Coase Conjecture," *Journal of Economic Theory*, 39(1), 155–190.

Hart, O. (1989): "Bargaining and Strikes," *Quarterly Journal of Economics*, 104(1), 25–43.

Hylton, K. (2002): "An Asymmetric-Information Model of Litigation," *International Review of Law and Economics*, 22(2), 153–175.

Iossa, E., and G. Palumbo (2007): "Information Provision and Monitoring of the Decision-Maker in the Presence of an Appeal Process," *Journal of Institutional and Theoretical Economics JITE*, 163(4), 657–682.

Landeo, C., M. Nikitin, and S. Baker (2007): "Deterrence, Lawsuits, and Litigation Outcomes Under Court Errors," *Journal of Law, Economics, & Organization*, 23(1), 57.

Levy, G. (2005): "Careerist Judges and the Appeals Process," *RAND Journal of Economics*, 36(2), 275–297.

Nalebuff, B. (1987): "Credible Pretrial Negotiation," *RAND Journal of Economics*, 18(2), 198–210.

Ottaviani, M., and P. Sørensen (2001): "Information Aggregation in Debate: Who Should Speak First?," *Journal of Public Economics*, 81(3), 393–421.

P'ng, I. (1987): "Litigation, Liability, and Incentives for Care," *Journal of Public Economics*, 34(1), 61–85.

Robson, A., and S. Skaperdas (2008): "Costly Enforcement of Property Rights and the Coase Theorem," *Economic Theory*, 36(1), 109–128.

Shavell, S. (1995): "The Appeals Process as a Means of Error Correction," *Journal of Legal Studies*, 24(2), 379–426.

——— (1999): "The Level of Litigation: Private Versus Social Optimality of Suit and of Settlement," *International Review of Law and Economics*, 19(1), 99–116.

——— (2006): "The Appeals Process and Adjudicator Incentives," *Journal of Legal Studies*, 35, 1–475.

——— (2007): "Optimal Discretion in the Application of Rules," *American Law and Economics Review*, 9(1), 175–194.

——— (2010): "On the Design of the Appeals Process: The Optimal Use of Discretionary Review Versus Direct Appeal," *Journal of Legal Studies*, 39(1), 63–108.

Spier, K. (1992): "The Dynamics of Pretrial Negotiation," *Review of Economic Studies*, 59(1), 93–108.

——— (1994): "Settlement Bargaining and the Design of Damage Awards," *Journal of Law, Economics, & Organization*, 10(1), 84–95.

Spitzer, M., and E. Talley (2000): "Judicial Auditing," *The Journal of Legal Studies*, 29(2), 649–683.